BAND SELECTION TECHNIQUES FOR
HYPERSPECTRAL CHEMICAL AGENT DETECTION

A Thesis Presented
by
Joseph Stephen Costa
to
The Department of Electrical and Computer Engineering
in partial fulfillment of the requirements
for the degree of
Master of Science
in
Electrical and Computer Engineering

Northeastern University
Boston, Massachusetts

April 2011

This work is sponsored by Defense Threat Reduction Agency under Air Force Contract FA8721-05-C-0002. Opinions, interpretations, conclusions, and recommendations are those of the author and not necessarily endorsed by the United States Government.
Abstract

In this thesis we explore the concept of information redundancy, in the form of additional spectral channels, provided by infrared hyperspectral sensors for chemical detection. The demand for enhanced features including frame rate, coverage area, and spatial resolution of today’s sensors has drastically increased the communications channel bandwidth and processing bandwidth needed to handle the immense amount of information being collected. The goal is to achieve dimensionality reduction allowing for a significant decrease in the data size with little sacrifice to the fidelity of our detection statistics. We focused our efforts on using four linear transformation algorithms, Principal Component Analysis, Smoothing Wiener Filtering, and Forward and Backward Wiener Filtering, to perform feature extraction and explore the information redundancy available in the frequency domain. The first three algorithms showed strong identification of the chemical absorption peaks as significant bands in the data indicating that the algorithms were viable for band selection. We were able to show that some of the bands identified were correlated to the atmospheric absorption bands, indicating that atmospheric correction algorithms may be necessary for improved detection. Using the identified bands to reduce the number of channels necessary for detection proved that a small sacrifice in detection statistics provides for a large reduction in the amount of data required to process.
I would like thank Prof. Vinay K. Ingle and Prof. Bahram Shafai of Northeastern University for reading this thesis and offering valuable suggestions and advice. I would also like to thank Dr. Dimitris G. Manolakis of MIT Lincoln Laboratory for his support in making this thesis possible. I extend my appreciation to Faith Crisley of Northeastern University for helping to coordinate my graduate degree schedule.

I would also like to thank Rob DiPietro and Eric Larrieux of MIT Lincoln Laboratory for passing along their knowledge about hyperspectral detection and for contributing to hours of conversation and debates about the topic. I would specifically like to thank Eric for taking the time to answer my many questions and working through problems with me on a daily basis.

I’m also grateful to my parents Joseph and Deborah Costa for their continued support throughout my life.

Finally, I would like to thank my wife and co-worker Dr. Juliette Costa for her assistance in helping me see this thesis through. Her patience and support has made this thesis possible. She inspires me to be a better person and challenge myself.
Contents

Abstract iii

Acknowledgements iv

1 Introduction 1

2 Background 3

2.1 Infrared Sensors 3

2.2 Infrared Spectroscopy 5

2.3 Radiative Transfer Model 7

2.4 Approximations 12

2.4.1 Thin Plume Approximation 12

2.4.2 Background Emissivity Approximation 12

2.4.3 Linear Planck Function 13

3 Hyperspectral Detection Algorithms 16

3.1 Signal Model 16

3.2 Detection Algorithms 17

3.3 Signal-to-Clutter Ratio 20
4 Linear Spectral Transformations

4.1 Principal Component Analysis .......................... 23
4.2 Smoothing Wiener Filter ............................... 26
4.3 Forward and Backward Wiener Filter ............... 30
4.3.1 Forward (MWIR-LWIR) Wiener Filter .......... 31
4.3.2 Backward (LWIR-MWIR) Wiener Filter .......... 32

5 Chemical Plume Embedding ............................. 37

5.1 Embedding Procedure .................................. 37
5.2 Performance Evaluation of Detection Algorithms .... 39

6 Results ...................................................... 43

6.1 Band Identification (Embedded Data) ............... 44
6.2 Detection and Associated ROC Curves ............... 51

7 Summary ..................................................... 58

A ROC Curves ................................................. 59

B Probability of Detection Tables ....................... 66

Bibliography .................................................. 68
## List of Figures

2.1 Fourier Transform Infrared Spectrometer .......................... 4
2.2 Off Plume Radiance Model ........................................ 8
2.3 Chemical Spectral Signature ...................................... 9
2.4 Planck Curve ....................................................... 10
2.5 On Plume Radiance Model ......................................... 11
2.6 MSE for Linear Approximation of Beer’s Law .................... 13
2.7 Chemical Spectral Signature ...................................... 15

3.1 Transformation from Original to Whitened Space ................. 19
3.2 Detection Examples (Dataset 1 Chemical A) ...................... 21
3.3 Detection Examples (Dataset 2 Chemical A) ...................... 22

4.1 Principal Component Image ........................................ 25
4.2 Principal Component Signal-to-Clutter ............................ 26
4.3 Eigenvector Values Plot .......................................... 27
4.4 Smoothing Wiener Filter Coefficients Matrix .................... 28
4.5 Smoothing Wiener Filter MSE Plot ............................... 29
4.6 Smoothing Wiener Filter Error Image ........................... 30
4.7 Forward Wiener Filter Coefficients Matrix ...................... 32
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.8</td>
<td>Forward Wiener Filter MSE Plot</td>
<td>33</td>
</tr>
<tr>
<td>4.9</td>
<td>Forward Wiener Filter Error Image</td>
<td>34</td>
</tr>
<tr>
<td>4.10</td>
<td>Backward Wiener Filter Coefficients Matrix</td>
<td>35</td>
</tr>
<tr>
<td>4.11</td>
<td>Backward Wiener Filter MSE Plot</td>
<td>36</td>
</tr>
<tr>
<td>4.12</td>
<td>Backward Wiener Filter Error Image</td>
<td>36</td>
</tr>
<tr>
<td>5.1</td>
<td>Atmospheric Transmission</td>
<td>39</td>
</tr>
<tr>
<td>5.2</td>
<td>Embedded Plume Detection (Dataset 1 Chemical A)</td>
<td>41</td>
</tr>
<tr>
<td>5.3</td>
<td>Embedded Plume Detection (Dataset 2 Chemical A)</td>
<td>42</td>
</tr>
<tr>
<td>6.1</td>
<td>Chemical Spectral Signatures A and B</td>
<td>44</td>
</tr>
<tr>
<td>6.2</td>
<td>Principal Component Analysis Signal-to-Clutter Ratio (Datasets 1 &amp; 2 Chemical A)</td>
<td>46</td>
</tr>
<tr>
<td>6.3</td>
<td>Principal Component Analysis (Datasets 1 &amp; 2 Chemical A)</td>
<td>47</td>
</tr>
<tr>
<td>6.4</td>
<td>Smoothing Wiener Filter (Dataset 1 &amp; 2 Chemical A)</td>
<td>48</td>
</tr>
<tr>
<td>6.5</td>
<td>Forward Wiener Filter (Dataset 1 &amp; 2 Chemical A)</td>
<td>49</td>
</tr>
<tr>
<td>6.6</td>
<td>Backward Wiener Filter (Dataset 1 &amp; 2 Chemical A)</td>
<td>50</td>
</tr>
<tr>
<td>6.7</td>
<td>Atmosphere Transmission and Smoothing Wiener Filter MSE</td>
<td>51</td>
</tr>
<tr>
<td>6.8</td>
<td>Bands Selected</td>
<td>53</td>
</tr>
<tr>
<td>6.9</td>
<td>ROC Curves (Dataset 1 Chemical A), PCA and SWF</td>
<td>55</td>
</tr>
<tr>
<td>6.10</td>
<td>ROC Curves (Dataset 1 Chemical A), Forward and Backward Wiener Filter</td>
<td>56</td>
</tr>
<tr>
<td>6.11</td>
<td>Probability of Detection for Constant Probability of False Alarm</td>
<td>57</td>
</tr>
<tr>
<td>A.1</td>
<td>ROC Curves (Dataset 1 Chemical B), PCA and SWF</td>
<td>60</td>
</tr>
<tr>
<td>A.2</td>
<td>ROC Curves (Dataset 1 Chemical B), Forward and Backward Wiener Filter</td>
<td>61</td>
</tr>
</tbody>
</table>
A.3 ROC Curves (Dataset 2 Chemical A), PCA and SWF .......................... 62
A.4 ROC Curves (Dataset 2 Chemical A), Forward and Backward Wiener Filter ................................................................. 63
A.5 ROC Curves (Dataset 2 Chemical B), PCA and SWF ...................... 64
A.6 ROC Curves (Dataset 2 Chemical B), Forward and Backward Wiener Filter ................................................................. 65
List of Tables

2.1 Molecular Degrees of Freedom ........................................... 6

B.1 Probability of Detection Table (Dataset 1 Chemical A), .......... 66
B.2 Probability of Detection Table (Dataset 1 Chemical B) .......... 66
B.3 Probability of Detection Table (Dataset 2 Chemical A) .......... 67
B.4 Probability of Detection Table (Dataset 2 Chemical B) .......... 67
Chapter 1

Introduction

In this thesis we explore the concept of information redundancy, in the form of additional spectral channels, provided by modern infrared hyperspectral sensors for chemical detection. The demand for enhanced features including frame rate, coverage area, and spatial resolution of today’s sensors has drastically increased the communications channel bandwidth and processing bandwidth required to handle the immense amount of information being collected, on the order of hundreds of megabytes per second. We will focus our efforts on exploiting the information redundancy available in the frequency domain in order to intelligently select the spectral bands to be collected allowing for a significant data reduction with little sacrifice to the fidelity of our detection statistics. Other work in this area includes wavelet-based and matching pursuit feature extraction algorithms, Pai-Hui Hsu [6] and spectral binning optimization based on a limited number of target signatures, Charles C. Wamsley, et al [2]. Genetic search algorithms for band selection have also been extensively explored, Chaichoke Vaiphasa et al [1].
In Chapter 2 we introduce a passive infrared (IR) sensor commonly used in chemical detection, the Fourier transform spectrometer (FTS). Next, a brief look at the fundamental physics of IR spectroscopy will be given. An extensive exploration of the radiative transfer model will be conducted, including review of common approximations to the model. In Chapter 3 we derive three common linear detection algorithms, matched filter (MF), adaptive matched filter (AMF), and adaptive coherence/cosine estimator (ACE). In order to provide insight into the nature and importance of the various spectral channels produced by our hyperspectral imagers four linear band selection algorithms will be introduced in Chapter 4 and tested with field collected data. Chapter 5 introduces a chemical embedding procedure for creating synthetic data to test the algorithms under more controlled conditions. In Chapter 6 we threshold the band selection results to create hyperspectral data cubes with a subset of spectral channels. We conclude with a comparison of receiver operating characteristic (ROC) curves produced by the various band selection techniques, and plots of probability of detection for a constant probability of false alarm.
Chapter 2

Background

In this chapter we introduce the Fourier transform spectrometer (FTS), a passive infrared (IR) sensor commonly used in chemical detection. We will review the fundamentals physics of infrared spectroscopy. An extensive exploration of the radiative transfer model will be conducted, including review of common approximations to the model. These approximations will result in a simplified linear model that can be used for development of chemical detection algorithms.

2.1 Infrared Sensors

In this section we describe the Fourier Transform Infrared Spectrometer (FTS) operating in the long wave infrared (LWIR) portion of the electromagnetic spectrum and used to collect the data presented in this thesis.

The particular type of FTS used is known as a Michelson Interferometer invented by Albert Abraham Michelson in 1879. The FTS works by dividing a beam of radiation into two paths. One path travels a fixed length while the other path length is varied by a moving mirror. The light from both paths is then recombined. The optical
path difference (OPD) between the two paths, causes an interference resulting in a change in intensity of the light that can be measured by an optical detector, typically Mercury Cadmium Telluride (MerCad) material for LWIR detection. The wavelength dependent varying of light intensity as a function of OPD can be used to determine the frequency of light passing through the interferometer [4]. A simple block diagram of the FTS can be found in Figure 2.1. The laser and control electronics comprise the metrology that is responsible for keeping track and controlling mirror movement.

Figure 2.1: Fourier Transform Infrared Spectrometer
2.2 Infrared Spectroscopy

Infrared spectroscopy is the measurement of radiant energy from the infrared frequency portion of the electromagnetic spectrum. Frequency, $\tilde{\nu}$, is the number of wave cycles per unit time cycles/second; in SI units, the unit of frequency is the hertz (Hz). Wavelength, $\lambda$, is the physical length of one cycle and is often measured in cm. Frequency and wavelength are inversely related

$$\tilde{\nu} = \frac{c}{\lambda}$$  \hspace{1cm} (2.1)

where $c$ is the speed of light, $3 \times 10^{10}$ cm/sec. The energy of the light is related to the frequency by

$$E = h\tilde{\nu} = \frac{hc}{\lambda}$$  \hspace{1cm} (2.2)

where $h$ is Planck’s constant, $6.6 \times 10^{-34}$ joules-sec. Infrared refers to the frequency range below the lowest frequency of the visible spectrum, below red. The portion of the electromagnetic spectrum from 0.7um to 15um is considered to be the infrared region. This portion of the spectrum is typically further divided into 4 regimes, near-infrared (NIR) 0.7 - 1.4 um, short-wavelength infrared (SWIR) 1.4 - 3 um, mid-wavelength infrared (MWIR) 3 - 5 um, and long-wavelength infrared (LWIR) 8 - 15 um. This paper focuses on detection of chemicals using the LWIR portion of the electromagnetic spectrum.

Infrared spectra are a result of transitions between molecular vibrational energy states. These vibrational energy states are commonly divided into two categories: stretching and bending. A molecule consisting of $n$ atoms has $3n$ degrees of freedom since each of the $n$ atoms moves in 3 dimensions. These degrees of freedom can be
vibrational, rotational, or translational. Linear molecules have only 2 degrees of rotational freedom, due to symmetry about the molecular axis. They have 3 translational degrees of freedom, one for each dimension that can be translated in space, and the remaining of the $3n$ degrees of freedom are vibrational. Nonlinear molecules, which lack axial symmetry, have 3 rotational degrees of freedom, 3 translational degrees of freedom, and the remaining are vibrational [15]. Therefore, the net number of fundamental vibrations for nonlinear and linear molecules is

<table>
<thead>
<tr>
<th>molecules</th>
<th>degrees of freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>nonlinear</td>
<td>$3n-6$</td>
</tr>
<tr>
<td>linear</td>
<td>$3n-5$</td>
</tr>
</tbody>
</table>

Table 2.1: Molecular Degrees of Freedom for molecule with $n$ number of atoms.

The stretching frequency of the atoms of a molecule can be approximated by a simple harmonic motion equation known as Hooke’s law, Equation 2.3.

$$E = \frac{1}{2} kx^2 = \hbar \tilde{\nu}$$

Because we are working at the atomic level, these rules of motion are based on quantum mechanics and therefore these vibrational states are quantized, as described in Equation 2.4, and are referred to as vibrational modes.

$$E = (n + \frac{1}{2})\hbar \tilde{\nu}$$

If a particular vibration leads to a change in dipole moment in the molecule then a transition between vibrational states can be excited by absorption of a photon at the frequency which corresponds to the energy difference between vibrational states, according to Equation 2.2. Because the energy difference between states is highly
specific to the structure of the molecule, scientists can very precisely classify different compounds by their chemical absorption spectra. Passing a monochromatic beam of light through a chemical and studying the amount of energy that was absorbed at various wavelengths, provides information about the chemical’s absorption spectrum.

2.3 Radiative Transfer Model

In this section we will show how at-sensor radiance can be put in a form such that pattern recognition algorithms can be used to detect the chemical of interest [8]. Ideally, the ability to detect should be independent of auxiliary information that is often not available, such as plume concentration and plume temperature. The following will describe how measurements on the plume and off the plume can be combined to result in an expression approximately linear in the library signature, which naturally lends itself to detection using algorithms such as matched filter.

Radiance measured at the sensor is given by the sum of radiances along the sensor line of sight. The background emits radiation depending on its temperature and emissivity. The plume and atmosphere both absorb that radiation in the spectral regions that they have high absorption coefficients. The plume and atmosphere emit radiation at those wavelengths as well. The ability to detect the attenuation due to the plume relies on a temperature difference between the background and the plume. Without a temperature difference the plume would emit exactly as much as it absorbs, leaving only the radiance of the background.

Two-Layer Transfer Model We will first describe the radiance measured by the sensor without the presence of the chemical in the scene. As shown in Figure 2.2, the $L_{\text{off}}(\lambda)$ radiance can be simply described by the radiance of the background, $L_b(\lambda)$,
attenuated by the transmission of the atmosphere, $\tau_p(\lambda)$, plus the radiance of the atmosphere $L_a(\lambda)$ [14].

![Two layer plume free radiative transfer model](image)

**Figure 2.2: Two layer plume free radiative transfer model**

\[
L_{\text{off}}(\lambda) = L_b(\lambda)\tau_a(\lambda) + L_a(\lambda)
\]  

(2.5)

To better understand the model we will further explore the parameters $\tau(\lambda)$ and $L(\lambda)$ respectively. The Bouguer-Lambert-Beer law (typically refereed to as Beer’s law) describes the transmittance, $\tau$, of any material at wavelength $\lambda$. Transmittance is the ratio of emerging radiant power, $I$, to that of the incident power, $I_o$ for any material. In the case of gases, this is an exponential function that is dependent on the absorption of the gas at each wavelength $\tau(\lambda)$, see Figure 2.3, and the concentration pathlength product, $\gamma$, of the gas. The concentration pathlength product is described as the concentration of the gas, $C$, times the length of the gas, $l$ measured along the line of sight of the sensor.

\[
\tau_p(\lambda) = \frac{I(\lambda)}{I_o(\lambda)} = e^{-\gamma(\lambda)}
\]  

(2.6)
The spectral radiance, $L(\lambda)$, of any material can be described by its emissivity, $\epsilon$ times its blackbody emission, $B(\lambda, T)$, at wavelength $\lambda$ and temperature $T$.

$$L(\lambda) = \epsilon(\lambda)B(\lambda, T)$$ (2.7)

If we assume reflectance terms to be minimal we can describe emissivity of a material at any wavelength as

$$\epsilon(\lambda) = 1 - \tau(\lambda).$$ (2.8)

Planck’s law, Equation 2.9, describes the power emitted per unit area, at wavelength $\lambda$, from the surface of a material at temperature $T$, [5]. A sample black body curve
for a 300 Kelvin source is shown in Figure 2.4.

\[
B(\lambda, T) = \frac{C_1}{\lambda^8 [e^{C_2/(\lambda T)} - 1]}
\]  

(2.9)

![Planck Curve](image)

**Figure 2.4: Planck Curve**

We can now proceed to write the radiance as seen by the sensor with the presence of the plume in the scene [14].

\[
L_{on}(\lambda) = L_b(\lambda)\tau_a(\lambda)\tau_p(\lambda) + L_p(\lambda)\tau_a(\lambda) + L_a(\lambda)
\]  

(2.10)

**Three-Layer Transfer Model** The three layer radiative transfer model described in Equation 2.10 is illustrated in Figure 2.5. The ability to detect the plume relies
on the difference in radiance measured on and off the plume

\[ \Delta L(\lambda) = L_{\text{on}}(\lambda) - L_{\text{off}}(\lambda) \] (2.11)

Equation 2.11 can be expanded by substituting in 2.5 and 2.10

\[ \Delta L(\lambda) = L_b(\lambda)\tau_a(\lambda)e^{-\gamma_a(\lambda)} + L_p(\lambda)\tau_a(\lambda) + L_a(\lambda) - L_b(\lambda)\tau_a(\lambda) - L_a(\lambda) \] (2.12)

and simplified to

\[ \Delta L(\lambda) = \tau_a(\lambda)[(\tau_p(\lambda) - 1)B(\lambda, T_p) + L_p(\lambda)]. \] (2.13)

Using Equation 2.7, \( L_b \) and \( L_p \) become

\[ L_b(\lambda) = \epsilon_b(\lambda)B(\lambda, T_b) \] (2.14)

\[ L_p(\lambda) = \epsilon_p(\lambda)B(\lambda, T_p) = (1 - \tau_p(\lambda))B(\lambda, T_p) \] (2.15)
Substituting 2.14 and 2.15 into 2.13 yields our final \( \Delta L \) equation,

\[
\Delta L = \tau_a(\lambda)[(1 - \tau_p(\lambda))(B(\lambda, T_p) - \epsilon_b(\lambda)B(\lambda, T_b))].
\] (2.16)

### 2.4 Approximations

Some common approximations of 2.16 can be made that will further simplify the model and allow us to use linear detection algorithms such as matched filter and adaptive cosine estimation [11].

#### 2.4.1 Thin Plume Approximation

Assuming only optically thin plumes, \( \gamma \ll 1 \), allows for a linear approximation of Beer’s law, 2.6, using the first two terms of it’s Taylor series expansion. The typical error incurred by assuming an optically thin plume can be found in Figure 2.6.

\[
\tau_p(\lambda) = e^{-\gamma a(\lambda)} \approx 1 - \gamma a(\lambda)
\] (2.17)

#### 2.4.2 Background Emissivity Approximation

The next approximation that will simplify 2.16 involves an approximation of \( \epsilon(\lambda) \). If we assume the reflectivity of the background to be small and hence the emissivity to be large, \( \epsilon(\lambda) \approx 1 \) then the resulting \( \Delta L \) equation becomes

\[
\Delta L = \gamma a(\lambda)\tau_a(\lambda)[B(\lambda, T_b) - B(\lambda, T_p)].
\] (2.18)
Typically the chemical plumes thermal contrast relative to the background $|T_b - T_p|$, is assumed to be small, less than 5°C. This assumption allows for a linear approximation of the Planck function about $T_b$,

$$B(\lambda, T_b) - B(\lambda, T_p) \approx \Delta T \frac{\partial B(\lambda, T)}{\partial T} \bigg|_{T=T_b}$$

(2.19)

where $\Delta T = |T_b - T_p|$ and the partial derivative of the Planck function is

$$\frac{\partial B(\lambda, T)}{\partial T} = B(\lambda, T_b) \left( \frac{C_2}{\lambda T^2} \right) (1 - e^{-C_2/\lambda T})^{-1}.$$  

(2.20)
The result is smooth varying functions of wavelength for any two blackbody curves with a small $\Delta T$. The removal of the dependence on $\lambda$ allows for a replacement of Equation 2.20 with a constant, $B_0$.

Using the previous stated approximations we can rewrite Equation 2.16 as

$$\Delta L = \gamma \alpha(\lambda)\tau_a(\lambda)(B_0 \Delta T).$$

(2.21)

The above equation is now linear with respect to the chemical absorption coefficient spectrum $\alpha(\lambda)$, unique to each chemical and more easily matched to a library of chemical spectral signatures [16]. Note that $\gamma$ acts as a linear scaling of the signature; its effect on the signature is shown in Figure 2.7. This scaling has little effect on the algorithms presented in this paper. In order to account for $\tau_a(\lambda)$ each library signature is typically multiplied by an approximation of $\tau_a(\lambda)$ using a computer program such as MODTRAN [3].

In this chapter we reviewed a passive IR sensor commonly used in chemical detection, the Fourier transform spectrometer (FTS). We explored the fundamentals physics of infrared spectroscopy. The radiative transfer model was introduced including the common approximations to the model. These approximations resulted in a simplified linear model that can be used for development of chemical detection algorithms. In the next chapter we will introduce the signal model and its relationship to the radiative transfer model. We will also derive three detection algorithms and explain the signal-to-clutter ratio.
Figure 2.7: The effect of $\gamma$ on a chemical signature
Chapter 3

Hyperspectral Detection Algorithms

In this chapter we will introduce the common form of the signal model and its relationship to the previously derived radiative transfer model. A brief introduction to the generalized likelihood ratio test (GLRT) will be given in order to derive three common detection algorithms, matched filter (MF), adaptive matched filter (AMF), and adaptive coherence/cosine estimator (ACE) [10]. The signal-to-clutter ratio (SCR) will be introduced as a performance metric for the detection results.

3.1 Signal Model

Each spectrum collected by our IR sensor can be described as a random vector $x$, commonly assumed to be normally distributed with mean $\mu$ and covariance matrix $\Sigma$, as specified in Equation 3.1.

$$x \sim N(\mu, \Sigma) \quad (3.1)$$
The probability density function of $x$ is defined as

$$ f(x) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp \left( -\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu) \right) $$

(3.2)

where $p$, is the vector dimensionality.

We will next redefine the off and on plume radiances $L_{off}(\lambda)$ and $L_{on}(\lambda)$, as two hypotheses, $H_0$ and $H_1$ respectively

$$ H_0 : x = v $$

(3.3)

$$ H_1 : x = sa + v $$

(3.4)

where $s$ is the wavelength dependent chemical signature including atmospheric modulation, $v$ is the off plume radiance $L_{off}(\lambda)$ also known as the background radiance, and $a$ is the abundance of the chemical. The relationship between 3.4 and 2.16 is

$$ v(\lambda) = L_{off}(\lambda) $$

(3.5)

$$ s(\lambda) = \tau_a(\lambda)\alpha(\lambda) $$

(3.6)

$$ a = B_0 \Delta T_p \gamma $$

(3.7)

### 3.2 Detection Algorithms

Determining the probability density function of $x$ under previously stated hypothesis $H_0$ and $H_1$, we can express the likelihood ratio test (LRT) as defined in Equation 3.8. The exceedance of some threshold $\eta$ by LRT is considered a detection of the chemical.

$$ \Lambda(x) = \frac{f(x|H_1)}{f(x|H_0)} > \eta $$

(3.8)
It’s important to note that the LRT approach assumes that the statistics, i.e., $\mu$ and $\Sigma$, are explicitly known, which they are not. This leads to a more practical application of LRT which involves using the maximum likelihood estimates for $f(x|H_0)$ and $f(x|H_1)$ producing the generalized likelihood ratio test GLRT, Equation 3.9. The GLRT approach allows us to use statistics estimated from the available data.

$$y_{\text{GLR}} = \frac{((x - \mu)^T \Sigma^{-1} s)^2}{(s^T \Sigma^{-1} s)(\psi_1 + \psi_2 (x - \mu)^T \Sigma^{-1} (x - \mu))}$$ \hspace{2cm} (3.9)$$

The $\psi_1$ and $\psi_2$ are variables that vary depending on the specific form of the GLRT detector being implemented. The three common detectors MF, AMF, and ACE are derived from the GLRT assuming Gaussian statistics. In order to evaluate Equation 3.9 we must first compute $\mu$ and $\Sigma$ from $x$, the data.

$$\mu = \frac{1}{N} \sum_{n=1}^{N} x(n)$$ \hspace{2cm} (3.10)$$

$$\Sigma = \frac{1}{N} \sum_{n=1}^{N} (x_n - \mu)^T (x_n - \mu)$$ \hspace{2cm} (3.11)$$

Evaluating 3.9 by setting $\psi_1 = ((x - \mu) \Sigma^{-1} s)$ and $\psi_2 = 0$ we derive one of the most common detection algorithms, matched filter (MF). MF is simply an orthogonal projection of the data onto the signature.

$$y_{\text{MF}} = \frac{((x - \mu)^T \Sigma^{-1} s)}{(s^T \Sigma^{-1} s)}$$ \hspace{2cm} (3.12)$$

where $\Sigma^{-1} = \Sigma^{-1/2} \Sigma^{-1/2}$ is the inverse of the covariance matrix used to whiten the data and signature, $x$ and $s$. A simplified 2D representation its effect on $x$ is shown in Figure 3.1.
The detection results of MF can be either positive or negative requiring two thresholds to be set for $\eta$. Setting $\psi_1 = 1$ and $\psi_2 = 0$ in Equation 3.9 produces a squared form of MF, adaptive matched filter (AMF). This produces strictly positive detection statistics allowing for a single threshold to be set.

\[
y_{\text{AMF}} = \frac{((x - \mu)^T \Sigma^{-1} s)^2}{(s^T \Sigma^{-1} s)}
\]  

(3.13)

The last detection algorithm we derive from 3.9 is adaptive coherence/cosine estimator (ACE), determined by setting $\psi_1 = 0$ and $\psi_2 = 1$. The result is a detection algorithm that calculates the angle separation between the pixel being tested and the target signature. A small angle implies that the pixel has large similarity to the library signature. A simplified 2D representation of ACE is also shown in Figure 3.1.

\[
y_{\text{ACE}} = \frac{((x - \mu)^T \Sigma^{-1} s)^2}{(s^T \Sigma^{-1} s)(x - \mu)^T \Sigma^{-1} (x - \mu)}
\]  

(3.14)

Sample detection statistic images from the three previously discussed algorithms are

Figure 3.1: Transformation from original to whitened space. MF and ACE 2D representations
displayed in Figure 3.2 and 3.3 for two of our hyperspectral data sets which we will refer to as Dataset A and Dataset B.

### 3.3 Signal-to-Clutter Ratio

The signal-to-clutter ratio (SCR) is a common metric used to determine the performance of the previously mentioned algorithms. The SCR compares target signature to background statistics as described in Equation 3.16.

\[
SCR = a^2 s^T \Sigma^{-1} s
\]  

(3.15)

where \( a = \gamma \Delta T_p \). Because the value of \( a \) is not typically known for real data we use an alternative version of SCR which can be estimated from the detection statistics \([8]\).

\[
\hat{SCR} = \frac{< y_p >^2}{\text{var}[y_{bg}]}
\]  

(3.16)

where \( y_p \) are the detection statistics and \( \text{var}[y_{bg}] \) is the variance of the background pixels. This approach to measuring algorithm performance will be used to derive a SCR for PCA in Chapter 4.

In this chapter we introduced the common form of the signal model and its relationship to the radiative transfer model. A brief introduction to the generalized likelihood ratio test (GLRT) was presented in order to derive the three common detection algorithms, MF, AMF, and ACE. A detection performance metric known as signal-to-clutter ratio (SCR) was presented. In the next chapter we will investigate four linear transformations that will assist us in deciding which bands are the most significant bands to be used for chemical detection.
Figure 3.2: Detection Examples (Dataset 1 Chemical A)
Figure 3.3: Detection Examples (Dataset 2 Chemical A)
Chapter 4

Linear Spectral Transformations

In this chapter four linear spectral transformation methods will be introduced that will provide insight into the nature and importance of the various wavelengths collected by our hyperspectral imagers. These algorithms are Principal Component Analysis (PCA), Smoothing Wiener Filter (SWF), Forward Wiener Filter (FWF), and Backward Wiener Filter (BWF). These algorithms will allow us to explore the spectral redundancy in the data and assist in identifying the bands needed for detection. Each of the band identifying algorithms will be derived and implemented using our field collected hyperspectral data cube.

4.1 Principal Component Analysis

The first of the algorithms is Principal Component Analysis (PCA), one of the most common data analysis methods in modern signal processing. It’s a simple, easily implemented method for extracting relevant information, specifically when there is redundancy in the data [7].

PCA is a mathematical procedure that transforms the data to a new basis that is
a linear combination of the original basis. An orthogonal transformation of the data is performed that results in maximum variance of the data along the first principal component. Each succeeding principal component will have maximum variance in that direction while being orthogonal to the preceding component. The data will be transformed from a spectral space, where each band corresponds to a spectral wavelength, to a principal component space where each dimension corresponds to one of the eigenvectors of the hyperspectral data global covariance.

The covariance of the data is first computed as shown in Equation 3.11. The eigenvalue/eigenvector decomposition of the covariance matrix is given by

$$
\Sigma = QQ^T = \sum_{i=1}^{L} \lambda_i q_i q_i^T
$$

(4.1)

where \( \lambda_i \) is the \( i \)th eigenvalue and corresponds to the \( i \)th eigenvector \( q_i \). The eigenvectors are orthogonal to each other and the eigenvalues are arranged in decreasing order. The eigenvectors indicate the direction of variability of the hyperspectral data, while the magnitude of the corresponding eigenvalue represents the amount of variability in the data. The eigenvectors can now be used to transform the data

$$
y = Q^T x
$$

(4.2)

where \( x = (x_1, x_2, \ldots, x_n) \) is the \( n \)-dimensional hyperspectral data in spectral space and \( y \) is the transformed data, now in principal component space. The eigenvectors are used to provide a linear weighting of the bands of \( x \) as follows

$$
y_i = \sum_{j=1}^{P} q_{ji} x_j
$$

(4.3)
where $y_i$ is the $ith$ principal component and $j$ are band numbers in units of cm$^{-1}$.

A graphical representation of the transform at the 3rd principal component is shown in Figure 4.1. This principal component image clearly displays the presence of the chemical plume. The 3rd principal component was selected by calculating the Singal-to-Clutter Ratio (SCR), defined in Equation 4.4. This form of SCR was derived from Equations 3.16 and 4.1. The SCR of our transformed data is plotted in Figure 4.2, with principal component 3 exhibiting the highest SCR.

$$SCR = \sum_{k=1}^{K} \frac{[q_k^T s]^2}{\lambda_k} \quad (4.4)$$

![Principal Component Image (PC 3)](image)

Figure 4.1: Principal Component Image (PC 3)

Closer inspection of the eigenvector weighting associated with the 3rd principal component plotted against its corresponding wavelength, (Figure 4.3), shows a large correlation to the absorption spectrum of the chemical released during the experiment, (Figure 2.3), specifically 1049 cm$^{-1}$. The chemical released will be referred to as chemical A for the remainder of this discussion.

Using Figure 4.3 we can determine which bands are the largest contributors to
the SCR and therefore most important to keep when creating a hyperspectral data cube with a reduced number of spectral bands.

4.2 Smoothing Wiener Filter

The next approach in investigating the importance of the spectral bands of our data involves attempting to estimate the $i$th band’s radiance based on a linear combination of all the other available bands [9]. The hypothesis is that if the band can be estimated from the other bands with a relatively small error then the additional information provided by that band has only minor importance to the overall spectral signature of the target and therefore can be removed. The error in the estimation of the $i$th band
Figure 4.3: Eigenvector Weight: Eigenvector 3

is determined by

$$e_i = x_i - \hat{x}_i = \sum_{k=1}^{p} a_{ik}x_k = a_i^T x_i, \text{ with } a_i = 1$$ (4.5)

where

$$a_i = [a_1, ..., a_p]^T$$ (4.6)

$$x_i = [x_1, ..., x_p]^T.$$ (4.7)

The variance of the error is given by

$$\text{var}(e_i) = a_i^T \Sigma a_i.$$ (4.8)
The minimization of 4.8, subject to the constraint $a_i = 1$, yields the optimum spectral Smoothing Wiener Filter (SWF). The $p$ filter vectors $a_i, i = 1, \ldots, p$ can be arranged in matrix form, $A = [a_1, \ldots, a_p]$. The error is given as

$$e = A^T x.$$ (4.9)

The SWF $A$ is calculated from the data by

$$A = (\text{diag}(\Sigma^{-1}))^{-1} \Sigma^{-1} = D(1/\sigma_{ii}) \Sigma^{-1}$$ (4.10)

where $\sigma_{ii}$ are the diagonal elements of the inverse covariance matrix, $\Sigma^{-1}$. The coefficients of the SWF from one of the data sets are shown in Figure 4.4. The

![Figure 4.4: Smoothing Wiener Filter Coefficients Matrix](image)

28
minimum error variance or mean squared error (MSE) of the $i$th band is calculated by

$$\sigma_i^2 = 1/\sigma_{ii}. \quad (4.11)$$

By plotting the MSE for each band we can examine the bands that were most difficult to predict and according to our hypothesis the most important to keep when creating a hyperspectral data cube with fewer bands.

![Figure 4.5: Smoothing Wiener Filter MSE](image)

Figure 4.5 clearly identifies $1049 \text{ cm}^{-1}$ as being the most difficult band to predict, again we see a large correlation to the absorption spectrum of chemical A.

Transforming the data using the SWF coefficients, as described in Equation 4.9, allows us to visually inspect the image associated with the $1149 \text{ cm}^{-1}$ band. This image clearly shows the presence of the chemical plume.
We were able to determine bands of interest by performing Smoothing Wiener Filtering the data and looking at the MSE in the prediction of the bands. We found a high correlation among the bands that were most difficult to predict and the absorption peaks of chemical A.

### 4.3 Forward and Backward Wiener Filter

Next we will attempt to estimate the radiance the $i$th band by using a linear combination of a portion of the other bands called predictors [9]. We achieve this by starting at the shorter wavelengths and introducing one new band until we reach the $(i - 1)$ band at the longer wavelengths (MWIR-LWIR). We can also take the reverse approach and start at the longer wavelengths and work towards the shorter wavelengths (LWIR-MWIR), both techniques are presented in this section.
4.3.1 Forward (MWIR-LWIR) Wiener Filter

Forward Wiener Filtering (FWF) is computed by starting at the shorter wavelengths and moving to the longer wavelengths. The error in the linear prediction can be represented by

\[ e_i = x_i - \hat{x}_i = \sum_{k=1}^{i-1} b_k x_k = b^T_i x_i, \text{ with } b_i = 1 \]  

(4.12)

where

\[ b_i = [b_1, ..., b_{i-1}]^T \]  

(4.13)

\[ x_i = [x_1, ..., x_{i-1}]^T \]  

(4.14)

The variance of the error is given by

\[ \text{var}(e_i) = b^T_i \Sigma_i b_i \]  

(4.15)

where \( \Sigma_i \) is the \((i \times i)\) submatrix of \( \Sigma \). According to [12] the rows of \( B = L^{-1} \), where \( L \) is the lower Cholesky decomposition triangular matrix computed from the covariance matrix, are shown in Equation 4.16. The coefficients matrix image \( B \), is displayed in Figure 4.7.

\[ \Sigma = LDL^T \]  

(4.16)

where \( D = \text{diag}[d_1, d_2, ..., d_p] \), the diagonal elements of \( L \). The minimum prediction error variance is determined below and plotted in Figure 4.8 for our data.

\[ \sigma_i^2 = d_i \]  

(4.17)

Figure 4.8 indicates a correlation to chemical A’s absorption peaks. Inspection
of the error image, 4.9, shows that we are able to distinguish the plume from the background; however the statistics are similar to the values along the horizon.

### 4.3.2 Backward (LWIR-MWIR) Wiener Filter

Next we will attempt to estimate the radiance of the $i$th band by adding one new band at a time starting at the longer wavelengths and moving to the shorter wavelengths [9]. This is achieved by performing an upper Cholesky decomposition of the covariance matrix

$$\Sigma = UDU^T. \quad (4.18)$$

The Backward Wiener Filter (BWF) coefficients matrix is displayed in Figure 4.10. We are able to plot the error in our prediction, seen in Figure 4.11. Similar to the
FWF results, there is a correlation to the absorption peaks of chemical A however, a few of the bands above 1100 cm\(^{-1}\) are excessively difficult to predict, and therefore may be important.

Displaying the error image for the 1049 cm\(^{-1}\) band, Figure 4.12, shows that the plume is again visible but not significantly different than a large majority of sky pixels.

Using both the Forward and Backward Wiener Filters we were able to demonstrate a small correlation between band prediction error and the absorption peaks of chemical A. However, a more detailed study of the usefulness of this algorithm using synthetic data is required and will be performed in Chapter 6.

In this chapter four band identifying methods were introduced, Principal Component Analysis, Smoothing Wiener Filter, and Forward and Backward Wiener Filters.
Each of the band identifying algorithms was mathematically derived and implemented using our field collected hyperspectral data cube. The first two algorithms were able to strongly identify the peak absorption feature of the chemical at 1049 cm$^{-1}$. The next two algorithms also identified the major absorption feature, but identified a significant number of additional bands as well. In order to perform a more controlled analysis of the performance of the band identifying algorithms we will introduce a chemical embedding procedure in the next chapter.

Figure 4.9: Forward Wiener Filter Error Image for $\tilde{\nu} = 1149$cm$^{-1}$
Figure 4.10: Backward Wiener Filter Coefficients Matrix
Figure 4.11: Backward Wiener Filter MSE

Figure 4.12: Backward Wiener Filter Error Image for $\tilde{\nu} = 1149\text{cm}^{-1}$
Chapter 5

Chemical Plume Embedding

During experimental chemical releases there are numerous unknowns that can affect the hyperspectral data being collected. Plume characteristics such as concentration, path-length, temperature, size, and location are only a few factors that can significantly alter the performance of our algorithms. A more controlled method of producing hyperspectral data is required. We have developed a plume embedding tool that allows us to modify the chemicals, previously mentioned, physical characteristics and embed that chemical into a hyperspectral data cube.

5.1 Embedding Procedure

The goal is to develop an equation for $L_{on}$ that is a function of $L_{off}$. The $L_{off}$ used will be a chemical-free hyperspectral data cube [11]. This approach is preferred over creating an entirely synthetic cube because chemical free cubes maintain the scene variability encountered in real plume data sets. We begin with our derivation of $L_{on}$ developed from our three layer plume model, restated below for convenience.
$$L_{on}(\lambda) = L_b(\lambda)\tau_a(\lambda)\tau_p(\lambda) + L_p(\lambda)\tau_a(\lambda) + L_a(\lambda) \quad (5.1)$$

Next, we rewrite Equation 2.5, our equation for $L_{off}$.

$$L_b(\lambda)\tau_a = L_{off}(\lambda) - L_a(\lambda) \quad (5.2)$$

Substituting 5.2 into 5.1 provides

$$L_{on}(\lambda) = [L_{off}(\lambda) - L_a(\lambda)]\tau_p(\lambda) + L_p(\lambda)\tau_a(\lambda) + L_a(\lambda) \quad (5.3)$$

Finally, we can write 5.3 in terms of $\tau(\lambda)$ and $B(\lambda, T)$

$$L_{on}(\lambda) = L_{off}(\lambda)\tau_p(\lambda) + B(\lambda, T_a)(1-\tau_a(\lambda))(1-\tau_p(\lambda)) + B(\lambda, T_p)(1-\tau_p(\lambda))\tau_a(\lambda). \quad (5.4)$$

By using Equation 2.6 for $\tau_p(\lambda)$ we will be able to embed various chemicals into the data, $L_{off}$, and adjust the concentration pathlength of the chemical. Also, an arbitrary plume temperature can be input to calculate $B(\lambda, T_p)$. Using a computer program such as MODTRAN enables us to compute an estimate for $\tau_a(\lambda)$, Figure 5.1. An estimate for the atmospheric temperature, $T_a$, can be made from the data to compute $B(\lambda, T_p)$, if we assume the chemical plume is at equilibrium with the surrounding atmosphere.
A commonly used metric in evaluating the performance of detection algorithms is known as the Receiver Operating Characteristic curve or ROC curve [13]. The ROC curve is a plot of probability of detection versus probability of false alarm for threshold $\eta$, $P_D(\eta)$ vs. $P_{FA}(\eta)$. The ability to determine $P_D$ and $P_{FA}$ requires that we explicitly know the spatial extent of the plume. Typically embedded data is used to create ROC curves in order to evaluate the performance of our detection algorithms.

We have chosen a reasonable concentration path length, $\gamma = 100 ppm - m$, and a plume temperature equal to the surrounding atmosphere. These parameters will need to be varied in order to understand the limitations of our algorithms. Figures 5.2 and 5.3 are the result of using MF and ACE detection on the synthetic data set.
created by embedding a chemical plume into Datasets 1 and 2. The corresponding
ROC curves are also shown. Chosen plume location is obvious from the detection
map. As expected, given our choices for $\gamma$ and $T_p$, the probability of detection vs.
probability of false alarm is reasonably high.

In order to perform a more systematic study of the band identifying algorithms
a chemical embedding procedure was introduced. The plume characteristics such as
concentration path-length product, temperature, size, and location are the factors
chosen when performing plume embedding. Two detection algorithms were imple-
mented on our synthetic data and their associated ROC curves were created.

In the next chapter a study of the previously discussed spectral transformation
algorithms, using synthetic data, will be conducted. The tests will be expanded to
include an additional chemical. Detection algorithms will be performed and ROC
curves will be created to assess the performance of the transformation algorithms for
band selection.
Figure 5.2: Embedded Plume Detection (Dataset 1 Chemical A)
Figure 5.3: Embedded Plume Detection (Dataset 2 Chemical A)
Chapter 6

Results

Using the embedding procedure discussed in the previous chapter we can perform a systematic study of the various band selection techniques discussed in Chapter 4. In order to better characterize the robustness of the band selection algorithms we created four synthetic hyperspectral data cubes using the embedding procedure. We used the two previously mentioned datasets, Dataset 1 and Dataset 2 and embed two different chemicals, Chemical A and Chemical B, into the chemical free cubes. In this chapter we will display the band identifying plots, their corresponding error images, and associated histograms from the synthetic data cubes. Next, we will define a procedure for identifying the bands to be used to create various band decimated hyperspectral data cubes. We then implemented two of our standard detection algorithms, MF and ACE and compare the ROC curves generated from using the complete hyperspectral cube to those of the band decimated cubes. Finally, we calculated the probability of detection for a constant probability of false alarm, results are summarized in four x-y plots.
6.1 Band Identification (Embedded Data)

In this section we will review the results of the four band selection algorithms using embedded data created by embedding Chemical A into Dataset 1 and 2. The results of embedding Chemical B into both datasets will be summarized in Section 6.2. We expect that the band identification algorithms will identify the maximum absorption features of the embedded chemical as one some of the most significant bands. Therefore, for ease of reference we have displayed the absorption spectrum of Chemical A and Chemical B in Figure 6.1. We begin our analysis with the aforementioned PCA algorithm. The PCA SCR plots for Datasets 1 and 2 can be found in Figures 6.2(a) and 6.2(b). The eigenvector associated with the principal component with the highest SCR is shown in the upper left of Figure 6.3(a) and 6.3(b) for Datasets 1 and 2,
respectively. The normalized band weighting of the particular eigenvector plot clearly identifies 1045 cm$^{-1}$ as the major band of interest in both datasets, consistent with the peak absorption band for Chemical A. We have also displayed the corresponding principal component images; the embedded plume is clearly identifiable in Datasets 1 and 2.

Next we review the band selection results for the Smoothing Wiener Filter algorithm, Figure 6.4(a) and 6.4(b). The MSE plot, upper left, clearly identifies the maximum absorption feature of Chemical A at 1045 cm$^{-1}$ as one of the most significant bands in each of the embedded datasets. The location of the embedded plume is visible in both images displayed in Figure 6.4.

Figures, 6.5(a), 6.5(b), 6.6(a), and 6.6(b) show the results of band identification using the Forward and Backward Wiener Filtering methods. Both algorithms were able to identify the 1045 cm$^{-1}$ band as significant. However, many additional bands were also determined to be important.
Figure 6.2: Principal Component Analysis Signal-to-Clutter Ratio (Datasets 1 & 2 Chemical A)
Figure 6.3: Principal Component Analysis (Datasets 1 & 2 Chemical A)
Figure 6.4: Smoothing Wiener Filter (Dataset 1 & 2 Chemical A)
(a) Forward Wiener Filter Results (Dataset 1 Chemical A)

(b) Forward Wiener Filter Results (Dataset 2 Chemical A)

Figure 6.5: Forward Wiener Filter (Dataset 1 & 2 Chemical A)
(a) Backward Wiener Filter Results (Dataset 1 Chemical A)

(b) Backward Wiener Filter Results (Dataset 2 Chemical A)

Figure 6.6: Backward Wiener Filter (Dataset 1 & 2 Chemical A)
In addition to detecting the major absorption peak of the embedded chemical, the four band selecting algorithms consistently selected bands above 1100 cm$^{-1}$ to be important. A closer look at this spectral region shows a significant increase in atmospheric absorption, possibly accounting for the large error in band prediction of the algorithms. Figure 6.7 shows the atmospheric absorption plotted with the Smoothing Wiener Filter MSE for Dataset 1 with Chemical A embedded. The results of performing atmospheric correction prior to performing band selection will need to be studied further.

![Atmosphere Transmission and Smoothing Wiener Filter MSE](image)

Figure 6.7: Atmosphere Transmission and Smoothing Wiener Filter MSE.

### 6.2 Detection and Associated ROC Curves

For each band selection algorithm/dataset/chemical combination, we create 5 new hyperspectral cubes, each retaining a different number of bands. Therefore there will be a total of 20 new cubes for each dataset/chemical combination. The bands chosen
to create each new cube will be determined from the results of the previous section. By placing the associated band importance metrics in descending order, and choosing the top 50%, 40%, 30%, 20%, and 10%, the five new cubes for each band selection method were constructed. The top 50% of bands selected by each algorithm using Dataset 1 Chemical A are shown in Figure 6.8. Using MF and ACE detection algorithms we calculated detection statistics and created the associated ROC curves.

Both MF and ACE were used in order to study how each algorithm performed as a function of number of bands used for detection. We also decided to use two datasets taken in drastically different environments, Dataset 1 is a dessert landscape while Dataset 2 is lakeside. The chemicals used, shown in Figure 6.1, allowed us to explore the performance of the band selection algorithms under varying chemical concentrations at different wavelengths. The ROC curves for MF (left) and ACE (right) detection results for Chemical A embedded into Dataset 1 are shown in Figures 6.9 and 6.10. The ROC curves for Dataset 1 Chemical A and Dataset 2 Chemicals A and B can be found in Appendix A.

Reviewing Figures 6.9 and 6.10 it is clear that three of the four band selection algorithms showed promising results. Using PCA to select the most important bands caused only a minor degradation in detection performance using as little as 17 of the initial 68 bands. The SWF and FWF algorithms performed well down to 21 bands, but quickly degraded at 17 bands and below. The BWF was unable to identify the necessary bands for detection.

Next, in order to better compare the band selection algorithms, we arbitrarily select a $P_{FA} = 1e^{-3}$ and determine the $P_D$ from the previously generated ROC curves. The results are provided in table format in Appendix B; the results are also summarized in Figure 6.11 for easier comparison across the multiple datasets and
multiple chemicals.

The plots of Figure 6.11(a) and 6.11(b) show that MF and ACE algorithms have similar performance for Chemicals A and B embedded into Dataset 1. However, ACE was able to provide better detection results for the Dataset 2/Chemical B combination. It is also clear that Chemical A consistently provided better detection statistics for both Datasets.

In addition to the previously discussed work we selected every other band of our original data cube to make a cube with 34 bands. We repeated this procedure down to every 10th band to construct a cube with 7 bands. We then compared the detection results to the ones we got by using PCA for band selection. Due to
the broad absorption peak of Chemical A the every other band method resulted in many samples on the peak absorption feature and therefore high probably of detection results. These detection statistics were superior to the ones acquired when using PCA for band selection. As we continued to decrease the number of bands used for detection the PCA method of selecting the bands provided superior results, specifically when using only 21 or 7 bands.

Of the four band selection algorithms PCA consistently performed the best at identifying important bands, providing only a minor decrease in probability of detection for a significant decrease in number of bands used. The Dataset 2/Chemical A combination showed no change in $P_D$ when the number of bands used for detection was decreased from 68 to 34 at a $P_{FA} = 1e^{-3}$. The average drop in $P_D$ at a $P_{FA} = 1e^{-3}$ using 50% of the bands for all dataset/chemical combinations was 18% for MF and 12% for ACE. SWF and FWF also provided strong results for band identification. The average drop in $P_D$ using 50% of the bands for all dataset/chemical combinations was 34% for MF and 19% for ACE using SWF and 33% for MF and 31% for ACE using the FWF method. The BWF method consistently failed to properly identify the necessary bands for quality detection statistics.
Figure 6.9: MF and ACE ROC curves for Dataset 1 Chemical A using PCA and SWF band selection techniques. Number of bands used, Blue=68, Green=34, Red=28, Cyan=21, Purple=17, and Yellow=7.
Figure 6.10: MF and ACE ROC curves for Dataset 1 Chemical A using the FWF and BWF band selection techniques. Number of bands used, Blue=68, Green=34, Red=28, Cyan=21, Purple=17, and Yellow=7.
Figure 6.11: Probability of Detection for Constant Probability of False Alarm
Chapter 7

Summary

The objective of this work was to explore the concept of information redundancy, in the form of additional spectral channels. Using four linear transformation algorithms for band identification, Principal Component Analysis, Smoothing Wiener Filtering, and Forward and Backward Wiener Filtering we were able to explore the importance of the various spectral bands first from actual field collected data and then more controlled synthetic data. The band identifying algorithms showed promising results, specifically strongly identifying the peak absorption bands of the in-scene chemical. Through thresholding of the identified bands decimated hyperspectral data cubes were created. Performing detection using MF and ACE and subsequent ROC curve generation we were able to demonstrate varying levels of impacted performance for a significant reduction in the number of spectral bands used.
Appendix A

ROC Curves
Figure A.1: MF and ACE ROC curves for Dataset 1 Chemical B using PCA and SWF band selection techniques. Number of bands used, Blue=68, Green=34, Red=28, Cyan=21, Purple=17, and Yellow=7.
Figure A.2: MF and ACE ROC curves for Dataset 1 Chemical B using the FWF and BWF band selection techniques. Number of bands used, Blue=68, Green=34, Red=28, Cyan=21, Purple=17, and Yellow=7.
Figure A.3: MF and ACE ROC curves for Dataset 2 Chemical A using PCA and SWF band selection techniques. Number of bands used, Blue=68, Green=34, Red=28, Cyan=21, Purple=17, and Yellow=7.
Figure A.4: MF and ACE ROC curves for Dataset 2 Chemical A using the FWF and BWF band selection techniques. Number of bands used, Blue=68, Green=34, Red=28, Cyan=21, Purple=17, and Yellow=7.
Figure A.5: MF and ACE ROC curves for Dataset 2 Chemical B using PCA and SWF band selection techniques. Number of bands used, Blue=68, Green=34, Red=28, Cyan=21, Purple=17, and Yellow=7.
Figure A.6: MF and ACE ROC curves for Dataset 2 Chemical B using the FWF and BWF band selection technique. Number of bands used, Blue=68, Green=34, Red=28, Cyan=21, Purple=17, and Yellow=7.
Appendix B

Probability of Detection Tables

<table>
<thead>
<tr>
<th></th>
<th>MF</th>
<th>ACE</th>
<th>MF</th>
<th>ACE</th>
<th>MF</th>
<th>ACE</th>
<th>MF</th>
<th>ACE</th>
<th>MF</th>
<th>ACE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bands</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>0.67</td>
<td>0.70</td>
<td>0.67</td>
<td>0.70</td>
<td>0.67</td>
<td>0.70</td>
<td>0.67</td>
<td>0.70</td>
<td>0.67</td>
<td>0.70</td>
</tr>
<tr>
<td>34</td>
<td>0.63</td>
<td>0.63</td>
<td>0.61</td>
<td>0.62</td>
<td>0.59</td>
<td>0.59</td>
<td>0.03</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>0.61</td>
<td>0.61</td>
<td>0.60</td>
<td>0.59</td>
<td>0.56</td>
<td>0.55</td>
<td>0.03</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>0.59</td>
<td>0.57</td>
<td>0.56</td>
<td>0.51</td>
<td>0.52</td>
<td>0.48</td>
<td>0.02</td>
<td>0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.58</td>
<td>0.52</td>
<td>0.44</td>
<td>0.39</td>
<td>0.18</td>
<td>0.16</td>
<td>0.01</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.27</td>
<td>0.12</td>
<td>0.29</td>
<td>0.15</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table B.1: Probability of Detection Table (Dataset 1 Chemical A), $P_{FA} = 1e^{-3}$

<table>
<thead>
<tr>
<th></th>
<th>MF</th>
<th>ACE</th>
<th>MF</th>
<th>ACE</th>
<th>MF</th>
<th>ACE</th>
<th>MF</th>
<th>ACE</th>
<th>MF</th>
<th>ACE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bands</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>0.11</td>
<td>0.10</td>
<td>0.11</td>
<td>0.10</td>
<td>0.11</td>
<td>0.10</td>
<td>0.11</td>
<td>0.10</td>
<td>0.11</td>
<td>0.10</td>
</tr>
<tr>
<td>34</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.07</td>
<td>0.09</td>
<td>0.08</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>0.09</td>
<td>0.07</td>
<td>0.07</td>
<td>0.07</td>
<td>0.09</td>
<td>0.06</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>0.08</td>
<td>0.06</td>
<td>0.03</td>
<td>0.03</td>
<td>0.07</td>
<td>0.05</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.06</td>
<td>0.03</td>
<td>0.02</td>
<td>0.01</td>
<td>0.03</td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.04</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table B.2: Probability of Detection Table (Dataset 1 Chemical B), $P_{FA} = 1e^{-3}$
<table>
<thead>
<tr>
<th>Bands</th>
<th>PCA MF</th>
<th>PCA ACE</th>
<th>SWF MF</th>
<th>SWF ACE</th>
<th>FWF MF</th>
<th>FWF ACE</th>
<th>BWF MF</th>
<th>BWF ACE</th>
</tr>
</thead>
<tbody>
<tr>
<td>68</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
</tr>
<tr>
<td>34</td>
<td>0.97</td>
<td>0.97</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>28</td>
<td>0.94</td>
<td>0.96</td>
<td>0.94</td>
<td>0.95</td>
<td>0.95</td>
<td>0.95</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>21</td>
<td>0.92</td>
<td>0.95</td>
<td>0.92</td>
<td>0.95</td>
<td>0.95</td>
<td>0.94</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>17</td>
<td>0.89</td>
<td>0.71</td>
<td>0.68</td>
<td>0.71</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>7</td>
<td>0.63</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table B.3: Probability of Detection Table (Dataset 2 Chemical A), $P_{FA} = 1e^{-3}$

<table>
<thead>
<tr>
<th>Bands</th>
<th>PCA MF</th>
<th>PCA ACE</th>
<th>SWF MF</th>
<th>SWF ACE</th>
<th>FWF MF</th>
<th>FWF ACE</th>
<th>FWF MF</th>
<th>FWF ACE</th>
</tr>
</thead>
<tbody>
<tr>
<td>68</td>
<td>0.54</td>
<td>0.85</td>
<td>0.54</td>
<td>0.85</td>
<td>0.54</td>
<td>0.85</td>
<td>0.54</td>
<td>0.85</td>
</tr>
<tr>
<td>34</td>
<td>0.34</td>
<td>0.73</td>
<td>0.01</td>
<td>0.56</td>
<td>0.01</td>
<td>0.13</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>28</td>
<td>0.14</td>
<td>0.62</td>
<td>0.00</td>
<td>0.31</td>
<td>0.00</td>
<td>0.09</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>21</td>
<td>0.11</td>
<td>0.49</td>
<td>0.00</td>
<td>0.03</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>17</td>
<td>0.05</td>
<td>0.28</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>7</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table B.4: Probability of Detection Table (Dataset 2 Chemical B), $P_{FA} = 1e^{-3}$
Bibliography


