PHYSICS-BASED DETECTORS APPLIED TO LONG-WAVE INFRARED HYPERSONTICAL DATA

Joshua Broadwater, Rama Chellappa
University of Maryland
Center for Automation Research
College Park, MD 20742

ABSTRACT

Long-wave infrared (LWIR) hyperspectral image (HSI) data presents an interesting challenge for automatic target detection algorithms. LWIR HSI data is useful for both day and night operations, but weak signatures like disturbed soil can be problematic for standard matched-filter techniques (Bowman, et al. 1998). In this paper, we augment the standard matched-filter techniques with physics-based information particular to HSI data. Our results show that these physics-based detectors provide improved detection performance with quick processing times.

1. INTRODUCTION

LWIR HSI data provide a useful sensor for the Army’s Future Combat Systems. LWIR HSI data offer the capability to conduct image collection in both daylight and nighttime operations. Additionally, LWIR HSI data show the capability to detect disturbed soil. The raw imagery, however, may not be the most suitable format for the operator. Instead, Automatic Target Recognition (ATR) algorithms need to be employed so the war fighter can quickly make sense of the scene and identify possible targets or threats.

Because of this need, a number of ATR algorithms have been developed over the years. In some of the earliest work for LWIR HSI data, Bowman, et al. made some significant findings about detection in the LWIR region (Bowman, et al., 1998). Their paper divides the problem into strong targets (surface mines) and weak targets (buried mines). For surface mines, they found that standard ATR algorithms such as spectral matched filtering performed well. The main difficulty was the generation of emissivity signatures for LWIR data when given an emissivity signature. Therefore, they recommended anomaly detection approaches like RX (Reed and Yu, 1990). For buried mines, they noted that matched filter techniques performed poorly because of the marginal separation between the disturbed soil signature and mixed background pixels.

A few other researchers have worked on LWIR HSI data since Bowman’s paper in 1998. Winter is probably the most notable of these extending his anomaly detection method with the N-FINDR algorithm (Winter, 1999) to adaptively identify target-like signatures in the imagery (Winter, 2004). Healey has also extended his invariant subspace work (Healey and Slater, 1999) to the LWIR regions in an attempt to overcome the problems found with generation of target radiance signatures.

For this paper, we focus on detecting weak targets in LWIR HSI data. It is these weak targets that Bowman and Lucey argued were problematic for standard matched-filter techniques. To overcome this issue, we augment these techniques with physics-based information. The key to our approach is the linear mixing model and its associated parameters (Hapke, 1993).

The linear mixing model assumes a pixel is made up of endmembers, each with its own abundance. Endmembers are the spectra representing the unique materials in a given image. For instance, in an image that contains dirt, grass, and road, the endmembers would be the corresponding unique spectral signatures for each of these materials. Abundances are the percentage of the area of each material within a given pixel. Mathematically, these concepts are expressed as

\[ x = Ea + n, \quad a_i \geq 0 \forall i, \quad \sum_{i=1}^{M} a_i = 1 \]  

where \( x \) is an \( D \times 1 \) vector that represents the spectral signature of the current pixel, \( M \) is the number of endmembers within the image, \( E \) is an \( D \times M \) matrix where each column represents the \( i^{th} \) endmember, \( a \) is an \( M \times 1 \) vector where the \( i^{th} \) entry represents the abundance value \( a_i \), and \( n \) is noise typically modeled as a normal distribution.

Nearly all hyperspectral detectors use the linear mixing model as the basis for their hypothesis test. The difference we are promoting is in the estimation of the linear mixing model parameters – namely, the endmembers and their abundances. Standard techniques like AMSD do not model the endmembers in a physical fashion (Manolakis, Siracusa, and Shaw, 2001). Instead, the endmembers and abundances are simply vectors and magnitudes spanning an abstract mathematical subspace.

For our physics-based methods, we use algorithms that extract physically meaningful endmembers that represent materials like grass and sand. Because we use
Physics-Based Detectors Applied To Long-Wave Infrared Hyperspectral Data

See also ADM002075., The original document contains color images.
physically meaningful endmembers, the associated abundances also have a physical interpretation. Namely, the abundances measure the amount of endmember material within a given pixel. This approach is fundamentally different from the work that Bowman originally did with matched-filters. Taking this physics-based approach, we will show significantly improved detector performance with faster processing times.

The rest of this paper is as follows. Section 2 describes the classical detection algorithms used in our analysis. This includes derivation of each detector and how the endmembers and abundances are estimated. Section 3 describes the physics-based detectors and how they differ specifically from their classical counterparts. Section 4 details the experimental results. Section 5 concludes our paper and makes recommendations for further study.

2. CLASSICAL DETECTION ALGORITHMS

This section identifies the classical detection algorithms used in our analysis. This is not meant to be a totally inclusive list. It only identifies the two main types of statistical algorithms: structured and unstructured. The unstructured detector is the ACE algorithm which models the background as a multivariate normal distribution (Kraut, Scharf, and McWhorter, 2001). The structured detector is the AMSD algorithm which models the background as an eigenvalue decomposition of the image correlation matrix (Manolakis, Siracusa, and Shaw, 2001). Both detectors are based entirely on statistical inference techniques.

2.1. Adaptive Coherent Estimate (ACE)

The ACE detector is one of the most powerful detectors available for HSI processing, yet it has one of the simplest models. The ACE algorithm uses only a statistical distribution (namely the multivariate normal distribution) to model the background. The resulting hypothesis test is

\[ H_0 : x = n \]
\[ H_1 : x = Sa + n \]  

(2)

where \( x \) is the pixel under test, \( S \) are the target endmembers, \( a \) are the target abundances, and \( n \) is a zero-mean multivariate normal distribution with covariance \( \sigma^2 \Gamma \). Note the interesting scale factor \( \sigma^2 \) on the covariance matrix. The scaling term is introduced to make the detector scale-invariant even though this assumption (like the normality assumption) is not based on empirical data. Since the algorithm does not use any background endmembers, the sum-to-one and non-negativity constraints of (1) cannot be met as they require a background subspace. Despite these simplifying assumptions, the ACE detector performs well on HSI data (Manolakis and Shaw, 2002).

For this derivation, we follow the work by Kraut and Scharf (Kraut and Scharf, 1999) based on the work of Kelly (Kelly, 1986). Besides the information we have in (2), we also assume that we have an independent data set \( Y \) such that

\[ Y = \{ y_i | y_i = N(0, \Gamma), i = 1, \ldots, N \} \].  

(3)

Combining (2) and (3) provides the joint likelihood equation under the null hypothesis

\[
L(x, Y | H_0) = (2\pi)^{\frac{1}{2}(N + 1)} | \Gamma |^{\frac{1}{2}(N + 1)} (\sigma^2)^{\frac{1}{2}L} \times \exp \left\{ -\frac{1}{2\sigma^2} x^T \Gamma^{-1} x - \frac{1}{2} \sum_{i=1}^{N} y_i^T \Gamma^{-1} y_i \right\}
\]

(4)

and the joint likelihood equation under the alternate hypothesis

\[
L(x, Y | H_1) = (2\pi)^{\frac{1}{2}(N + 1)} | \Gamma |^{\frac{1}{2}(N + 1)} (\sigma^2)^{\frac{1}{2}L} \times \exp \left\{ -\frac{1}{2\sigma^2} (x - Sa)^T \Gamma^{-1} (x - Sa) \right\} \times \exp \left\{ -\frac{1}{2} \sum_{i=1}^{N} y_i^T \Gamma^{-1} y_i \right\}
\]

(5)

If we assume that \( N \) is very large, the covariance estimate from these likelihoods can be simplified to

\[
\hat{\Gamma} = \sum_{i=1}^{N} y_i y_i^T
\]

(6)

which is a standard assumption made in the literature. Note that under this assumption, the covariance under the null hypothesis and alternate hypothesis are equal and greatly simplify the mathematics.

We use the likelihood equations in (4) and (5) to find the maximum likelihood estimates (MLEs) of the unknown abundance and variance terms. These estimates are substituted back into the original likelihood equations. The updated likelihoods are taken as a ratio to obtain the generalized likelihood ratio test (GLRT) statistic. After some algebra and simplification, the ACE detector is

\[
D_{\text{ACE}}(x) = \frac{x^T \hat{\Gamma}^{-1} S(\hat{\Gamma}^{-1} S)^{-1} S^T \hat{\Gamma}^{-1} \hat{\Gamma}_x}{x^T \hat{\Gamma}_x}
\]

(7)

2.2. Adaptive Matched Subspace Detector

The AMSD algorithm is fundamentally different from the ACE detector. This detector uses the linear
mixing model directly in its hypothesis modeling both target and background endmembers. The hypothesis used to generate this particular detector is

\[ H_0 : x = Ba_b + n \]
\[ H_1 : x = Sa_a + Ba_b + n \]  \hspace{1cm} (8)

where \( x \) is the pixel under test, \( B \) are the background endmembers, \( S \) are the target endmembers, \( a_b \) are the background abundances, \( a_a \) are the target abundances, and \( n \) is a zero-mean multivariate normal distribution with covariance \( \sigma^2 I \).

To begin the AMSD derivation, we first calculate the likelihood equations for the null hypothesis

\[ L(x | H_0) = (2\pi\sigma^2)^{-\frac{n}{2}} \times \exp\left\{-\frac{1}{2\sigma^2} (x - Ba_b)^T (x - Ba_b) \right\} \]  \hspace{1cm} (9)

and the alternate hypothesis

\[ L(x | H_1) = (2\pi\sigma^2)^{-\frac{n}{2}} \times \exp\left\{-\frac{1}{2\sigma^2} (x - Ea)^T (x - Ea) \right\}. \]  \hspace{1cm} (10)

Using the above likelihood equations, the maximum likelihood estimates of the abundance and variance terms can be found as was done with ACE. These are substituted back into the likelihood equations. Having calculated the likelihoods for each hypothesis and using some simple algebra, the GLRT takes the ratio of the two to calculate the AMSD detection statistic

\[ D_{AMSD}(x) = \frac{x^T (I - B(B^T B)^{-1} B^T) x}{x^T (I - E(E^T E)^{-1} E^T) x} = \frac{x^T P_{a_b} x}{x^T P_{a} x}. \]  \hspace{1cm} (11)

In the original work (Manolakis, Siracusa, and Shaw 2001), they decided to use a fully statistical approach when implementing the AMSD algorithm. Instead of using physically meaningful endmembers to construct the background \( B \), they chose to use an eigenvalue decomposition of the image correlation matrix. This is an intelligent choice as it provides the minimal least squared error between the model and the real image. Unfortunately, it has to forego any physical meaning to do this. The endmembers are no longer endmembers that identify materials. They are only vectors that span a subspace of the background and thus have no physical meaning. Therefore, their corresponding abundances also have no physical meaning and hence cannot conform to the sum-to-one and non-negativity constraints.

3. PHYSICS-BASED DETECTORS

These detectors use the underlying physics of the linear mixing model in their solutions. The simplest is the Fully Constrained Least Squares (FCLS) algorithm which estimates the abundances with all their constraints when given physically meaningful endmembers. The next algorithm is the AMSD algorithm that uses physically meaningful endmembers. The final detector is a culmination of the AMSD and FCLS algorithms.

3.1. Fully Constrained Least Squares (FCLS) Detector

The FCLS algorithm directly estimates the abundances in (1). This makes this algorithm more an estimator than a detector; however, the abundance estimates it produces can be used for detection purposes. While other algorithms had been developed that handle both the non-negativity and sum-to-one constraints (Boardman, 1990), (Settle and Drake, 1993), (Ashton and Schaum, 1998), these algorithms tended to be very computationally intensive – especially as the number of endmembers increased. The FCLS algorithm was developed to meet both abundance constraints as well, but in a very efficient manner (Heinz and Chang, 2001). Additionally, the FCLS method is optimal in terms of least squares error. Because of these reasons, we chose to use it in our algorithms. Unfortunately, FCLS does not allow a closed-form mathematical solution due to the non-negativity constraints. Instead, a numerical solution is required.

To calculate the FCLS solution, we begin with the non-negativity constraints. The idea is to minimize the least squares error (LSE) by estimating the non-negative abundance values. Mathematically this is expressed as

\[ \min_{a} (x - Ea)^T (x - Ea), a_i \geq 0 \forall i \]  \hspace{1cm} (18)

where \( E \) is the concatenation of the target \( S \) and background \( B \) signatures. Using Lagrange multipliers, a new loss function \( J \) is developed where

\[ J = \frac{1}{2} (x - Ea)^T (x - Ea) + \lambda^T (a - c), \]  \hspace{1cm} (19)

\( a = c, \) and each member of the unknown constant \( M \times 1 \) vector \( c \) is non-negative to enforce the non-negativity constraint. This construction allows the use of Lagrange multipliers because the non-negativity constraints have been substituted by equality constraints with the unknown vector \( c \). To calculate the estimate of \( a \), we take the partial derivative of \( J \) with respect to \( a \) to obtain

\[ \frac{\partial J}{\partial a} = E^T E\dot{a} - E^T x + \lambda = 0. \]  \hspace{1cm} (20)
Equation (20) contains two unknowns: the abundance estimates and the Lagrange multipliers. Solving for these unknown results in

$$\hat{a} = (E^T E)^{-1} E^T x - (E^T E)^{-1} \lambda$$

(21)

and

$$\lambda = E^T (x - E\hat{a}).$$

(22)

Iterating (21) and (22) provides the numerical solution for the non-negativity constraints.

To handle the sum-to-one constraints, an easy modification of the aforementioned algorithm was developed to retain the optimality guaranteed under the Kuhn-Tucker conditions (Haskell and Hansen, 1981). In the modification, the endmember matrix and pixel signatures are extended such that

$$\tilde{E} = \begin{bmatrix} \delta E^T \\ 1^T \end{bmatrix}$$

(23)

is the new endmember matrix and

$$\tilde{x} = \begin{bmatrix} \delta x \\ 1 \end{bmatrix}$$

(24)

is the new pixel signature where $\delta$ is a small number (typically $1 \times 10^{-5}$). The $\delta$ variable controls how tightly the solution will sum to one so that smaller values provide a better solution, but may need longer convergence time. The new endmember matrix and pixel signature are then used in (21) and (22) to obtain an abundance solution that meets both the non-negativity and sum-to-one constraints.

### 3.2. AMSD/IEA

This detector is the same detector as AMSD except the background endmembers have been extracted using the Iterative Error Analysis (IEA) algorithm (Neville, et al., 1999). The IEA algorithm extracts physically meaningful endmembers that are based on minimizing the least squared error using all abundance constraints. The algorithm begins with the target signature and unmixes the image using the FCLS algorithm. An error image is then created between the original image and the unmixed image. The pixel that contains the most error in a least squared error sense is then chosen as the next endmember. This continues until $N$ number of endmembers is extracted.

The IEA algorithm has a few advantages. First, the algorithm produces physically meaningful endmembers that are well matched to the FCLS algorithm. Second, the algorithm provides endmembers that are significantly different from the target signature minimizing the chance of background signatures “bleeding” into the target subspace. Third, the algorithm runs quickly taking only a few minutes to extract 30 endmembers. Fourth, the IEA algorithm was identified as one of the best performing endmember extraction techniques (Plaza, et al., 2004).

### 3.3. AMSD/IEA/FCLS

The last detector is one we have presented as the hybrid detector (Broadwater, Meth, and Chellappa, 2004). This detector is a culmination of FCLS, IEA, and AMSD. Besides using IEA to generate background endmembers, this detector uses the FCLS abundance estimates directly in the detector. In other AMSD versions, the abundances have been calculated using a maximum likelihood estimate which may not meet the sum-to-one and non-negativity constraints identified in (1).

To utilize the FCLS abundance estimate, we must redefine the AMSD detection statistic. In this derivation, we simply replace the MLE abundance estimates $a$ with their FCLS counterparts. If we do this, the new hybrid detection statistic is

$$D_{HYBRID}(x) = \frac{(x - B\hat{a}_x)^T (x - B\hat{a}_x)}{(x - E\hat{a})^T (x - E\hat{a})}$$

(27)

where the abundance estimates are taken from the FCLS solution.

### 4. LWIR HSI EXPERIMENTS

This section applies the aforementioned detectors to real world collected LWIR HSI data. The section is broken down into two parts. The first part describes the experiments and how they were implemented. The second part contains the experimental results and our conclusions.

#### 4.1. Experimental Design

The imagery used in this analysis comes from the U.S. Army Night Vision Electronic Sensors Division (RDECOM CERDEC NVESD). Over a number of experiments they collected LWIR HSI data using the AHI sensor (Lucey, et al., 1998). We are using two images taken at a U.S. Army test site. The images are of desert scenes with dry brush and vegetation. Additionally, the images contain dry stream beds with weak disturbed soil signatures. These signatures are older and weaker than the target signatures, but nonetheless present an additional challenge for the target detection algorithms.

The targets for the test are disturbed soil signatures. We used disturbed soil because it was identified in Bowman’s work as some of the hardest targets to detect in LWIR HSI data. We also used it to show that matched-
filters based on physical information can separate such targets from similar background signatures.

As for the target exemplars used in our detectors, we obtained them directly from the imagery. Ten target signatures were pulled from each image. The mean of these ten signatures was used as the target exemplar. Because we extracted our target exemplar this way, we did not have to worry about atmospheric compensation algorithms that could induce artifacts. While obtaining our target signatures this way is not ideal from a real-world operational perspective, it does provide the most unbiased way to analyze the detectors because we know the targets signatures are perfectly matched to the imagery.

We were able to locate ten target signatures due to the pixel-level ground truth we received. For each image, there is a target mask that identifies which pixels contain which types of targets. This allows us to measure the performance of the algorithms on each individual pixel and generate statistically meaningful Receiver Operating Characteristic (ROC) curves.

The last and probably most important parameter for our experiments is the estimation of the background. In Section 2 and 3, each algorithm uses a different estimate of the background. How those estimates are obtained though needs to be described. For instance, the ACE detector uses a covariance matrix to model the background, but no indications are given for how to obtain that covariance matrix. This is what the following paragraphs will detail for each detector. In all cases, we used the background estimate that provided the best performance for each detector as will be explained.

For ACE, we used both a global and local model of the covariance. The global model estimated the covariance matrix of the image using all pixels except those that contained targets. This approach makes sure the covariance matrix is a true estimate of the background and not tainted by any target signatures. The local model used a sliding window that was 21 pixels squared with the middle 11 pixels removed. We chose this window because it was large enough to encompass a target, but small enough that it fit in between multiple targets. Again, this is an ideal solution as we would not know the best size to make this window in real-world operations.

For AMSD, we used a global correlation matrix as defined in the original paper. As with the global ACE detector, we removed all target signatures from the image before generating the correlation matrix. This guaranteed that the correlation matrix was a pure estimate of the background. Eigenvalue decomposition was performed on the correlation matrix. The 30 most significant eigenvectors were extracted to model the background according to the magnitude of their respective eigenvalues. The AMSD algorithm was then run 30 times. Each time the number of endmembers was increased from a single eigenvector to all 30 eigenvectors. Thirty ROC curves were generated and the ROC curve with the best performance was identified. This curve is the one shown in the experimental results section.

For the physics-based detectors, IEA was used to model the background. Thirty endmembers were extracted starting with the endmembers that differed significantly from the target. Each of the three physics-based detectors was run 30 times in the same fashion as AMSD. Again, 30 ROC curves were generated for each detector and the ROC curve that provided the best result was used. Table 1 provides a summary of all the background models used for each detector.

<table>
<thead>
<tr>
<th>Detector</th>
<th>Type</th>
<th>Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global ACE</td>
<td>Statistical</td>
<td>Global Covariance</td>
</tr>
<tr>
<td>Local ACE</td>
<td>Statistical</td>
<td>Local Covariance</td>
</tr>
<tr>
<td>AMSD</td>
<td>Statistical</td>
<td>Eigenvectors</td>
</tr>
<tr>
<td>FCLS</td>
<td>Physics</td>
<td>IEA</td>
</tr>
<tr>
<td>AMSD/IEA</td>
<td>Physics</td>
<td>IEA</td>
</tr>
<tr>
<td>AMSD/IEA/FCLS</td>
<td>Physics</td>
<td>IEA</td>
</tr>
</tbody>
</table>

4.2. Experimental Results

Fig. 1 and Fig. 2 show the detection images generated by the six algorithms for each LWIR HSI image. The images show the raw detection scores with no threshold applied. White x symbols mark the true location of the targets in each set of images. The images are color coded such that red is target-like while blue is background.

Fig. 3 and Fig. 4 show the ROC curves for the detectors applied to each LWIR HSI image. The dashed lines represent the classical detection algorithms. The solid lines represent the physics-based algorithms. Because we have pixel level ground truth, the ROC curves are calculated for the probability of detection and probability of false alarm. Therefore, the axes are scaled from zero to one. While this information does not provide the number of false alarms per unit area, it does provide a good way to compare the detectors’ performances.
From these figures, we first note that the global ACE does not perform well. The detection algorithm uses a global covariance matrix to model the background. While the covariance matrix was estimated from background pixels only, the background pixels include the dry stream beds. These dry stream beds have signatures very close to the desired target signature. Because of this and the way ACE works, the global ACE detector ends up suppressing both the background and target signatures. This is why few targets are seen in the detection images and leads to the poor ROC performance.

The other two statistical based methods, local ACE and AMSD, perform better than expected. The AMSD algorithm uses the eigenvectors of the background correlation matrix to model the background. In Image 1 where large areas of undisturbed soil are present, this provides enough of a difference to push the eigenvectors farther away from the target signatures. The net effect of this is improved performance. However, in Image 2, the image is comprised mostly of dry stream beds which are very similar to the target. These signatures push the eigenvectors closer to the target and result in slightly worse detection performance.

An interesting side note of the local ACE detector is its implicit spatial filtering. Because the sliding window is sized to fit around a target, the ACE detector will not detect targets that are larger than the window size. This is the proverbial double edged sword. If you size the sliding window correctly, the local ACE algorithm basically performs spectral and spatial filtering simultaneously. If you size the sliding window incorrectly, you run the chance of corrupting your background samples with target signature and degrading performance significantly. Because these sliding windows were sized perfectly to the data, the local ACE algorithm performs well and rejects the dry stream beds which are found in the physics-based detectors.

The physics-based detectors provide improved performance over the classical detectors. Not only do the algorithms identify the targets in the undisturbed soil regions, they also are able to isolate targets within the dry stream beds. This gives them a significant advantage over the other methods. The downside is that they also detect some of the dry stream beds as well. However, a spatial filter could significantly improve the detectors by removing anything that does not match the circular structure of the targets.
This leads to the interesting conclusion that the physically meaningful endmembers are the key to the physics-based detectors’ success. The FCLS based detectors have slightly better performance than the other two detectors, but this makes sense since the FCLS algorithm is used in the IEA endmember extraction method. What is particularly interesting is the AMSD/IEA algorithm. This method does not use the sum-to-one or non-negativity abundance constraints, but does nearly as well as the algorithms that do. It is this finding that suggests that the endmembers are the critical parameters that improve LWIR HSI target detection.

The final note of this section focuses on processing times. The physics-based methods all provide detection scores in less than a minute. The local ACE detector, however, takes well over an hour to process one image. This is due to over 100,000 covariance estimates and inversions that are needed. Considering these algorithms have not been optimized, this is still a significant increase in processing time when using local ACE. Therefore, the physics-based detectors provide improved performance with significantly improved processing times over the local ACE detector.

5. SUMMARY

We presented a physics-based solution to LWIR HSI detection problems. Our analysis concluded that using physically-based endmembers leads to improved detection performance – especially for weak target signatures like disturbed soil. Additionally, the physics-based method took less than a minute to process the images compared to over an hour for local ACE methods.

For future work, fundamental research needs to be focused on estimating the number of endmembers for a particular detector. We found that the number of endmembers used was highly dependent on the type of detector. For example, FCLS used over 20 endmembers while the AMSD/IEA algorithm tended to use less than ten. Identifying the correct number of endmembers for a particular detector is a necessary step before this technology can be applied in real-world applications.

ACKNOWLEDGMENTS

The authors would like to thank Ms. Miranda Schatten and Mr. David Hicks of the US Army RDECOM CERDEC NVESD for the hyperspectral data used in our analysis.

REFERENCES


