Temperature Robust Longwave Infrared Hyperspectral Change Detection

Nicholas A. Durkee
Wright State University

Follow this and additional works at: https://corescholar.libraries.wright.edu/etd_all
Part of the Electrical and Computer Engineering Commons

Repository Citation
https://corescholar.libraries.wright.edu/etd_all/2224

This Thesis is brought to you for free and open access by the Theses and Dissertations at CORE Scholar. It has been accepted for inclusion in Browse all Theses and Dissertations by an authorized administrator of CORE Scholar. For more information, please contact library-corescholar@wright.edu.
Temperature Robust Longwave Infrared Hyperspectral Change Detection

A Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Electrical Engineering

by

NICHOLAS A. DURKEE
B.S.E.E., Wright State University, 2017

2018
Wright State University

Joshua Ash, Ph.D.
Thesis Director

Fred Garber, Ph.D.
Interim Chair, Department of Electrical Engineering

Committee on Final Examination

Joshua Ash, Ph.D.

Fred Garber, Ph.D.

Arnab Shaw, Ph.D.

Barry Milligan, Ph.D.
Interim Dean of the Graduate School
ABSTRACT


In this thesis, we develop and evaluate change detection algorithms for longwave infrared (LWIR) hyperspectral imagery. Because measured radiance in the LWIR domain depends on unknown surface temperature, care must be taken to prevent false alarms resulting from in-scene temperature differences that appear as material changes. We consider four strategies to mitigate this effect. In the first, pre-processing via traditional temperature-emissivity separation yields approximately temperature-invariant emissivity vectors for use in change detection. In the second, we utilize alpha residuals to obtain robustness to temperature errors. Next, we adopt a minimax approach that minimizes the maximal spectral deviation between measurements. Finally, we reduce our minmax approach to solve with fewer variables. Examples using synthetic and measured data quantify the computational complexity of the proposed methods and demonstrate orders of magnitude reduction in false alarm rates relative to existing methods.
List of Symbols

Chapter 2

$L$  Radiance term represented in $W m^{-1} cm^{-2} sr^{-1}$
$L_\downarrow$  Downwelling radiance term, reflected radiance coming down from atmosphere, reflected by the ground into the sensor
$L_\uparrow$  Upwelling radiance, radiance originating within the atmosphere directed into the sensor
$\tau$  Path loss of radiance traveling through the atmosphere
$\lambda_k$  Wavelength in $\mu m$, $k$ indicates the index of wavelength
$\epsilon$  Emissivity of material
$T$  Temperature of material
$q_{BB}(T, \lambda)$  Planck’s Function (Blackbody Function)
$c$  speed of light = 299702547 m/s
$h$  Planck’s constant = $6.625 \times 10^{-34} m^2 kg/s$
$k_b$  Boltzmann’s constant = $1.38064852 \times 10^{-23} m^2 kg s^{-1} K^{-1}$
$\sigma^2$  Variance of data
$\sigma^2_\epsilon$  Variance of emissivity, indexed by $\lambda_k$
$\mathcal{F}(\mathbf{x})$  Convolutional operator on vector $\mathbf{x}$ with boxcar filter
$f_1/f_2/f_{comp}$  Additional error terms in smoothness-TES
$\frac{\partial f}{\partial T}$  Partial derivative of undetermined function $f$ with respect to temperature
$\hat{f}$  Hat refers to value approximated from radiance data
$N_c$  Number of channels
$\alpha_\epsilon$  Ideal alpha residual computed from emissivity
$\alpha$  Alpha Residual
$L_{gl}$  Ground leaving radiance with path and part of downwelling radiance subtracted
$c_1$  First constant in Planck’s function = $2hc^2$
$c_2$  Second constant in Planck’s function = $\frac{hc}{k_b}$
$a$  Full exponential term in Planck’s function = $\frac{c_2}{X_T}$
$b$  Full constant multiplier in Planck’s function = $\frac{c_1}{X_T}$
$T_a$  Custom function for use in alpha residuals
$t$  Temperature modifier = $\frac{T}{T_0}$
$\alpha_\hat{\epsilon}$  Alpha residual calculated from radiance using Diani’s method prior to temperature elimination
$\alpha_M$  Alpha residuals post temperature elimination
$\mu$  Sample mean
$(k)$  Superscript indicating time of data
$\Sigma_{jj}$  Sample covariance matrix of datacube
$\Sigma_{jk}$  Sample cross-covariance matrix between datacubes
$\| \|_2^2$  $\ell_2$ norm or Euclidean distance between two vectors

Chapter 3
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0$</td>
<td>Null hypothesis</td>
</tr>
<tr>
<td>$H_1$</td>
<td>Alternative hypothesis</td>
</tr>
<tr>
<td>$P$</td>
<td>Projection matrix for use in orthogonal projection</td>
</tr>
<tr>
<td>$\epsilon_M$</td>
<td>Orthogonally projected emissivity</td>
</tr>
<tr>
<td>$\Sigma$</td>
<td>Covariance matrix of data</td>
</tr>
<tr>
<td>$s$</td>
<td>Change detection error</td>
</tr>
<tr>
<td>$R$</td>
<td>Jacobian of JTO</td>
</tr>
<tr>
<td>$r_0$</td>
<td>Error vector for JTO</td>
</tr>
<tr>
<td>$q^{-1}$</td>
<td>Inverse blackbody function</td>
</tr>
</tbody>
</table>
## Contents

1 Introduction 1
   1.1 Motivation ........................................... 1
   1.2 Objective ........................................... 4
   1.3 Organization and Contributions ......................... 4

2 Background 6
   2.1 LWIR Physical Model ................................. 6
   2.2 Temperature Emissivity Separation ...................... 9
   2.3 Alpha Residuals ..................................... 13
   2.4 Image Registration ................................... 16
   2.5 Existing Anomalous Based Methods ....................... 22

3 LWIR Change Detection Algorithms 24
   3.1 Smoothness-TES ....................................... 26
      3.1.1 Mean Subtraction ................................ 27
   3.2 Alpha Residuals ..................................... 29
      3.2.1 Covariance Matrix for Alpha Residuals ............ 31
      3.2.2 Downside of Alpha Residuals ..................... 36
   3.3 Joint Temperature Optimization (JTO) .................. 38
      3.3.1 JTO-Linear .................................... 39
      3.3.2 JTO-Newton ................................... 41
   3.4 Single Temperature Dependence ......................... 43
      3.4.1 Single Temperature Variance Optimization ......... 45
      3.4.2 Cross Radiance Optimization ...................... 46
      3.4.3 Closed Form Using Known Temperatures ............ 47
   3.5 LWIR Data Challenges .................................. 48
      3.5.1 Smoothness-TES Dependence on Atmosphere .......... 48
      3.5.2 Invalid Denominators ............................ 49
      3.5.3 Runaway Temperatures ............................ 50
4 Results
4.1 Synthetic Data 
4.1.1 Experiment Setup 
4.1.2 Anomalous Detection Methods 
4.1.3 Smoothness-TES 
4.1.4 Alpha Residuals 
4.1.5 Joint Temperature Optimization 
4.1.6 Single Temperature Dependence 
4.1.7 All Methods 
4.1.8 Noise Scales 
4.2 Measured Data 
4.3 Semi-Synthetic Data 
4.4 Timing Results

5 Conclusions
5.1 Summary 
5.2 Future Work

Bibliography
# List of Figures

1.1 Example hyperspectral representation of outdoor scene .......................... 2  
1.2 Significant and non-significant changes ............................................. 3  
1.3 Comparison of two images with nuisance parameters ............................ 4  

2.1 Effects on measured radiance .......................................................... 8  
2.2 Smoothness error vs. temperature .................................................... 10  
2.3 Types of affine distortion in imagery ................................................ 18  
2.4 Example of non-affine deformation ................................................... 19  
2.5 Manual selection of points .............................................................. 20  
2.6 Correlation of points via SURF ....................................................... 20  

3.1 Comparison of two measurements with only atmosphere and temperature  
Changes ......................................................................................... 25  
3.2 View of two recovered emissivity vectors compared to truth ............... 28  
3.3 Alpha residual as a function of emissivity ........................................... 37  
3.4 Example emissivity estimates produced for joint temperature optimization  
3.5 Atmosphere rigidness vs. smoothness-TES Performance ..................... 39  
3.6 Critical temperature locations vs. downwelling radiance ................. 49  
3.7 Critical temperature locations in error functions .............................. 50  
3.8 Error function of original smoothness-TES ...................................... 51  

4.1 ROC curve results for anomalous change detection methods ............... 56  
4.2 Results for smoothness-TES change detection with orthogonal projection  
improvements .................................................................................. 57  
4.3 Results for smoothness-TES change detection with standardized Euclidean  
error metric .................................................................................... 58  
4.4 Results for smoothness-TES change detection with mean subtraction element  
4.5 Results for smoothness-TES with all variants .................................... 60  
4.6 Results for alpha residuals all methods ............................................. 61  
4.7 Results for joint temperature optimization all methods ..................... 62  
4.8 Results for single temperature optimization all methods ................... 63  
4.9 All methods on synthetic data ......................................................... 64  
4.10 Noise stress testing at $P_{fa} = 10^{-1}$ .............................................. 65
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.11</td>
<td>Noise stress testing at $P_{fa} = 10^{-2}$</td>
<td>66</td>
</tr>
<tr>
<td>4.12</td>
<td>Noise stress testing at $P_{fa} = 10^{-3}$</td>
<td>66</td>
</tr>
<tr>
<td>4.13</td>
<td>Raw radiance data for each time</td>
<td>69</td>
</tr>
<tr>
<td>4.14</td>
<td>Detection statistic for all methods on real data</td>
<td>70</td>
</tr>
<tr>
<td>4.15</td>
<td>ROC results for all methods on real data</td>
<td>71</td>
</tr>
<tr>
<td>4.16</td>
<td>Raw radiance data for each time</td>
<td>73</td>
</tr>
<tr>
<td>4.17</td>
<td>Detection statistic for all methods on semisynthetic data</td>
<td>74</td>
</tr>
<tr>
<td>4.18</td>
<td>ROC results for all methods on semisynthetic data</td>
<td>75</td>
</tr>
</tbody>
</table>
List of Tables

2.1 Error terms used in smoothness-TES ........................................ 12
4.1 Runtime for all methods per 100,000 pixels in seconds ............. 76
Acknowledgment

I would like to thank Dr. Ash for helping(hauling) me through this academic process. You have taught me how to be a better thinker, writer, reader, debater, and researcher than I ever would have thought possible. Additionally, I would like to thank Dr. Meola and AFRL for sponsoring this thesis and providing data and tools necessary for this project.

I would also like to thank Dr. Garber for an incredible amount of encouragement throughout this process. Secondly, table 2 at the ATR center for sparking the real speedup of this process and providing the camaraderie to make this project enjoyable.

Last but not least, I would like to thank my mom and dad for providing the motivation to make it through this process. Without you, I would never have made it and never have the motivation to finish writing.
Dedicated to

Rocky, for without you, I don’t know where I’d be
Introduction

1.1 Motivation

Consider a normal picture taken from a color camera. This picture contains radiance measurements from the 3 primary colors of red, green, and blue (RGB). These colors correspond to certain wavelengths on the electromagnetic spectrum of 700, 550, and 475 nanometers respectively. Hyperspectral imagery (HSI) captures information from multiple locations on the electromagnetic spectrum. Rather than sampling at just 3 wavelengths, a hyperspectral camera could capture images at 440, 450, 460 nm all the way up to 12.5 micrometers and thus obtain hundreds if not thousands of individual single-band images. While many of these wavelengths may not be not visible to the human eye, they still contain information about the underlying objects. Together, these hundreds/thousands of images can contain a very large amount of information. Within this information, a single pixel contains a vector of radiance values covering a large spectra as shown in Figure 1.1. Due to this wealth of spectral information, various image processing operations may be done on a single pixel rather than relying on groups of pixels. Other modalities such as RADAR or electro-optic (EO) perform change detection via blocks of pixels which can diminish the accuracy of a change detection map [18]. For this reason, hyperspectral imagery (HSI) is of great interests to many researchers.

In this work, the goal is to detect changes between two long-wave infrared (LWIR)
HSI images taken at different times. Example changes include detecting diseases in crops and detecting enemy combatants in denied areas. These changes are considered “significant” changes or changes that are pertinent to a human operator. Significant changes comprise of inserted or deleted objects. For example, this could be due to a vehicle entering or exiting a scene, as well as significant material changes due to explosions or other human activity. For example, in Figure 1.2, the time 1 or reference image represents an overhead view of Wright State. The true significant change is represented by the blue circles in which a building has been torn down and replaced with a field. Meanwhile, on the right side, nothing has actually changed in the image. However, the right side of the second image looks significantly different from the reference image. These changes are considered non-significant or false changes.

These false changes, or nuisance changes, corrupt HSI and make simple comparison between scenes to find significant changes challenging. These nuisance changes consist of two main elements in the longwave range: atmospheric and temperature effects. The atmospheric effects add and block radiance from reaching the camera whereas the temperature is an additive effect dependent on the temperature of the underlying object. Together, these nuisance parameters obscure the underlying elements necessary to find true changes within
Figure 1.2: Significant and non-significant changes

Consider Figure 1.3, in which two images are compared directly against each other. One of the images has been modified with moderate atmospheric and temperature effects. If the images are compared directly against each other using either an $\ell_2$ norm or structural similarity metric [23], the change detection algorithm will output a large error for much of the image even if there is no true change between the two images. Therefore, in order to perform change detection properly, various methods must be utilized to eliminate these nuisance effects.
1.2 Objective

Previous work in HSI change detection predominantly focused on visible-shortwave infrared (VIS-SWIR) data. In this spectral range, the primary challenges exist due to unknown changes in atmosphere and shadows. The current state-of-the-art uses linear transformations to account for atmospheric and illumination changes. In LWIR imagery, the spatially-varying temperature must be compensated for in addition to atmospheric changes. Our primary objective in this work is to develop new methods which compensate for these temperature differences and compare them to existing state-of-the-art methods utilized in VIS-SWIR data.

1.3 Organization and Contributions

In Chapter 2, we explore existing models and literature relating to LWIR HSI. Our primary contributions are presented in Chapter 3, which includes the development of multiple new change detection algorithms for LWIR imagery. The first two algorithms obtain robustness to temperature by utilizing existing TES and alpha residual methodologies. The third and fourth algorithms are based on minimax strategies which attempt to minimize the maxi-
mal spectral deviation between measurements. In Chapter 4, we use both synthetic and measured LWIR data to quantify the performance of existing algorithms and our proposed algorithms from Chapter 3. All methods are evaluated in terms of detection performance and computational complexity. Finally, in Chapter 5, we summarize the thesis and propose potential future research.
Background

This chapter discusses existing work utilized for LWIR signal exploitation. This comprises of the model and temperature emissivity separation processes as well as image registration preprocessing. Finally, this chapter will cover the most popular method of change detection utilized in the visible and SWIR domains.

2.1 LWIR Physical Model

In order to approach the problem of LWIR hyperspectral change detection, one must understand the underlying physical model for the data. The generalized forward model describes the physics of radiative transfer in the LWIR spectral region [9]

\[
L(\lambda_k) = q_{BB}(T, \lambda_k)\epsilon(\lambda_k)\tau(\lambda_k) + (1 - \epsilon(\lambda_k))L_{\downarrow}(\lambda_k)\tau(\lambda_k) + L_{\uparrow}(\lambda_k) + n_k, \quad (2.1)
\]

where \(\lambda_k\) indicates the measurement at wavelength number \(k = 1, 2 \ldots N_c\) and the black-body radiance is computed as

\[
q_{BB}(T, \lambda_k) = \frac{2hc^2}{\lambda_k^5 \left( e^{\frac{hc}{\lambda_k T}} - 1 \right)}. \quad (2.2)
\]
In this equation \( c \) is the speed of light \( 3 \times 10^8 \, ms^{-1} \), \( h \) is Planck’s constant \( 6.625 \times 10^{-34} \, m^2 kgs^{-1} \), \( k_b \) is Boltzmann’s constant \( 1.380 \times 10^{-23} m^2 kgs^{-2} K^{-1} \), and \( T \) is the ground surface temperature measured in Kelvins (K).

Emissivity, \( \epsilon \), is a measure of how much energy a material emits relative to a perfect blackbody. The radiance from the surface of a high emissivity material is composed of mainly surface emission. Likewise, the radiance from a low emissivity material is composed of mainly reflected radiance from other sources.

Path transmission loss, \( \tau \), describes signal loss due to atmospheric absorption between the ground and the sensor. This effect varies as a function of wavelength based on the absorption properties of water vapor, ozone, and other atmospheric gases. Gases will have critical wavelengths for which strong absorption occurs.

Downwelling radiance, \( L_{\downarrow} \), is the radiance originating from other sources within the scene such as sky radiance or emission from nearby background. As stated previously, if a material has a high reflectance, this component will come into play and the sensor will receive a high amount of reflected downwelling radiance.

Upwelling radiance (path radiance), \( L_{\uparrow} \), represents radiance emitted directly from the sky into the sensor’s field-of-view. Together, transmission loss, upwelling radiance, and downwelling radiance comprise the atmospheric conditions or TUD (transmission, upwelling, downwelling) of the scene. In this work, the atmospheric conditions are assumed to be constant throughout the entire scene.

The noise, \( n_k \), represents the additive noise within the scene. In this work, the noise is assumed to be additive identically distributed zero mean Gaussian noise \( \mathcal{N}(0, \sigma^2) \) in which the variance is constant throughout the entire spectra. Together, all of these effects can be represented as shown in Figure 2.1.

For the purposes of autonomous intelligence, material changes, as opposed to atmospheric or temperature changes, are generally more important. Therefore, this work focuses on detecting changes in material emissivity as these are indicative of object changes. In or-
Figure 2.1: Effects on measured radiance
order to retrieve the emissivity, Equation (2.1) can be rewritten as

\[
\epsilon(\lambda_k, T) = \frac{L(\lambda_k) - L_u(\lambda_k)\tau(\lambda_k) - L_T(\lambda_k)}{q_{BB}(T, \lambda_k) - L_u(\lambda_k)}\tau(\lambda_k).
\] (2.3)

The process of retrieving emissivity and temperature from the radiance is referred to as temperature emissivity separation (TES). As long as all of the parameters described in (2.3) are known, the emissivity can be recovered perfectly if there is no noise in the radiance dimension. Various atmospheric compensation (AC) methods exist to estimate the unknown TUD parameters. Examples of AC include Fast Line-of-sight Atmospheric Analysis of Hypercubes Infrared (FLAASH-IR) [5], In Scene Atmospheric Correction (ISAC) [14], and the empirical line method. However, this leaves an unknown temperature element which can vary throughout an entire scene. In this thesis, it is assumed that TUD estimates are available, this leaves the spatially varying temperature as the only unknown element.

## 2.2 Temperature Emissivity Separation

The problem with estimating unknown temperatures is that there are too many unknown variables for the given measurement. Equation (2.3) describes \(N_c\) unknown emissivities as well as an unknown \(T\). Meanwhile, there are only \(N_c\) measurements contained within \(L\). This means that any system of recovering emissivity must recover \((N_c + 1)\) values from \(N_c\) measurements which makes the entire system under-determined. Therefore, barring physically measuring the ground temperature of every single pixel, some assumptions need to be made concerning the physical model in order to extract the temperature and emissivity.

Smoothness-TES [4] looks at the attributes of the expected recovered emissivity and makes assumptions about the emissivity properties in order to reduce the dimensionality. The model states that the recovered emissivity will contain residual atmospheric absorption features at all temperatures besides the true temperature. These atmospheric absorp-
tion features are generally very rigid meaning they possess very sharp spectral features. Meanwhile, the emissivity of solid materials varies much more slowly as a function of wavelength. Therefore, a way to gauge the accuracy of a temperature estimate is to measure the rigidness of the recovered emissivity. A high-pass filter measures this rigidness. Then, a $\ell_2$ norm quantifies the residual of the high-pass filter to calculate the total rigidness of the emissivity. In practice, a low-pass filtered recovered emissivity is subtracted from the unfiltered recovered emissivity to emulate a high-pass filter. The suggested low-pass-filter is a 3-tap boxcar filter [4] indicated by $\mathcal{F}$. An iterative process scans through a temperature range in order to find the optimal temperature [4]

$$\hat{T} = \min_T ||\hat{\epsilon}(T) - \mathcal{F}(\hat{\epsilon}(T))||_2.$$ \hspace{1cm} (2.4)

Using simulated data, this approach towards the smoothness minima can be seen in Figure 2.2

---

This method encounters problems due to the lack of temperature bounds. Additionally, this method does not correctly account for the noise covariance matrix. This noise can be
obtained reversing Equation (2.1)

\[ \epsilon(\lambda_k) = \left( \frac{L(\lambda_k) - L_\downarrow(\lambda_k)\tau(\lambda_k) - L_\uparrow(\lambda_k) + n_k}{q_{BB}(T, \lambda_k) - L_\downarrow(\lambda_k))\tau(\lambda_k)} \right). \]  \quad (2.5)

From there, in order to obtain the effects of noise, eliminate constant values. Finally, when multiplying a Gaussian random variable by a constant value \( a \), the variance of the new variable is \( a^2\sigma^2 \). Therefore, the noise of the emissivity is

\[ \sigma_\epsilon(\lambda_k) = \left( \frac{\sigma}{q_{BB}(T, \lambda_k) - L_\downarrow(\lambda_k))\tau(\lambda_k)} \right)^2. \]  \quad (2.6)

As stated previously, the noise is assumed to be constant across all wavelengths in the radiance space. Therefore, the emissivity noise is widely varying due to the rigid nature of the atmospheric conditions as shown in Equation (2.6). However, Equation (2.4) treats all bands equally as if the noise remained the same. To adjust for this, Borell released a new paper which re-posed the problem in the radiance domain [3]

\[ T = \min_T ||L - L(\mathcal{F}(\hat{\epsilon}(T)))||^2_T. \]  \quad (2.7)

In this method, \( L(\mathcal{F}(\hat{\epsilon}(T))) \) represents a reconstructed radiance utilizing a smoothed recovered emissivity

\[ L(\mathcal{F}(\hat{\epsilon}(T, \lambda_k))) = q_{BB}(T, \lambda_k)\mathcal{F}(\hat{\epsilon}(\lambda_k, T))\tau(\lambda_k) + (1 - \mathcal{F}(\hat{\epsilon}(\lambda_k, T)))L_\downarrow(\lambda_k)\tau(\lambda_k) + L_\uparrow(\lambda_k). \]  \quad (2.8)

Additionally, the new method added multiple additional error terms in order to regulate the results as shown in Table 2.1. These error terms prevent unrealistic conditions and serve to improve the overall performance of smoothness-TES.
\[ f_1(T, \hat{\epsilon}) = \sqrt{\sum_{k=1}^{N_e} (L(\lambda_k) - L(\mathcal{F}(\hat{\epsilon}(T, \lambda_k))))^2} \]

\[ f_2(T, \hat{\epsilon}) = \frac{1}{N_e} \sum_{k=1}^{N_e} \text{sgn}(\hat{\epsilon} < 0) + \text{sgn}(\hat{\epsilon} > 1) \]

Table 2.1: Error terms used in smoothness-TES

Finally, a minimization function solves for temperatures

\[ \hat{T} = \min_T ||L - L(\mathcal{F}(\hat{\epsilon}(T))))||_2^2 + f_2(T). \]  

This method allows for TES with relatively fast with reasonably high accuracy. Testing with synthetic data reveals that this method is capable of extracting temperatures with a RMS error of \( \sim 0.8K \). However, it is important to note that smoothness-TES does not provide a guarantee of an exact solution. There are still more unknown variables than measured radiance values.
2.3 Alpha Residuals

However, the process of estimating a temperature via an optimization process for every single pixel within a scene can be extraordinarily expensive, therefore, an alternative is desired. Due to the mathematical properties of Planck’s function, the temperature and emissivity are inseparable. However, usage of Wein’s approximation removes the offending $-1$ portion of the denominator of Equation (2.2) resulting in minimal approximation error for the LWIR and associated relevant temperature range. It then becomes possible to rewrite the emissivity equation in an alternate form independent of temperature. This reformulation is referred to as alpha residuals or alpha emissivity.

Alpha residuals has undergone several stages of development. The general history involves 4 main stages:

1. Approximating Planck’s function as Wein’s approximation so that a new emissivity realization can be generated without knowledge of temperature [15].

2. Using an estimate of temperature, it becomes possible to avoid the Wein’s approximation. However, this method still ignores the downwelling radiance component [11].

3. A new method considers the down-welling radiance component and adds it to the alpha residual equation [7].

4. Diani’s work also considers temperature estimation error. The paper proposes that the temperature error can be eliminated by decorrelating temperature artifacts with the alpha residuals [7].

Alpha emissivity is related to the true emissivity via the following equation

$$\alpha_e(\lambda_k) = \lambda_k \log(\epsilon(\lambda_k)) - \frac{1}{N_c} \sum_{i=1}^{N_c} \lambda_i \log(\epsilon(\lambda_i)).$$

(2.10)
The goal is to represent the received radiance in this form such that the library alpha emissivity and recovered alpha emissivity can be compared in a one-to-one comparison. To start, the forward model can be re-represented in a way that removes elements not pertaining to the emissivity or temperature

\[
\frac{L(\lambda_k) - L_\uparrow(\lambda_k)}{\tau(\lambda_k)} = q_{BB}(T, \lambda_k)e(\lambda_k) + (1 - e(\lambda_k))L_\downarrow(\lambda_k).
\] (2.11)

From there, the lefthand side of the equation (2.11) is redefined as \(L_{gl}\). Additionally, the downwelling radiance effect is assumed to be reasonably low because emissivity is generally high. Therefore, as \(e \to 1\), \((1 - e) L_\downarrow \to 0\). This allows Equation (2.11) to be rewritten as

\[
L_{gl}(\lambda_k) = q_{BB}(T, \lambda_k)e(\lambda_k) = e(\lambda_k) \frac{a(\lambda_k)}{e^{b(\lambda_k)T_0} - 1},
\] (2.12)

where \(a(\lambda_k) = \frac{2hc}{\lambda_k^5}\), \(b(\lambda_k) = \frac{hc}{k_BT_0\lambda_k}\) and \(T_0\) is a central temperature estimate for the scene.

In order to solve (2.10) \(\to\) (2.12), a new system must be defined

\[
T_a[x(\lambda_k)] = \frac{2hc^2}{b(\lambda_k)} \log(x(\lambda_k)) - \frac{2hc^2}{N_c} \sum_{i=1}^{N_c} \frac{\log(x(\lambda_i))}{b(\lambda_i)}.
\] (2.13)

By neglecting the \(-1\) in the denominator of Planck’s function, this system becomes solvable in closed form via the following system \([15]\)

\[
\alpha_{W_eins}(\lambda_k) = T_a[L_{gl}(\lambda_k)] - T_a[a(\lambda_k)].
\] (2.14)

However, this method can have large inaccuracies particularly when \(|T - T_0|\) becomes very large. Therefore, a new method was proposed which included a temperature estimate \(t = \frac{T}{T_0}\) \([11]\)

\[
\alpha_{T_{approx}}(\lambda_k) = T_a[L_{gl}(\lambda_k)] - T_a[a(\lambda_k)] + T_a[e^{b(\lambda_k)t} - 1].
\] (2.15)
For low emissivity cases, this method becomes invalid as the downwelling radiance element cannot be ignored. Therefore, a new method was proposed to include the downwelling radiance in the calculations [7]

\[ \alpha(\lambda_k) = T_a[L_{gl}(\lambda_k) - L_{\downarrow}(\lambda_k)] - T_a[a(\lambda_k)] + T_a[e^{b(\lambda_k)t} - 1] - T_a \left[ 1 - \frac{L_{\downarrow}(\lambda_k)(e^{b(\lambda_k)t} - 1)}{a(\lambda_k)} \right]. \]

(2.16)

If \( |T_{\text{estimated}} - T_{\text{true}}| \) is very small, then the approximation in (2.16) is reasonably accurate. However, obtaining accurate initial temperature estimates is challenging. Therefore, further steps were taken [7] in order to annihilate the temperature effects by decorrelating the temperature effects from the alpha residuals. This can be done by computing the projection of the derivative of the recovered alpha emissivity with respect to \( t \) onto the recovered alpha emissivity. The derivative with respect to temperature can be defined as

\[ \frac{\partial \alpha(\lambda_k)}{\partial t} = \alpha_{dt}(\lambda_k) = T_a \left[ \exp \left( \frac{b(\lambda_k)}{e^{b(\lambda_k)t - 1}} + \frac{b(\lambda_k)L_{\downarrow}(\lambda_k)e^{b(\lambda_k)t}}{a(\lambda_k)} - L_{\downarrow}(\lambda_k)(e^{b(\lambda_k)t} - 1) \right) \right]. \]

(2.17)

This allows the true alpha emissivity to be set up as a first order Taylor Series

\[ \alpha(\lambda_k) = \alpha(\lambda_k) - \alpha_{dt}(\lambda_k) \Delta t. \]

(2.18)

This Taylor Series can then be used to decorrelate \( \alpha_{dt} \) from \( \alpha_\varepsilon \) via orthogonal projection [7]

\[ \alpha_{\delta t} = \alpha_\varepsilon - \frac{\alpha_\varepsilon \cdot \alpha_{dt}}{\alpha_{dt} \cdot \alpha_{dt}} \alpha_{dt}. \]

(2.19)

However, by decorrelating the derivative, other elements are also removed from the resulting alpha emissivity. Therefore, the same operation must be applied to all library elements...
in order to get a one-to-one comparison

\[ \alpha_M = \alpha_e - \frac{\alpha_e \cdot \alpha_{dt}}{\alpha_{dt} \cdot \alpha_{dt}} \cdot \alpha_{dt}. \]  \hspace{1cm} (2.20)

For the purpose of target detection, this approximation will result in relatively large error values for higher valued alpha residuals. The reason for this is because this method has a relative magnitude dependent on the log of emissivity. For emissivities with a low magnitude, the log of the emissivity decreases in value exponentially thus also exponentially magnifying the noise. Therefore, higher valued alpha residuals must be normalized by taking the relative error of the function for the purpose of target detection

\[ s_{RE} = \frac{||\alpha^{(1)}_M - \alpha^{test}_M||^2_2}{||\alpha^{test}_M||^2_2}. \]  \hspace{1cm} (2.21)

### 2.4 Image Registration

In order to compute a change detection statistic, images need to first be co-registered. This allows pixels to be compared in a one-to-one basis. Most current hyperspectral change detection methods rely on pixels compared directly against the same location in the corresponding scene. If this is not properly done, the scenes will be shifted relative to each other. This introduces false changes into the scene and the edges of the scene will show up in a change detection map.

In order to adjust for this, there exists some transform/deformation which will transfer the coordinate system of one scene onto the coordinate system of another scene. This transform can be either affine or non-affine. In affine transforms, the change may be represented
by the affine matrix $A$ [17]

$$A = \begin{bmatrix}
S_x \cos(\theta) & \sin(\theta + H_x) & T_x \\
-S(\theta + H_y) & S_y \cos(\theta) & T_y \\
0 & 0 & 1
\end{bmatrix}. \quad (2.22)$$

In this transform $T$ represents a shift in the x and y dimensions in which the image is translated, $S$ represents a scaling factor which changes the size of an image in the x and y dimensions, $\theta$ represents a rotation in the image, $H$ represents a shear in the image in the x and y dimensions. This affine matrix can then be used to adjust an image. The affine matrix is multiplied by the coordinates of the image, $x$ and $y$, to obtain new coordinates, $x'$ and $y'$. From there, the new image is resampled via an interpolation method such as bicubic or bilinear interpolation. The coordinate transformation can be represented as follows

$$\begin{bmatrix}
x' \\
y' \\
1
\end{bmatrix} = A \begin{bmatrix}
x \\
y \\
1
\end{bmatrix}. \quad (2.23)$$

This matrix contains 4 levels of transformation: translation, rotation, scale, and shear. Figure 2.3 demonstrates these types of affine distortion.
Figure 2.3: Types of affine distortion in imagery
Additionally, images can be distorted in ways where pixels do not retain an affine relationship to each other. These images can be distorted via a distortion field in which pixels are sampled differently throughout the entire image. An example is shown in Figure 2.4.

In Figure 2.4, each pixel has its own drift value and in order to correct the image, the entire deformation matrix must be solved for. Various solvers such as Demons, B-Spline, and Morphon exist for this purpose.

In this thesis, only the affine transformation as defined by Equation (2.22) is considered. The main difficulty within image registration lies in estimating this transformation. One possibility to approach the optimal solution is to find corresponding points between images. This can be done by manually selecting corresponding points via a point selection program or autonomously via algorithms such as SURF [1] [22].
Once the corresponding points have been found, a procrustes analysis [12] can compute the least squares transformation to align the images. The algorithm takes in 4 sets of points $\begin{bmatrix} x^{(1)} & y^{(1)} & x^{(2)} & y^{(2)} \end{bmatrix}$ which represent vectors of points corresponding to locations in each image. The procrustes algorithm then works via the following steps. Note that this algorithm only considers rotation and translation. For the more complex affine transforms, a more sophisticated algorithm must be used.
1. Subtract the mean from each time

\[ \hat{x}^{(t)} = x^{(t)} - \sum_{i=1}^{N_p} x_i^{(t)} \]  
\[ \hat{y}^{(t)} = y^{(t)} - \sum_{i=1}^{N_p} y_i^{(t)} \]  
(2.24)  
(2.25)

2. Calculate the Frobenius norm for each time

\[ f^{(t)} = \sqrt{\sum_{i=1}^{N_p} x_i^{(t)2} + \sum_{i=1}^{N_p} y_i^{(t)2}} \]  
(2.26)

3. Scale the points using the Frobenius norm for the x and y coordinates

\[ \hat{x}_s^{(t)} = \frac{\hat{x}^{(t)}}{f^{(t)}} \]  
(2.27)

\[ \hat{y}_s^{(t)} = \frac{\hat{y}^{(t)}}{f^{(t)}} \]  
(2.28)

4. Concatenate the x and y coordinates for each time into a single matrix

\[ M^{(t)} = \begin{bmatrix} \hat{x}_s^{(t)} & \hat{y}_s^{(t)} \end{bmatrix} \]  
(2.29)

5. Calculate rotation via a Singular Value Decomposition

\[ D = M^{(1)^T} M^{(2)} \]  
(2.30)
Let $USV^T = D$ denote the SVD of $D$.

$$\theta = VU \quad (2.31)$$

6. Using the rotation matrix, the translation can be calculated by comparing the center of points in time 1 vs the rotated points in time 2

$$
\begin{bmatrix}
T_x \\
T_y
\end{bmatrix} = \left[ \frac{1}{N_p} \sum_{i=1}^{N_p} x_i^{(1)} \frac{1}{N_p} \sum_{i=1}^{N_p} y_i^{(1)} \right] - \theta \times \left[ \frac{1}{N_p} \sum_{i=1}^{N_p} x_i^{(2)} \frac{1}{N_p} \sum_{i=1}^{N_p} y_i^{(2)} \right] \quad (2.32)
$$

7. Using these values, the affine matrix can be assembled

### 2.5 Existing Anomalous Based Methods

Previous HSI change detection methods attempt to compensate for atmospheric changes as via normalization methods in visible-SWIR data. These methods postulate that atmospheric effects and illumination differences can be encapsulated via linear transformations. If one dataset is linearly transformed, then the HSI data from both times will be in the same domain and thus can be compared in a meaningful way. If $x^{(t)}$ represents a measured radiance vector at time $t$, and $\hat{x}^{(2)} = f(x^{(1)})$ is a time-2 prediction based on measured time-1 radiance, then $d = x_2 - \hat{x}_2$ represents a residual upon which detection statistics may be formed. Two popular linear transformations, covariance equalization (CE) [19] and chronochrome (CC) [20] are given by

$$CE: \hat{x}^{(2)} = f(x^{(1)}) = \Sigma_{22}^{1/2} \Sigma_{11}^{-1/2}(x^{(1)} - \mu_1) + \mu_2 \quad (2.33)$$

$$CC: \hat{x}^{(2)} = f(x^{(1)}) = \Sigma_{21} \Sigma_{11}^{-1}(x^{(1)} - \mu_1) + \mu_2. \quad (2.34)$$
where $\mu_t$, $\Sigma_{tt}$ represent the mean vector and covariance matrix, respectively, of the measured radiance at time $t$, and $\Sigma_{21}$ is the cross-covariance between $x^{(2)}$ and $x^{(1)}$. These methods predict the hyperspectral datacube utilizing the covariance matrices between the cubes. In order to compute a change detection statistic, the $\ell_2$ norm may be taken between the predicted datacube and the true datacube on a pixel level

$$s_{CC/CE} = \left| \left| \tilde{x}^{(2)} - x^{(2)} \right| \right|_2^2.$$  (2.35)

Implicit in the development of the prediction methods is the assumption that the statistics of the radiance vectors are spatially invariant over the scene, and estimates of the required statistical quantities may be computed as sample statistics of the measured data cubes. This makes these methods well-suited to modeling global effects, such as illumination levels and atmospheric transmission loss, but less well suited to modeling local effects, such as spatially varying shadows [10]. In the LWIR regime, varying surface temperatures give rise to spatially non-stationary statistics and therefore limit the effectiveness of global prediction-based methods.

Previous work on LWIR change detection considered anomalous methods [6]. However, in these works, the images were taken just an hour apart. In this short period of time, the temperatures and atmospheres are unlikely to change significantly. In the present work, considerations are made for cases several days or longer apart. In these scenarios, the scenes may change temperatures in significant non-uniform ways which cannot be accounted for by globally estimated statistics.
LWIR Change Detection Algorithms

This chapter presents new methods for LWIR change detection which address some of the shortcomings of applying pre-existing SWIR methods to LWIR imagery. The benefit of hyperspectral data is that it has such a wealth of information in the spectral bands such that a single pixel from one time can be compared directly against a pixel from another time. However, if pixels are compared directly against each other, false changes can be detected from nuisance changes related to temperature and atmospheric effects. This effect is shown particularly in Figure 3.1 where changes in atmosphere and temperature result in different measurements for the same material. Therefore, other methods must be developed in order to compensate for these nuisance parameters.

In order to truly compare pixels from scenes, detection of object changes is desired. As stated previously, there are three elements to the measured radiance: emissivity, ground surface temperature, and atmospheric effects. Only the emissivity is an intrinsic property of objects whereas the object temperature and atmosphere can change between times and are not necessarily indicative of material changes. Therefore, in order to determine if an object has entered or left a scene, the object temperature and atmosphere must be compensated for via the reverse model stated in Equation (2.3) or other methods such as alpha residuals in Equation (2.10) or anomalous based methods in Equations [(2.34),(2.33)] can be used. In this thesis, it is assumed that the atmosphere is a global operator and is known via existing methods such as FLAASH-IR. However, the object temperature is unknown and must be
approximated in order to make an accurate change detection statistic.

Once the measured radiance values have been modified to eliminate nuisance parameters, the resulting data must be compared in such a way to quantify the separation between the resulting data. For this, a two class hypothesis test is used. In this test, the data from each time at a given spatial position \((x_i, y_j)\) are compared to test for changes.

\[
H_0 : \epsilon^{(1)}(x_i, y_j) = \epsilon^{(2)}(x_i, y_j) \tag{3.1}
\]

\[
H_1 : \epsilon^{(1)}(x_i, y_j) \neq \epsilon^{(2)}(x_i, y_j). \tag{3.2}
\]

In this test, the null hypothesis, \(H_0\), indicates that there is no true change in the pixel between the reference pixel, \(\epsilon^{(1)}\), and the test pixel, \(\epsilon^{(2)}\). The alternate hypothesis, \(H_1\)
indicates change due to structural changes or insertion/deletion of objects.

3.1 Smoothness-TES

In Section 2.2, smoothness-TES was described as a method of estimating temperatures with known atmospheres. This method allows the emissivity to be extracted as opposed to alpha residuals which extracts an abstract representation of emissivity. The simplest method of change detection with this method is to simply take the $\ell_2$ norm of the difference of emissivities of each time. This was presented in an earlier version of this work [8]

\[
\text{error} = \left\| \hat{\epsilon}^{(1)} - \hat{\epsilon}^{(2)} \right\|^2 .
\]

(3.3)

However, this method may suffer from inaccurate emissivity estimates and does not consider the noise profile in the computation of the error value.

The temperature element is inspired by [7]. For alpha residuals, temperature elements were annihilated via an orthogonal projection of the derivative of alpha residuals with respect to temperature. The same type of operation can be done with emissivity by taking the derivative of emissivity with respect to temperature. Additionally, in order to keep both emissivities in the same domain, the orthogonal projection needs to occur with respect to both derivatives simultaneously

First, compute the derivative of emissivity with respect to temperature

\[
\frac{\partial \epsilon(\lambda_i)}{\partial T} = \frac{(L(\lambda_i) - L_\downarrow(\lambda_i)\tau(\lambda_i) - L_\uparrow(\lambda_i)) \frac{\partial q(\lambda_i)}{\partial T}}{\tau(\lambda_i) (q(T, \lambda_i) - L_\downarrow(\lambda_i))^2} \quad (3.4)
\]

\[
\frac{\partial q(\lambda_i)}{\partial T} = \frac{2h^2c^3}{k\lambda_i^6T^2} \frac{e^{\frac{hc}{\lambda_i kT}}}{(e^{\frac{hc}{\lambda_i kT}} - 1)^2} . \quad (3.5)
\]
Next, remove the mean of emissivity to account for real world effects as will be described in Section 3.1.1. Finally, compute the orthogonal projection. The $M$ subscript indicates that the emissivity is modified to remove the derivative components.

$$P = A^T (AA^T)^{-1} A \quad (3.6)$$

$$A = \begin{bmatrix} \frac{\partial \tilde{\epsilon}(1)}{\partial T} & \frac{\partial \tilde{\epsilon}(2)}{\partial T} \end{bmatrix} \quad (3.7)$$

$$\tilde{\epsilon}_M^{(t)} = \tilde{\epsilon}^{(t)} - \tilde{\epsilon}^{(t)} P. \quad (3.8)$$

In addition, the noise profile can be computed as described by Equation (2.6). Note that the noise of every wavelength is independent from other wavelengths. This gives a covariance matrix with only diagonal components represented as $\Sigma$ which is the covariance matrix for $\hat{\epsilon}$ and $\tilde{\epsilon}$. In the next section, the full covariance matrix with orthogonal projection will be calculated. However, for the purposes of this thesis, only the emissivity noise is used as the full Mahalanobis distance is expensive to compute and does not improve overall performance. Computation of the covariance matrix as well as Mahalanobis distance requires 6 additional matrix multiplications which make the process nearly 500 times more expensive than the proposed methodology. Therefore, a standardized Euclidean distance is used

$$s_{Euc} = \left( \tilde{\epsilon}_M^{(1)} - \tilde{\epsilon}_M^{(2)} \right)^T \left( \text{diag} (\Sigma^{(1)}) + \text{diag} (\Sigma^{(2)}) \right)^{-1} \left( \tilde{\epsilon}_M^{(1)} - \tilde{\epsilon}_M^{(2)} \right). \quad (3.9)$$

### 3.1.1 Mean Subtraction

Upon examination of real world data, estimated emissivity between scenes appear to visually match but give poor results when compared via distance metrics. Consider Figure 3.2 which displays recovered emissivities of two different times. Visual inspection indicates that there is no true change occurring between times. All of the peaks and valleys occur in the same locations at approximately the same magnitudes. However, the overall emissivity
appears to be offset by a constant value across the entire emissivity. Therefore, by subtracting the mean of the emissivity, this effect may be partially avoided. The tilde above the emissivity indicates that the emissivity/derivative is mean subtracted

\[ \tilde{\epsilon}(\lambda_i) = \hat{\epsilon}(\lambda_i) - \frac{1}{N_c} \sum_{k=1}^{N_c} \hat{\epsilon}(\lambda_k) \]  

(3.10)

\[ \frac{\partial \tilde{\epsilon}(\lambda_i)}{\partial T} = \frac{\partial \hat{\epsilon}(\lambda_i)}{\partial T} - \frac{1}{N_c} \sum_{k=1}^{N_c} \frac{\partial \hat{\epsilon}(\lambda_k)}{\partial T}. \]  

(3.11)

Figure 3.2: View of two recovered emissivity vectors compared to truth
3.2 Alpha Residuals

In Section 2.3, alpha residuals were described as a new method for comparing two times in absence of temperature and atmospheric effects. This work originally focused on computing an emissivity-like vector without knowledge of temperature via Wein’s approximation of Planck’s function. However, newer developments added in the use of approximate temperatures in order to refine the computation.

In this thesis, the alpha residuals are substituted in for the emissivity values in the hypothesis test to give

\[ H_0 : \alpha^{(1)}_{M}(x_i, y_j) = \alpha^{(2)}_{M}(x_i, y_j) \]  
\[ H_1 : \alpha^{(1)}_{M}(x_i, y_j) \neq \alpha^{(2)}_{M}(x_i, y_j). \]  

In the work of Diani et al. [7], a new method for annihilating temperature effects from an emissivity. However, by removing the derivative for each time from the alpha residual separately, the alpha residuals are placed into different domains. This is because orthogonal projection only decorrelates temperature elements from the alpha residual rather than actually separating temperature. Because the values of the derivatives are different for each time, different vectors are decorrelated resulting in different planes.

Therefore, there must be some way to simultaneously remove both temperature derivatives from each alpha residual. This can be done via the general definition of orthogonal projection. Consider the derivatives with respect to temperature for each time as defined by (2.17). In order to remove both temperatures simultaneously, the generalized version of
projection may be used

\[ \bar{P} = A^T (AA^T)^{-1} A \]  

\[ A = \begin{bmatrix} \frac{\partial \alpha^{(1)}}{\partial T} & \frac{\partial \alpha^{(2)}}{\partial T} \end{bmatrix} \]  

\[ \alpha^{(t)} = \alpha^{(t)}_\epsilon - \alpha^{(t)}_\epsilon \bar{P} . \]  

The work of Diani et al. [7] focused primarily on the difference between a pixel in an image and a reference emissivity and computed the relative error between the two as defined in Equation (2.21). This method takes into account that if an alpha residual has more variance throughout the vector, then there is a greater chance for more noise to be incorporated into the final result. By normalizing the error by the magnitude, the expected noise can be dampened.

For the purposes of change detection, Equation (2.21) can reasonably be adapted to take into account the magnitude of both times equally

\[ s_{RE} = \left( \frac{\left\| \alpha^{(1)}_M - \alpha^{(2)}_M \right\|^2}{\left\| \alpha^{(1)}_M \right\| \left\| \alpha^{(2)}_M \right\|} \right) . \]  

(3.17)

Alternatively, a more naive approach would be to simply take the $\ell_2$ norm between the alpha residuals

\[ error = \left\| \alpha^{(1)}_M - \alpha^{(2)}_M \right\|^2 . \]  

(3.18)

In this method, there is no “truth” to consider. Therefore, both times must be considered in the total magnitude.

However, these methods do not take into account the system noise after applying the alpha residual transform nor do they account for the correlation present in the data. The alpha residual transform is akin to taking the logarithm of emissivity. As emissivity ap-
proaches 0, the log of emissivity becomes exponentially larger. As the emissivity becomes exponentially larger, the noise associated with the emissivity must also become exponentially larger. In addition, alpha residuals are by definition a mean subtracted statistic. This means that every single element in the alpha residual is dependent upon every other element within the alpha residual. Therefore, as detailed in the next subsection, the Mahalanobis distance is a much more appropriate method to compute the error metric of alpha residuals.

3.2.1 Covariance Matrix for Alpha Residuals

In order to compute the covariance matrix, there are 4 steps to consider: the noise of radiance, the noise of emissivity, the noise of alpha residuals, and the noise of modified alpha residuals. In this thesis, it is assumed that noise in the radiance domain is white Gaussian noise of constant variance. Therefore, noise of radiance can be represented as follows

\[
\hat{L}(\lambda_i) \approx N(L(\lambda_k), \sigma^2).
\] (3.19)

Next, the emissivity equation, (2.3), can be written to include noise as

\[
\hat{\epsilon}(\lambda_k, T) = \frac{L(\lambda_k) - L_{\downarrow}(\lambda_k) \tau(\lambda_k) - L_{\uparrow}(\lambda_k) + n}{(q_{BB}(T, \lambda_k) - L_{\downarrow}(\lambda_k)) \tau(\lambda_k)}. \tag{3.20}
\]

In order to compute the distribution of noise in the emissivity estimate, the constants can be ignored and the expected emissivity and be represented as Equation (3.21)

\[
\hat{\epsilon}(\lambda, T) \approx N(\epsilon(\lambda_k, T), \sigma^2(\lambda_k)). \tag{3.21}
\]

For simplicity, Equation (3.21) can be rewritten as

\[
\hat{\epsilon}(\lambda, T) \approx N(\epsilon(\lambda_k, T), \sigma^2(\lambda_k)). \tag{3.22}
\]
From there, the noise needs to propagate through the alpha emissivity equation as described in Equation (2.10). In order to compute this statistic, the log of a Gaussian distribution must be calculated. Consider the following system

\[ n_e(\lambda_k) \approx N(\epsilon(\lambda_k), \sigma^2) = \frac{1}{\sqrt{2\pi \sigma^2(\lambda_k)}} \exp \left( \frac{(x - \epsilon(\lambda_k))^2}{2\sigma^2(\lambda_k)} \right). \] (3.23)

Next, by taking the log of a normal distribution, a distribution transformation must be defined

\[ n_x(\lambda_k) \approx \log(n_e(\lambda_k)) \] (3.24)

\[ P_y(y) = P_x \left( g^{-1}(y) \right) \left| \frac{dg^{-1}(y)}{dy} \right| \] (3.25)

\[ g(x) = \log(x) \] (3.26)

\[ g^{-1}(y) = e^y \] (3.27)

\[ \left| \frac{dg^{-1}(y)}{dy} \right| = e^y \] (3.28)

\[ P_y(y) = \frac{1}{\sqrt{2\pi \sigma^2_x}} \exp \left( -\frac{(e^y - \epsilon)^2}{2\sigma^2_x} \right) e^y. \] (3.29)

From there, the mean and variance of the system can be calculated as follows

\[ \mu = \int_{-\infty}^{\infty} y P_y(y) dy \] (3.30)

\[ \sigma^2 = \int_{-\infty}^{\infty} y^2 P_y(y) dy - \mu^2. \] (3.31)
However, this integral is not solvable in closed form. In order to solve Equations (3.30) and (3.31), integration by parts must be used which requires a solution of the following integral

\[
\int_{-\infty}^{\infty} \text{erf} \left( \frac{k_1 - e^{y}}{k_2} \right) dy.
\] (3.32)

This particular error function integral is not found in any existing integral libraries. Therefore, some approximation must be used. If some random variable \( n_x \) with mean \( \mu_x \) and variance \( \sigma_x^2 \) goes through the function \( n_y = \log(n_x) \), then the approximate distribution of \( n_y \) is

\[
n_y \approx \mathcal{N} \left( \log(\mu_x) - \frac{\sigma_x^2}{2\mu_x}, \frac{\sigma_x^2}{\mu_x^2} \right).
\] (3.33)

This approximation is valid provided \( \frac{\mu_x}{\sigma_x} \gtrsim 3 \). This is because the log of a variable approaches negative infinity as the variable approaches 0. This effect can skew the distribution and make the distribution non-zero and even introduce imaginary numbers. However, if the conditions are met, the new distribution is relatively unaffected by this skewedness. For this problem, the mean is almost always several orders of magnitude greater than the variance so this is not a problem. Empirical testing indicates that this approximation gives at least 4 orders of magnitude of accuracy for this problem.

With this approximation, the covariance matrix for alpha residuals can be calculated. Consider Equation (2.10), this equation has two duplicated portions of the form

\[
x_k = \lambda_k \log(\epsilon(\lambda_k)).
\] (3.34)

The noise profile of \( x_i \) can be extrapolated from Equations (3.33) and (3.21) to give

\[
\hat{x}(\lambda_k) \approx \mathcal{N} \left( \lambda_k \left( \log(\epsilon(\lambda_k)) - \frac{\sigma_e^2(\lambda_k)}{2(\epsilon(\lambda_k))^2} \right), \sigma_k^2 \left( \frac{\sigma_e^2(\lambda_k)}{(\epsilon(\lambda_k))^2} \right) \right)
\] (3.35)

\[
\approx \mathcal{N} \left( \mu_x, \sigma_x^2 \right).
\] (3.36)
Next, the components must be combined to obtain the full covariance matrix of the alpha residual. Due to the mean subtracted nature of alpha residuals, the covariance can be represented as

\[
\text{cov}(y_i, y_j) = E[y_i y_j].
\] (3.37)

In this case, \(y_i\) and \(y_j\) indicate different wavelengths in the alpha residual. By substituting Equations (2.10) and (3.34) into Equation (3.37), the following system arises

\[
\text{cov}(y_i, y_j) = E\left[\left(x(\lambda_i) - \sum_{k=1}^{N_c} x(\lambda_k)\right)\left(x(\lambda_j) - \sum_{k=1}^{N_c} x(\lambda_k)\right)\right]
\] (3.38)

\[
A = \sum_{k=1}^{N_c} x(\lambda_k)
\] (3.39)

\[
\text{cov}(y_i, y_j) = E\left[(x(\lambda_i) - A)(x(\lambda_j) - A)\right]
\] (3.40)

\[
= E\left[x(\lambda_i)x(\lambda_j) - x(\lambda_i)A - x(\lambda_j)A + A^2\right]
\] (3.41)

\[
= E[x(\lambda_i)x(\lambda_j)] - E[x(\lambda_i)A] - E[x(\lambda_j)A] + E[A^2]
\] (3.42)

\[
E[A^2] = \frac{1}{N_c^2} \sum_{k=1}^{N_c} \sigma_x^2(\lambda_k)
\] (3.43)

\[
E[x(\lambda_i)A] = E\left[\frac{x(\lambda_i)}{N_c}\left(\frac{1}{N_c} x(\lambda_1) + \frac{1}{N_c} x(\lambda_2) + \ldots + \frac{1}{N_c} x(\lambda_n)\right)\right]
\] (3.44)

\[
= E\left[0 + 0 + \ldots + \frac{1}{N_c} x^2(\lambda_i) + 0\right] = \frac{1}{N_c} \sigma_x^2(\lambda_i).
\] (3.45)

If \(i = j\) then

\[
E[x(\lambda_i)x(\lambda_j)] = \sigma_x^2(\lambda_k).
\] (3.46)

Otherwise if \(i \neq j\), then the variables are uncorrelated per the definition of noise for this thesis

\[
E[x(\lambda_i)x(\lambda_j)] = 0.
\] (3.47)
By combining all of these elements, the covariance matrix can be computed as follows

$$
\Sigma(\lambda_i, \lambda_j) = \text{cov}(y_i, y_j) = \frac{1}{N^2} \sum_{k=1}^{N} \sigma_x^2(\lambda_k) - \frac{1}{N} \sigma_x^2(\lambda_i) - \frac{1}{N} \sigma_x^2(\lambda_j) \quad i \neq j
$$

(3.48)

$$
\Sigma(\lambda_i, \lambda_j) = \text{cov}(y_i, y_j) = \sigma_x^2(\lambda_i) + \frac{1}{N_c} \sum_{k=1}^{N} \sigma_x^2(\lambda_k) - \frac{1}{N} \sigma_x^2(\lambda_i) - \frac{1}{N_c} \sigma_x^2(\lambda_j)
$$

(3.49)

$$
= \left( \frac{N_c - 1}{N} \right)^2 \sigma_x^2(\lambda_i) + \frac{1}{N_c} \sum_{k \neq j}^{N_c} \sigma_x^2(\lambda_k) \quad i = j.
$$

(3.50)

Finally, the computed covariance matrix must be adjusted to account for the orthogonal projection operator $P$ described in Equation (3.16). To do this, rewrite Equation (3.52) as

$$
\alpha_M = (I - P)\alpha.
$$

(3.51)

This means that there is a matrix multiplied by a random vector, this means that the final covariance matrix can be written as follows

$$
\Sigma_{full} = \Sigma - P\Sigma P.
$$

(3.52)

Utilizing the covariance matrix computed in Equation (3.52), the change detection metric can be computed as follows

$$
\text{error} = \sqrt{\left( \alpha_M^{(1)} - \alpha_M^{(2)} \right) \Sigma_{full}^{-1} \left( \alpha_M^{(1)} - \alpha_M^{(2)} \right)}.
$$

(3.53)

To include information from both covariance matrices, the following modification could be
made

\[
\text{error} = \sqrt{\left(\alpha_M^{(1)} - \alpha_M^{(2)}\right) \left(\Sigma_{full}^{(1)} + \Sigma_{full}^{(2)}\right)^{-1} \left(\alpha_M^{(1)} - \alpha_M^{(2)}\right)}.
\] (3.54)

However, the full Mahalanobis distance is computationally and memory intensive as it requires 6 matrix multiplications potentially hundreds of thousands of times. Additionally, empirical evidence shows that the off-diagonal elements of the covariance matrix are at least 2 orders of magnitude less than the main diagonal. Therefore, Equation (3.54) can reasonably be reduced to a modified standardized Euclidean distance

\[
s_{Euc} = \left(\alpha_M^{(1)} - \alpha_M^{(2)}\right)^T \left(\text{diag} \left(\Sigma^{(1)}\right) + \text{diag} \left(\Sigma^{(2)}\right)\right)^{-1} \left(\alpha_M^{(1)} - \alpha_M^{(2)}\right).
\] (3.55)

In conclusion, the covariance matrix may be computed via the following steps:

1. Compute the emissivity and alpha residual from the temperature estimate via Equations (2.3) and (2.16)

2. Compute the noise distribution of emissivity via Equation (2.6)

3. Compute the noise of alpha residuals sub-component via Equation (3.35)

4. Compute the full covariance matrix from the sub-components via Equations (3.48) and (3.50)

5. (Optional) Finally, adjust the covariance matrix to account for orthogonal projection via Equation (3.52)

### 3.2.2 Downside of Alpha Residuals

As noted in Section 3.1.1, there appears to be a constant offset on emissivities. However, by definition, alpha residuals are already mean subtracted as a way to avoid the need for a good temperature estimate. Therefore, alpha residuals may not be adjusted in this manner.
Consider a pair of emissivities in which the reference emissivity has a constant value of 0.9 across the entire spectra. Under the assumption that environmental effects can offset the emissivity by a constant value, the emissivity from time 2 may have a constant value of 0.8 across the entire spectra. Due to the way the alpha residuals are calculated, the resulting alpha residuals will be irreconcilably different as shown in Figure 3.3. One possible solution to this problem is to force the slope of the alpha residual to be zero but that is beyond the scope of this thesis.

Figure 3.3: Alpha residual as a function of emissivity
3.3 Joint Temperature Optimization (JTO)

The problem with most tradition methods of processing LWIR hyperspectral data is that the methods are concerned with extracting individual cubes and comparing them or creating some generalized observation. The problem with this is that it doesn’t actually compare images directly but rather puts images onto a similar domain and hopes that the domains are close enough. In addition, as stated previously, generalized observations are incapable of dealing with the spatially varying temperature element. Therefore, in this section, a new method is developed which forces each cube into the same domain. This is done by assuming the null hypothesis that there is no true change between times. With this assumption, a process may be developed which attempts to bring two pixels together as a function of temperature. From there, the change detection statistic may be calculated from the residual error of this optimization process. This method, known as Joint Temperature Optimization (JTO), is represented as follows [8]

\[ \text{error} = \min_{T^{(1)}, T^{(2)}} \left\| \epsilon^{(1)}(T^{(1)}) - \epsilon^{(2)}(T^{(2)}) \right\|_2. \]  

(3.56)

In order to comply the mean subtraction element described in Section 3.1.1, Equation (3.56) can be modified as follows

\[ \text{error} = \min_{T^{(1)}, T^{(2)}} \left\| \epsilon^{(1)}(T^{(1)}) - \frac{1}{N_c} \sum_{k=1}^{N_c} \epsilon^{(1)}(T^{(1)}, \lambda_k) - \epsilon^{(2)}(T^{(2)}) + \frac{1}{N_c} \sum_{k=1}^{N_c} \epsilon^{(2)}(T^{(2)}, \lambda_k) \right\|_2. \]  

(3.57)

In theory, this method works on the principle that if the emissivity does not change, then an optimizer should be able to find some temperature pair to reduce the error to the point of noise as shown in Figure 3.4a. However, if there is a true change, an optimizer will fail to find a temperature pair and will have a large residual error as shown in Figure 3.4b.

Alternatively, the standardized Euclidean distance developed in Equation (3.9) could
be used. However, this type of computation is not recommended due to computational complexity, lack of derivatives, and minimal performance improvement.

The problem with JTO is that it requires a double variable optimization process for large numbers of pixels. Utilizing a naïve approach without derivatives (e.g., fmincon in MATLAB), this process can take upwards of an hour for 100,000 pixels. This means that it is impractical for use in scenarios where real-time on-board processing is desirable. Therefore, a new method must be developed to make JTO practical for real world implementation.

### 3.3.1 JTO-Linear

The first possibility is an aggressive approach which attempts to solve for both temperatures in closed form. Consider two scenes with temperatures varying between 300 and 330 degrees Kelvin. In this case the temperature range is approximately known and an approximate mean temperature of 315 K may be used as a rough starting point. In addition, the derivatives of emissivity with respect to temperature can be computed via Equation (3.4). If the derivative is “linear” with respect to temperature, then it should be possible to linearly traverse the derivative. Therefore, via linearization, a least squares temperature estimate
may be computed to find a new temperature pair which minimizes emissivity error. This process is better known as Gauss-Newton’s optimization. This process can be represented as follows [8]

1. Pick a mean temperature for each scene.

2. Compute emissivity of each pixel for both scenes as a function of the chosen temperatures per Equation (2.3).

3. Compute the derivatives with respect to temperature per Equation (3.4) centered at the chosen temperature

4. Subtract the means of emissivity and derivative of emissivity via Equations (3.10) and (3.11)

5. Compute base emissivity error as a function of the initial temperature

\[
\delta = \hat{e}^{(1)}(T_1) - \hat{e}^{(2)}(T_2)
\]  \hspace{1cm} (3.58)

6. Compute the jacobian using the derivatives of emissivity

\[
J = \begin{bmatrix}
\frac{\partial \hat{e}^{(1)}}{\partial T^{(1)}} & -\frac{\partial \hat{e}^{(2)}}{\partial T^{(2)}}
\end{bmatrix}
\] \hspace{1cm} (3.59)

7. In this system, the goal is to find some \( \Delta_T = \begin{bmatrix} \Delta_T^{(1)} & \Delta_T^{(2)} \end{bmatrix} \) which minimizes \( r_0 \).

Therefore, the system can be set up as follows

\[
J \Delta_T = \delta
\] \hspace{1cm} (3.60)
8. Compute the residual error

\[ s_{linear} = ||\delta - J\Delta_T||_2 \]  

(3.62)

If the system is linear throughout the entire function space, this solution should be able to very accurately describe the JTO function space and give an extremely accurate solution.

### 3.3.2 JTO-Newton's

One potential problem with JTO-Linear is that it assumes that the derivative of emissivity with respect to temperature is linear. This is a relatively poor assumption if temperatures vary from the mean temperature by more than 5 kelvin. The following method attempts to alleviate this effect by iterating portions of JTO-Linear.

In JTO-Newton's, the goal is to gradually approach the true temperature pair via Levenberg-Marquardt optimization [16]. JTO-linear was a single iteration of Gauss-Newton's method. This performed poorly due to the non-linear nature of Planck's function as well as low rank of the approximate hessian. This numerical instability occurs when the atmospheres for both scenes are very similar. When this occurs, the derivatives are also similar at all temperatures and makes the jacobian and approximate hessian low rank. Currently, there are several other conditions which result in poor convergence but these conditions are currently unknown. In addition, Gauss-Newton's method suffers from poor stability and tends towards saddle points with the only benefit being convergence speed. Therefore, it is desirable to modify the methodology of JTO-Linear into JTO-Newton's as follows:
1. Pick a starting temperature for each time as well as a set $\mu$ value between $10^{-6}$ and $10^{-7}$, higher $\mu$ results in a more stable hessian at the cost of slower convergence. This value is empirically derived and is approximately two orders of magnitude lower than the magnitude of the approximate hessian. Define minimum and maximum temperatures for each time. This bounds the system into realistic temperatures and prevents overfitting. Define number of iterations. While this step diverges from the standard Levenberg-Marquardt methodology, these parameters eliminate the need for additional function evaluations as well as many intermediate steps.

2. If the temperatures are greater than the maximum temperature or less than the minimum temperature, force the temperature to the maximum/minimum temperature respectively.

3. Compute $R$ and $r_0$ per steps 2 through 6 in JTO-linear.

4. Compute the damped least-squares step required for Levenberg-Marquardt

$$\Delta T = (JJ^T + \mu I)^{-1} J^T \delta$$  \hspace{1cm} (3.63)

5. Add the change in temperature to the temperature from the previous step to get a new updated temperature

$$T^{(n+1)} = T^{(n)} + \Delta T$$ \hspace{1cm} (3.64)

6. If the number of iterations does not exceed the maximum iterations, repeat steps 2 through 5, otherwise continue.

7. Ensure $\Delta T$ does not force $T$ to exceed boundaries, if it does, set $\Delta T$ such that $T$ is on the boundary.

42
8. Compute the final residual error. This residual error becomes the detection statistic.

\[ s_{\text{Newton}} = \|\delta - J\Delta_T\|_2^2 \]  

(3.65)

By taking the stability of gradient descent as well as the aggressiveness of Gauss-Newton’s optimization, this methodology allows a very quick approach to the same solution found in naïve JTO as well as an extremely fast solution. This methodology is able to solve for 100,000 pixels in just a half a second compared to over an hour with a naïve solver.

### 3.4 Single Temperature Dependence

In Section 3.3, a new method of change detection was proposed relying on the underlying assumption that if there is no true change between times, then the emissivities will be the same. However, this requires a double variable optimization process which is reasonably expensive. Therefore, it is desirable to reduce the number of variables to effectively reduce the computational complexity. This brings up the question: if an emissivity is chosen, can the temperatures be estimated in closed form from the emissivity? Consider Equation (2.3)

\[ \epsilon(\lambda_k) = \frac{L(\lambda_k) - L_\downarrow(\lambda_k)\tau(\lambda_k) - L_\uparrow(\lambda_k)}{q_{BB}(T, \lambda_k) - L_\downarrow(\lambda_k)} \tau(\lambda_k). \]  

(3.66)
If the emissivity and atmospheres are known, it should be possible to extract the temperatures

\[
(q_{BB}(T, \lambda_k) - L_\downarrow(\lambda_k)) \tau(\lambda_k) = \frac{L(\lambda_k) - L_\downarrow(\lambda_k) \tau(\lambda_k) - L_\uparrow(\lambda_k)}{\epsilon(\lambda_k)}
\]  
(3.67)

\[q_{BB}(T, \lambda_k) = L_{BB}(\lambda_k) = \frac{L(\lambda_k) - L_\downarrow(\lambda_k) \tau(\lambda_k) - L_\uparrow(\lambda_k)}{\tau(\lambda_k) \epsilon(\lambda_k)} + L_\downarrow(\lambda_k)
\]  
(3.68)

\[T(\lambda_k) = q_{BB}^{-1}(\frac{L(\lambda_k) - L_\downarrow(\lambda_k) \tau(\lambda_k) - L_\uparrow(\lambda_k)}{\tau(\lambda_k) \epsilon(\lambda_k)} + L_\downarrow(\lambda_k))
\]  
(3.69)

where

\[q_{BB}^{-1}(L_{BB}, \lambda_k) = \frac{2hc^2}{\log \left( \frac{hc}{L_{BB}(\lambda_k)k\lambda_k} + 1 \right)}.
\]  
(3.70)

Equation (3.69) gives a single temperature for each wavelength of what the temperature should be given the input emissivity. As a least squares estimate would be difficult to obtain, the approximated temperature can be obtained as

\[
\hat{T} = \frac{1}{N_c} \sum_{k=1}^{N_c} T(\lambda_k).
\]  
(3.71)

The temperature in Equation (3.71) can then be used for classification and change detection purposes.

The main problem with this method is that it cannot comply with Section 3.1.1 due to the nonlinear nature of the inverse blackbody function. In addition, as of now it is impossible to obtain a least squares estimate of temperature using this method.

The following sections will utilize this methodology in order to compute a change detection statistic. The primary methodology is to vary the temperature estimation for one time, compute the emissivity and enforce the emissivity onto the other time and compute
the temperature estimates for the second time. From there, check the model fit utilizing a
\(\ell_2\) norm or standardized Euclidean distance.

### 3.4.1 Single Temperature Variance Optimization

The original method for computing a change metric using single variable optimization is
as follows.

1. Compute the emissivity as a function of temperature for time 1

\[
\hat{\epsilon}^{(1)}(\lambda_k, T^{(1)}) = \frac{L^{(1)}(\lambda_k) - L_{\downarrow}^{(1)}(\lambda_k)\tau^{(1)}(\lambda_k) - L_{\uparrow}^{(1)}(\lambda_k)}{q_{BB}(T^{(1)}, \lambda_k) - L_{\downarrow}^{(1)}(\lambda_k)\tau^{(1)}(\lambda_k)} \tag{3.72}
\]

2. Enforce the emissivity from time 1 onto time 2 and compute the temperature vector
using Equation (3.69)

\[
\hat{T}^{(2)}(\lambda_k) = q_{BB}^{-1} \left( L^{(2)}(\lambda_k) - \frac{L_{\downarrow}^{(2)}(\lambda_k)\tau^{(2)}(\lambda_k) - L_{\uparrow}^{(2)}(\lambda_k)}{\tau^{(2)}(\lambda_k)\hat{\epsilon}^{(1)}(\lambda_k)} + L_{\downarrow}^{(2)}(\lambda_k) \right) \tag{3.73}
\]

3. Compute the variance of \(\hat{T}^{(2)}\) where \(\mu_{T^{(2)}}\) is the mean of the output temperatures

\[
\sigma^2(T^{(1)}) = \frac{1}{N_c} \sum_{k=1}^{N_c} \left( \hat{T}^{(2)}(\lambda_k) - \mu_{T^{(2)}} \right)^2 \tag{3.74}
\]

From there, minimize the variance obtained in Equation (3.74) to find the minimal error.

\[
s_{\text{var}} = \min_{T^{(1)}} \sigma^2(T^{(1)}) \tag{3.75}
\]

Ideally, if the materials are the same, the materials should have the same emissivity. As
a result the reverse operation should perfectly bring out the blackbody function in time
2. The variance measured in Equation (3.74) measures this conformance to the blackbody function. Thus, it is possible to perform an approximation of JTO utilizing only one temperature.

### 3.4.2 Cross Radiance Optimization

However, utilizing variance is a relatively weak statistic for determining whether a change has occurred. It does not take into account noise not does it consider the emissivity. In order to alleviate this, it is desirable to perform a change detection metric in radiance space via the following methodology.

1. Compute the temperature vector for time 2 using Equations (3.72) and (3.73) from a temperature estimate for time 1

2. Compute an estimate for the temperature in time 2 ($T^{(2)}$) using Equation (3.71)

3. Using $T^{(2)}$, compute the emissivity for time 2

$$\hat{\epsilon}^{(2)}(\lambda_k) = \frac{L^{(2)}(\lambda_k) - L^{(2)}(\lambda_k)\tau^{(2)}(\lambda_k) - L^{(2)}(\lambda_k)}{(q_{BB}(T^{(2)}, \lambda_k) - L^{(2)}(\lambda_k))\tau^{(2)}(\lambda_k)}$$  \hspace{1cm} (3.76)

4. Use the time 2 emissivity to recreate time 1 radiance using Equation (2.1)

$$\hat{L}^{(1)}(\lambda_k) = \hat{\epsilon}^{(2)}(\lambda_k)q_{BB}(T^{(1)}, \lambda_k)\tau^{(1)}(\lambda_k) + (1 - \hat{\epsilon}^{(2)}(\lambda_k))L^{(1)}(\lambda_k)\tau^{(1)}(\lambda_k) + L^{(1)}(\lambda_k)$$  \hspace{1cm} (3.77)
Utilizing Equation (3.77), a new minimization process can be set up as follows where \( f \left( \hat{T}^{(1)} \right) \) is the inverse operation described in steps 1 through 3

\[
s_{\text{cross}} = \min_{T^{(1)}} \left\| L^{(1)} - \hat{L}^{(1)} \left( f \left( \hat{T}^{(1)} \right) \right) \right\|^2_2. \tag{3.78}
\]

Alternatively, a more thorough approach would be to perform the process twice to eliminate bias depending on which time is chosen to be reference time. However, this process requires two minimization processes which somewhat defeats the purpose of using a single variable solver

\[
s_{\text{full}} = \min_{T^{(1)}} \left\| L^{(1)} - \hat{L}^{(1)} \left( f \left( \hat{T}^{(1)} \right) \right) \right\|^2_2 + \min_{T^{(2)}} \left\| L^{(2)} - \hat{L}^{(2)} \left( f \left( \hat{T}^{(2)} \right) \right) \right\|^2_2. \tag{3.79}
\]

### 3.4.3 Closed Form Using Known Temperatures

The final method assumes that temperatures have already been estimated for the reference cube using a temperature estimation method such as smoothness-TES. Under this method, no further temperatures need to be estimated. Therefore, the problem becomes completely closed form via the following process.

1. Perform steps 1 through 4 of Cross Radiance Optimization to obtain \( \hat{L}^{(1)} \)
2. Compute the error metric

\[
s_{\text{closed}} = \left\| L^{(1)} - \hat{L}^{(1)} \left( f \left( \hat{T}^{(1)} \right) \right) \right\|^2_2 \tag{3.80}
\]

In this way, a change detection metric can be computed in closed form assuming pre-processing has already been done. In the real world, it is desirable to compute statistics on-board as to remove the cost of sending data to the ground as well as reducing latency. Assuming temperature estimates exist, this closed form solution can be computed up to 20
times faster than traditional methods and is much more likely to run in real time giving it an edge over the other methods presented in this thesis.

3.5 LWIR Data Challenges

Longwave infrared data presents several challenges associated with the proposed formulation of the measured radiance as well as emissivity. In this section, smoothness-TES’s dependence on atmosphere, and invalid denominators for emissivity, and runaway temperatures will be discussed.

3.5.1 Smoothness-TES Dependence on Atmosphere

Smoothness-TES runs under the assumption that an atmosphere is corrupting the radiance. It assumes that the underlying emissivity is “smooth” and the atmosphere is extremely “jagged” due to absorption band. Therefore, smoothness-TES proposes that one can search a range of temperatures and find a temperature in which the resultant emissivity is “smooth.” However, this jaggedness is not always true for all atmospheres. Consider a case in which the atmosphere is transmissive and has low influence from external sources: \((\tau \rightarrow 1, L_\downarrow \rightarrow 0, L_\uparrow \rightarrow 0)\). In this case, there is no “jaggedness” to remove and therefore smoothness-TES will fail. In Figure 3.5, sets of measured radiance realizations were generated. In each set, a different atmosphere was chosen and all sets were run through smoothness-TES. From there, the temperature error of smoothness-TES was compared to the known temperatures is computed and compared against a “jaggedness” metric of the atmospheres on the y-axis. In the case of transmissive atmospheres, it may be better to consider alternative methods to smoothness-TES for temperature estimation.
3.5.2 Invalid Denominators

In Equation (2.3), the emissivity may be calculated assuming all other factors are known. However, there is a case where this is not valid. Consider the denominator of the emissivity equation

\[ \text{denom} = \frac{1}{q_{BB}(T, \lambda_k) - L_{\downarrow}(\lambda_k)}. \] (3.81)

In the denominator, there is a potential for the blackbody function to equal the downwelling radiance. If this is the case for any wavelength, the emissivity will go to infinity at that wavelength. This is an undesirable characteristic and will result in large error values for a range around temperatures \( q_{BB}(T, \lambda_k) \approx L_{\downarrow}(\lambda_k) \). These temperatures may be estimated via the inverse Planck function defined in Equation (3.70). The critical points for an example downwelling radiance are shown in Figure 3.6.
When combined with the error functions presented in this work, there are large errors presented at the critical locations as shown in Figure 3.7. Regrettably, in these conditions, an optimizer cannot traverse the temperature space and will likely get stuck in some local minima resulting an incorrect temperature estimate. In a worst case scenario, the true temperature is located in the midst of critical temperatures and will receive large error values even at the true location. There are a few options to avoid this type of error such as excluding contributions of bands near their critical points. However, this does not work well due to the error function becoming piece-wise and thus difficult to traverse. Overall, the best method to avoid this type of error is to just exclude the offending bands entirely.

3.5.3 Runaway Temperatures

Another problem associated with LWIR data is the problem of runaway temperatures. Consider the blackbody function in Equation (2.2). As $T$ increases, the exponential decreases and thus causes the radiance to increase for all wavelengths. Therefore, as $T \to \infty$, $q(T, \lambda_k) \to \infty$. By extension, this means that the emissivity will go to 0 because the denominator of the emissivity equation goes to infinity as well. When computing error val-
values, this becomes a problem in metrics where the emissivity is considered. If the emissivity is zero, then the error value is also zero. This is best illustrated in Figure 3.8 which illustrates the error function of the original smoothness-TES function as defined in Equation (2.4). In this figure, if the starting temperature of the optimization process is greater than 330K, then most optimization processes will simply increase the temperature to lower the error. This type of error only exists in the emissivity space and all methods operating in radiance or alpha residual space are immune to these effects.
Figure 3.8: Error function of original smoothness-TES
Results

4.1 Synthetic Data

In order to fully understand each algorithm presented in Chapter 3 and Section 2.5, it is important to consider each algorithm in a controlled situation. The benefit of this type of analysis is that there are no misregistratIon errors, incorrect atmospheres, mislabeled pixels, mixed pixels, or other nuisance parameters within the scene. This allows for an ideal scenario to test a large variety of algorithms with “unlimited” pixels. In addition, high quality labeled real data is difficult to come by in change pairs for LWIR sensors. Real-world data may also be noisy and/or possess sensor/calibration artifacts not accounted for in physical models.

Another benefit of synthetic data is the large range of possible scenarios. In real data, there is a relatively limited number of changing materials, atmospheres, and temperatures. In synthetic data, the entire range of change and non-change pairs may be explored. In the existing real dataset, there exists a maximum of 30 materials where most changes occur between a target and ground. In the synthetic data-set, an emissivity library with 3649 possible materials of which over 13 million change pairs may be explored. In addition, a database with over 10,000 possible atmospheres allowing for over $10^{15}$ combinations with several orders of magnitude more freedom for temperature combinations. All in all, synthetic data allows for thorough “unlimited” testing to determine which algorithms are viable for further testing and implementation.
4.1.1 Experiment Setup

As stated previously, there are nearly infinite combinations of synthetic data. However, due to obvious computational limitations and algorithm considerations, this data-set must be truncated. For this thesis, a data-set consisting of 10 million representative samples was created. These samples are created to emulate samples gathered by the Spatially-Enhanced Broadband Array Spectrograph System (SEBASS) sensor [13]. The SEBASS sensor gathers 128 bands between 7.6 and 13.5µm. However, due to limitations described in Section 3.5.2, the downwelling radiance is simply too high at some wavelengths. In order to adjust for this, the offending wavelengths are truncated to 88 spectral measurements between 8.2 and 12.4 µm. Within these parameters, the dataset is generated as follows:

- Two unique atmospheres drawn from a pre-generated dataset. This dataset consists of 10,000 atmospheres drawn from MODTRAN [2]. The MODTRAN parameters assume that a drone is flying at 1,500 feet with minimal cloud cover. The dataset was created by varying water vapor content, ozone content, and ground surface temperature over a broad range of potential values.

- An emissivity library is used consisting of 3649 spectra. If the pixel is labeled as non-change, the emissivity for each time will be the same. Otherwise, different emissivities will be chosen for each time. The chance of a pixel being a change target is 1%. The reasoning for this value is due to the way Chronochrome and Covariance-Equalization work. While all of the other methods in this paper work on a pixel-by-pixel basis, the anomalous detection methods work on globally estimated covariance matrices which are extremely sensitive to change pixels. If this number is raised, the synthetic scene will be unrealistic for these methods and give an unfairly poor performance.

- Unique temperatures are generated for each time. Each pixel receives a unique temperature drawn uniformly at random between 300 and 330 Kelvin. This range is
chosen because it avoids the problem stated in Section 3.5.2. In addition, it provides a realistic range of temperatures for a real scene.

- Independently identically distributed (IID) Gaussian noise is added to the scene. The standard deviation of this noise is 0.5 microflicks. This is in accordance with the estimated noise values for the SEBASS sensor.

- The synthetic radiance realizations are generated from the drawn parameters using equation (2.1).

In order to achieve a fuller realization, 1,000 sets of data with 10,000 measured radiance values were generated. In each set, a different atmosphere pair was chosen to allow for a full test with different atmospheres. The chosen atmosphere can have a drastic effect on how each algorithm performs. In order to compare the methods, the error value for every single pixel for every single atmosphere case was combined into a single receiver operating characteristic curve (ROC). This type of plot treats the data as a two class problem. The plot varies the threshold value for determining a change and compares how many true changes have been detected verses how many non-change pixels have been marked as changes as a function of the threshold value.

### 4.1.2 Anomalous Detection Methods

In Section 2.5, two pre-existing metrics were proposed for use in LWIR change detection: Covariance Equalization and Chronochrome. The results are shown in Figure 4.1 utilizing a ROC curve. As evidenced by the figure, the standard version of covariance equalization performs the best. However, the overall performance is much poorer than existing methods which will be discussed later.
4.1.3 Smoothness-TES

In Section 3.1, a variety of improvements and possibilities were suggested on how to perform change detection utilizing the temperatures outputted by Borell’s smoothness-TES algorithm. This consists of 3 different sub-classes for a total of 8 combinations. The sub-classes are orthogonal projection presented in Equation (3.8), standardized Euclidean distance presented in Equation (3.9), and the mean subtraction element presented in Section 3.1.1. Each of these elements can be combined in different ways and it is desirable to consider which processes improve performance the most.

In Figure 4.2, only the orthogonal projection element is considered. In this figure, a ROC curve is displayed. The line on top represents the performance of smoothness-TES if orthogonal projection is used to annihilate any residual temperature elements whereas the line on the bottom lacks this adjustment. This shows that applying orthogonal projection
has a significant impact on the overall performance of smoothness-TES for the purposes of change detection. In this approach, there is no compensation for noise or mean subtraction. The $\ell_2$ norm presented in Equation (3.3) is used.

![Graph showing ROC curve improvements](image)

Figure 4.2: Results for smoothness-TES change detection with orthogonal projection improvements

In Figure 4.3, only the noise adjustment via standardized Euclidean distance presented in Equation (3.9) is considered. The line on top represents the error metric defined in Equation (3.3). This shows that taking the standardized Euclidean distance has little to negative effect on the performance of the ROC curve and thus should not be considered for further development. This is somewhat expected as all emissivities are in the same frame between 0 and 1 whereas alpha residuals may vary between negative infinity and positive infinity. Therefore, it is reasonable to assume that band noise would likely have a much smaller effect on emissivity. These results do not include orthogonal projection or mean
subtraction.

Figure 4.3: Results for smoothness-TES change detection with standardized Euclidean error metric

Finally, in Figure 4.4, only the mean subtraction element described in Section 3.1.1 is considered. The line on top represents the improvements given by the mean subtraction elements. These results show that mean subtraction has a significant improvement on the overall performance. In this approach, there was no orthogonal projection or standardized Euclidean distance.

The results for all methods are combined in Figure 4.5. The lines are shown from top to bottom (First line is hidden by second line and third line is hidden by fourth line because performance is exactly the same). As evidenced, the orthogonal projection implemented in Equation (3.8) is the most significant improvement in this thesis. Overall,
standardized Euclidean distance has a slight improvement when combined with orthogonal projection. However, this improvement is extremely negligible compared to the overall performance as the difference is only seen below \( p_{fa} = 10^{-4} \). At that point, there are simply not enough points to generate a reliable ROC curve. The benefits of the mean subtraction element diminish with the use of orthogonal projection. Overall, the most optimal method as a function of difficulty, implementation, runtime, and performance is the variant of smoothness-TES with orthogonal projection, \( \ell_2 \) norm, and mean subtraction.
4.1.4 Alpha Residuals

In Section 3.2, 3 different methods for computing a change detection metric were presented: $\ell_2$ norm presented in Equation (3.18), relative error metric presented in Equation (3.17), and the standardized Euclidean distance described by Equation (3.55). The results of all methods are compiled in Figure 4.6. As expected, the standardized Euclidean distance outperforms the $\ell_2$ norm and relative error metrics.

4.1.5 Joint Temperature Optimization

In Section 3.3.1, two new methods were introduced specifically to compute a change detection metric for LWIR data: JTO-Linear and JTO-Newton. Naïve JTO is excluded as
JTO-Newton gives the same results. The comparison of these two methods is shown in Figure 4.7. Originally, JTO-linear was developed as an extremely fast approximation compared to JTO. At the time, the full JTO computation took over an hour whereas JTO-Linear took less than a second. However, with the various improvements of JTO-Newton, both methods are capable of running in real time with system appropriate hardware. Therefore, JTO-Newton, while somewhat slower, is the superior method for LWIR change detection.

### 4.1.6 Single Temperature Dependence

In Section 3.4, three new methods were presented under the guiding principle of joint temperature optimization. These methods were single temperature variance optimization, cross radiance optimization, and closed form. The results for all of these methods are shown
in Figure 4.8. The cross radiance version achieves superior performance at the expense of added computation time. However, it is important to note that the closed form version has the fastest computational time of all of the methods presented in this thesis if the emissivity of time one is already known.

### 4.1.7 All Methods

In this section, all of the best results for each method presented in sections 4.1.2 through 4.1.6 are considered. These are represented in Figure 4.9. The best overall performance is given by the variance adjusted alpha residuals followed close by smoothness-TES with orthogonal projection, variance correction, and mean subtraction. These methods are then followed by the joint temperature optimization and cross radiance optimization. Finally,
the worst performer is the anomalous based change detection method. An important thing to note, smoothness-TES has a lower area under curve (AUC) than the joint based methods for much of ROC curve but has a higher $P_d$ at lower $P_{fa}$ values.

With the improvements outlined in this work, these results are somewhat to be expected. The problem with the joint temperature based methods is that they bring change and non-change targets as close as possible together. This removes a layer of separation that exists in smoothness-TES and alpha residuals. Previously, smoothness-TES and alpha residuals suffered from poorly characterized noise and temperature estimates. However, with the improvements developed in this thesis, these errors are much less prominent and are properly accounted for. In addition, the cross radiance optimization version likely outperforms JTO because cross radiance has fewer degrees of freedom. It relies on only one variable which allows it much less opportunity to overfit the true change cases. In general,
Figure 4.9: All methods on synthetic data

if more unconstrained variables exist, an optimization method has much more opportunity
to overfit and compensate for true changes. Finally, the anomalous based method fails be-
cause it is simply unable to account for temperature differences because it relies on globally
estimated statistics whereas temperatures are locally varying parameters.

4.1.8 Noise Scales

As mentioned previously, 0.5 microflicks of IID Gaussian noise was added to the radiance
values of the synthetic data. This value was chosen because it best represents the value of
noise in SEBASS data. However, it is important to consider how each method performs
with varying degrees of noise. In this test, datasets were generated in accordance with Sec-
tion 4.1.1 with noise varying between $10^{-2}$ and $10^1$ microflicks. Each of the best methods
was then tested on each dataset and recorded the probability of detection at three probabilities of false alarm. These results are displayed in Figures 4.10 through 4.12. These results follow roughly with the results shown in Figure 4.9 with alpha residuals remaining the top performer for all noise levels. One important difference is that smoothness-TES which fails completely when noise goes above 3 microflicks. This result is somewhat unexpected as alpha residuals rely on the same temperature estimates given by the smoothness-TES algorithm in this thesis. One possibility for this difference is because alpha residuals rely on the Wein’s approximation and are therefore much less reliant on accurate temperatures.

In comparison with the expected noise value in real world data, all of the methods except covariance equalization have a reasonable tolerance to noise exceeding the expected value.

![Figure 4.10: Noise stress testing at $P_{fa} = 10^{-1}$](image)
Figure 4.11: Noise stress testing at $P_{fa} = 10^{-2}$

Figure 4.12: Noise stress testing at $P_{fa} = 10^{-3}$
4.2 Measured Data

With all of the algorithms tested for strengths and weaknesses on synthetic data, it is now time to test on real measured data. In this section we present change detection results obtained after applying the change detection algorithms to measured LWIR datacubes from the Spatially Enhanced Broadband Array Spectrograph (SEBASS) sensor. The original data consisted of 128 bands spanning 7.6–13.5 \( \mu m \), but was truncated to \( K = 88 \) bands in the 8.2–12.4 \( \mu m \) range in order to avoid low SNR bands as well as problems stated in Section 3.5. The time-1 and time-2 datacubes were first registered via a manual point selection process followed by Procrustes analysis [21] and transformation. Single band (\( \lambda = 8.3 \mu m \)) images of the time-1 and time-2 scenes are show in Figure 4.13. The change map was then manually generated and is shown in Figure 4.13c.

The two datacubes were collected five days apart and at different times of the day (cube 1 collected August 13th, 18:01 GMT, cube 2 collected August 18th, 19:19 GMT). The atmospheric parameter vectors for each datacube were previously estimated and supplied to the change detection algorithms. Ground surface temperatures ranged from 300 K to 350 K as estimated via smoothness-TES, however true pixel-level temperatures were not available for all pixels.

In Figure 4.15, the ROC curve is shown for the best 5 methods. In this figure, cross radiance, alpha residuals, smoothness-TES, and JTO all have approximately the same performance whereas covariance equalization lags far behind. This does not give a large amount of information about the new methods but it does indicate that the new methodologies presented in this work far outperform the existing methodologies.

In Figure 4.14, the change maps for all of the methods are shown. In Figure 4.14a, the error image for covariance equalization is shown. This figure is nearly completely incapable of finding any true changes and only shows noise pixels within the scene. The other figures all succeed at identifying the true changes within the scene. However, due to other artifacts within the scene, all of the algorithms identify many false changes.
One problem with this dataset is that there is no true change map. The change map was generated manually by examining RGB images of the scene and selecting the pixels which appeared to change. This means that there are likely several errors in the change map. In addition, there are likely to be errors due to subpixel misregistration. Overall, this dataset gives a somewhat poor representation of the potential of each of these algorithms. Future work may explore more datasets to test the algorithms on.
Figure 4.13: Raw radiance data for each time
Figure 4.14: Detection statistic for all methods on real data
Figure 4.15: ROC results for all methods on real data
4.3 Semi-Synthetic Data

As mentioned in the previous section, it is fairly difficult to determine which of the algorithms is best suited for use on real data. Therefore, due to lack of large amounts of real data, it became necessary to construct more data from the existing measured datasets. The semi-synthetic data is created from a $45 \times 40$ m area consisting of a dirt and grass background with four painted panels approximately $6 \times 6$ m each. As the region does not contain any actual changes, synthetic changes were introduced by inserting two rows of four panels into the time-2 data. The inserted pixels were composed of various paints and other man-made materials taken from other regions in the larger time-2 datacube but outside of the change detection region under consideration. By manually inserting change pixels, knowledge of the change locations is known but the scene retains the challenges of measured data. The modified time-2 scene is shown in Figure 4.16c and the change locations are shown in Figure 4.16d.

These results are much more conclusive due to a perfectly defined change mask and low registration errors. In Figure 4.18, JTO beats out the other methods in overall performance. However, all methods retain decent performance throughout the entire ROC curve. In this scenario, Covariance Equalization fails completely. This is likely due to the high ratio of change pixels within the scene. Because anomalous-based methods require a good estimate of the covariance matrices, if there is a significant number of change pixels, a good covariance matrix cannot be estimated. The change maps in Figures 4.17a through 4.17e validate these results with all of the new proposed methods having decent error maps. However, there is a slight misregistration artifact on the middle-left panel which is causing a significant amount of error in all of the new methods.
Figure 4.16: Raw radiance data for each time
Figure 4.17: Detection statistic for all methods on semisynthetic data
Figure 4.18: ROC results for all methods on semisynthetic data
4.4 Timing Results

In this work, five classes of change detection metrics have been proposed drawing upon previous state-of-the-art as well as new novel methods. While it is tempting to simply take the results presented in Figure 4.9 and take the best method, one must consider computational cost as well. Real-world scenarios may require on-board processing of data with real-time data products generated. This means that it is not always possible to process large amounts of data parallelized over thousands of cores on a High Performance Computing (HPC) machine. This section is concerned with the computational time of each method on a reasonable system that could be deployed to a platform.

All methods were run in MATLAB on a quad core 7th gen I7 computer processing unit (CPU) as well as a GTX 1060(6 GB) graphics processing unit (GPU). In the interest of fairness, a reasonable effort has been taken to reduce the time for each method. These timing results do not represent the absolute potential of each method but rather represent a reasonable attempt at optimizing each process with the stated hardware. The fastest obtained results per 100,000 pixels is shown in Table 4.1. (Alpha residuals without the processing time from smoothness-TES is given in parenthesis)

<table>
<thead>
<tr>
<th>Class</th>
<th>Subclass</th>
<th>Preprocessing</th>
<th>On-Board</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anomalous Methods</td>
<td>Chronochrome</td>
<td>0.01</td>
<td>0.10</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>Covariance Equalization</td>
<td>0.01</td>
<td>0.10</td>
<td>0.11</td>
</tr>
<tr>
<td>Smoothness-TES</td>
<td>Orth Projection/Std. Eucl./Mean Subtraction</td>
<td>0.57</td>
<td>1.04</td>
<td>1.61</td>
</tr>
<tr>
<td>Alpha Residuals</td>
<td>Orth Projection/Std. Eucl.</td>
<td>0.58 (0.02)</td>
<td>1.06 (0.50)</td>
<td>1.64 (0.52)</td>
</tr>
<tr>
<td>JTO</td>
<td>JTO-Linear</td>
<td>0</td>
<td>0.11</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>JTO-Newtons</td>
<td>0</td>
<td>0.53</td>
<td>0.53</td>
</tr>
<tr>
<td>Single Temperature</td>
<td>Single Temperature Variance</td>
<td>0</td>
<td>0.40</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>Cross Radiance</td>
<td>0</td>
<td>0.57</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>Closed Form</td>
<td>0.57</td>
<td>0.02</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Table 4.1: Runtime for all methods per 100,000 pixels in seconds

Certain operations are performed faster on a CPU (Eigenvalue Decomposition/Low-
volume Basic Linear Algebra Subprogram (BLAS) operations) whereas other operations are performed faster on a GPU (fast fourier transform (FFT)/High-volume BLAS operations). For example, in the anomalous methods, the covariance matrix can be computed on the GPU whereas the matrix inverse (via Eigenvalue Decomposition) as well as the matrix multiplication is done on the CPU via mtimesx. This balances the processing time and allows the processing time to be brought from ~7 seconds down to 0.11 seconds. In the future, better implementation of MATLAB’s pagefun command may allow this time to be sped up further.

For smoothness-TES, a vectorized minimization solver based on MATLAB’s fminbnd is used. The error computation is done on the GPU whereas the computation of the location of the new point is done on the CPU. This vectorization as well as CPU/GPU optimization brings the temperature estimation per 100,000 pixels from ~360 seconds down to 0.53 seconds. In the future, a more aggressive solver utilizing the derivative of smoothness error may be implemented to reduce this time further. In addition, the orthogonal projection is computed on the GPU utilizing MATLAB’s pagefun command which multithreads every matrix multiplication function out to individual cores bringing the orthogonal projection computation from ~3 seconds down to 0.5 seconds. Alpha residuals utilize the same temperature estimates from smoothness-TES as well as similar operations for orthogonal projection. It is possible to speed up alpha residuals by utilizing less computationally expensive temperature estimation methods at the cost of poorer performance.

The JTO based methods operate fully on the GPU with the matrix inversion and matrix multiplication components implemented via pagefun. This brings the total processing time down from ~3600 seconds for naïve JTO down to 0.53 seconds with JTO-Newtons. Regrettably pagefun’s backslash operator does not support division via non-square matrices. If a true backslash operator is found for the GPU, this could speed up JTO-linear by a factor of 3. JTO-Newtons would remain unchanged as the dampening element is required for Hessian stability.
The single temperature based methods rely on the same minimization procedure utilized by smoothness-TES. This process is unlikely to be sped up outside of a C implementation as the computation of the derivatives for these functions are currently incalculable.

Overall, all methods can likely be sped up by potentially several orders of magnitude. CUBLAS is still a relatively new software and has not had as much time to mature as BLAS particularly with MATLAB’s pagefun implementation. In addition, the atmospheres must be estimated which constitutes an additional processing step further increasing the processing time. With regards to processing time, the closed-form version provides the fastest processing time at the cost of performance. Likewise, JTO-Newton's is able to process data faster than the collection rate while providing results nearly as good as smoothness-TES/alpha residuals.

Of special note, many of the timings in Table 4.1 can be improved by simply reducing the accuracy of the optimization process. The current implementation uses tolerances to obtain the most optimal results. By reducing the tolerances or reducing the total number of iterations the runtime can be sped up by a factor of 2 at the cost of performance.
Conclusions

5.1 Summary

This thesis presented new methodologies for the detection of changes in LWIR hyperspectral imagery. Section 3.1 focused on implementing smoothness-TES for use in change detection. In addition, this section presented two new improvements to smoothness-TES which allow further removal of temperature and noise effects. Section 3.2 presented new methodologies for usage of alpha residuals in change detection. This section presented the computation of the covariance matrix of alpha residuals and proposed a new methodology for its use which allows for a fair comparison of two alpha residuals irregardless of the relative magnitude of the alpha residual. Section 3.3 presented a new methodology based on joint temperature optimization. In this section, two submethods were proposed which attempt to balance processing time and performance. Section 3.4 presented a new approach which manipulates the emissivity to compute the temperature of both times with knowledge of only one time by assuming the emissivity has not changed. This included a true closed-form method assuming one temperature has already been estimated.

Chapter 4 presented the results of all of the proposed methods. Section 4.1 presented all of the methods with purely synthetic data. In this section, the improvements presented in Chapter 3 for alpha residuals and smoothness-TES are validated and reduced the false alarm rate by 3 orders of magnitude over existing anomalous methodologies. Overall, smoothness-TES and alpha residuals had nearly optimal performance over the entire ROC
curve whereas the joint based methods perform nearly as well. All of the new work far out-performed the existing anomalous based methods as the anomalous based methods were incapable of capturing temperature disparities. In sections 4.2 and 4.3, all of the best methods were tested against real and semi-synthetic data. Due to the lack of variety of data and lack of overall data, it was difficult to make any definite conclusions over the performance of each individual algorithm. However, each of the newly proposed algorithms far outperformed the existing state-of-the-art methodologies for VIS-SWIR data. Finally Section 4.4 discussed computational time of each method. Overall, all methods are close to the desired processing time. However, the current implementations JTO-Newton and the closed form method offered competitive performance with one third of the processing time.

5.2 Future Work

Going forward, there are several areas that could be improved upon for this work. Currently, JTO utilizes attempts to minimize the emissivity. In terms of direct performance, alpha residuals slightly outperform smoothness-TES. Therefore, it is reasonable to assume that a proper implementation of JTO with alpha residuals has the potential to outperform the emissivity version.

Currently, all methods operate on a pixel-by-pixel basis and are therefore weak to subpixel misregistration errors. In the future, these methods could be adapted with existing spatial change detection methods utilized for RGB imagery to eliminate some of these subpixel errors. This could be potentially problematic as with LWIR data, the resolution is relatively small therefore a target could reasonably be represented by only one pixel. A spatial method could potentially overlook this single pixel whereas existing methods would mark it as a change.

In order to operate any of the new proposed methods, an atmosphere must be approximated via ISAC or other appropriate method. This atmosphere is not always simple to
compute and may change throughout the scene. In order to compensate for this, a new method based on JTO could optimize for both temperature and atmosphere to attempt to find a change detection statistic. The problem with this method is that the number of atmospheric conditions far outnumber the existing data and this must be adjusted for to allow the system to be appropriately solvable.

Finally, in recent years, large advances have been made to artificial neural networks (ANNs) which allow the processing and creation of new data. Due to the lack of existing data, a generative adversarial network (GAN) could be used to bridge the information gap and allow more test scenarios to be created to fully test all of the methods on real data in order to determine which methods work the best. In addition, an ANN could be trained on synthetic data to determine temperatures/changes and then tested on real input data to quickly and efficiently generate a change detection statistic.
Bibliography


[6] Dekker, Rob J and Schwering, Piet BW and Benoist, Koen W and Pignatti, Stefano and Santini, Federico and Friman, Ola. LWIR hyperspectral change detection for tar-


83


