

**MIST Final Report:
Multi-sensor Imaging Science
and Technology**

Michael A. Lind
Patricia A. Medvick
Michael G. Foley
Harlan P. Foote
Patrick G. Heasler
Sandy E. Thompson
Lisa L. Nuffer
Patrick S. Mackey
Jonathan L. Barr
Andrea S. Renholds

March 2008

Prepared for the U.S. Department of Energy
Under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, WA 98031

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.** Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service,
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161
ph: (800) 553-6847
fax: (703) 605-6900
email: orders@ntis.fedworld.gov
online ordering: <http://www.ntis.gov/ordering.htm>

Abstract

The Multi-sensor Imaging Science and Technology (MIST) program was undertaken to advance exploitation tools for Long Wavelength Infra Red (LWIR) hyper-spectral imaging (HSI) analysis as applied to the discovery and quantification of nuclear proliferation signatures. The program focused on mitigating LWIR image background clutter to ease the analyst burden and enable a) faster more accurate analysis of large volumes of high clutter data, b) greater detection sensitivity of nuclear proliferation signatures (primarily released gasses) , and c) quantify confidence estimates of the signature materials detected. To this end the program investigated fundamental limits and logical modifications of the more traditional statistical discovery and analysis tools applied to hyperspectral imaging and other disciplines, developed and tested new software incorporating advanced mathematical tools and physics based analysis, and demonstrated the strength and weaknesses of the new codes on relevant hyperspectral data sets from various campaigns. This final report describes the content of the program and the outlines the significant results.

Table of Contents

1	Introduction.....	1
1.1	Goals.....	1
1.2	Approach	3
1.2.1	Blind source separation	5
1.2.2	Known source separation	7
1.2.3	Site knowledge and multiple view separation	7
2	Basic Statistical Methods Investigated.....	8
2.1	Principal Components Analysis.....	8
2.2	Whitened Matched Filter	9
2.3	Independent Components Analysis.....	10
2.3.1	Enhancements using parallel processing	12
2.3.2	Enhancements using random sampling	12
2.3.3	Enhancements by combining with PCA	14
2.4	Fast Fourier Transform	15
2.5	Entropy Analysis.....	19
2.6	Sub-sampling for Adaptive Whitening	20
2.7	Blob Analysis for ROI	23
2.8	Error Estimates & Confidence	24
2.8.1	Signal to Noise	25
2.8.2	ROC curves.....	25
2.9	Bayesian Approaches.....	26
3	Physics Based Models Investigated	28
3.1	Model construction	29
3.2	Registration	31
4	Example Results	32
5	Summary and Recommendations.....	35
6	Papers, Publications, Presentations & Software.....	36

List of Figures

Figure 1. <i>Ideal processing of a hyperspectral data cube.</i>	2
Figure 2. <i>MIST program showing an implementation of the three different approaches with iteration for improving data analysis.</i>	4
Figure 3. <i>An illustration of the analysis process starting with the calibrated data cube and ending with a concentration profile and confidence estimate.</i>	5
Figure 4. <i>Compute time vs. number of processor CPU for a 600MB data cube.</i>	12
Figure 5. <i>Sub-sampling from 25% down to 1% of the original cube still yields valuable results with a large savings in computation burden.</i>	13
Figure 6. <i>ROC curve and S/N values for the ICAs generated from 25% , 10%, 5% and 1% randomly sub-sampled data cubes</i>	13
Figure 7. <i>The first ten principle components of a ground based hyperspectral image showing the stack plume in components 3,4,5,6,7,9 and 10.</i>	14
Figure 8. <i>Independent components 1 – 10 from principle components 1 – 10 with their associated histograms.</i>	15
Figure 9. <i>Overview of steps for using the 3DFFT as a Whitened 3D Matched Filter on a data cube, I is the imaginary and Q the real part of the Fourier transform.</i>	16
Figure 10. <i>Block diagram of pre-processing steps for implementation of processing steps for 3DFFT-WMF.</i>	17
Figure 11. <i>The effect of a FFT low pass filter with filter averaging diameters from 1 to 15 pixels.</i>	18
Figure 12. <i>Signal to noise of the plume in region 1 as a function of filter diameter.</i> ..	19
Figure 13. <i>The results of an entropy analysis of an urban scene with a high concentration gas release.</i>	20
Figure 14. <i>Ground based entropy image of near ground gas release (dark near ground region) courtesy of Harold Trease, PNNL</i>	20
Figure 15. <i>Illustrations of a) the signal to noise test region, b) the whitening sensitivity grid array showing the order in which the blocks contribute the most (1) to the least (64) in plume S/N, and c) the S/N performance curve derived from the individual and whitening blocks accumulated in order of sensitivity.</i>	21
Figure 16. <i>Illustration of a plume in an urban cluttered region, a) the WMF image of the plume, b) the partitioning scheme showing the areas that contributed the highest S/N (white) to the lowest (black), and c) the S/N from individual and cumulative background whitening.</i>	22
Figure 17. <i>The blob analysis process illustrating a) the object finder output, b) the signature vector scatter plot, and c) and final blob out put overlaid on the original image.</i>	23
Figure 18. <i>Relative probability of detection verses plume length for a) standard BMA, b) spatial BMA, and c) spatial and temporal BMA.</i>	28
Figure 19. <i>Aerial view of the industrial chemical production site that was modeled.</i> ..	29
Figure 20. <i>The site model that was constructed.</i>	30
Figure 21. <i>The broadband radiance from the MuSES model of the site.</i>	30

Figure 22. *Composite RGB image of the hyperspectral data from a portion of the site showing a plume in red. 31*

Figure 23. *Ground based LWIR HSI analysis images for three simultaneous gases using 3D-FFT (no spatial filtering) and the more traditional WMF analysis. 33*

Figure 24. *Ground based LWIR HSI analysis images for three simultaneous gases using 3D-FFT (no spatial filtering) and the more traditional WMF analysis. 33*

Figure 25. *Whitened images of two gas releases (upper and lower groupings) using the regions highlighted in red as the whitening area for clutter mitigation. 34*

1 Introduction

1.1 Goals

The goal of the MIST (Multi-sensor Imaging Science and Technology) program was the advancement exploitation tools for Long Wavelength Infra Red (LWIR) hyper-spectral imaging (HSI) applied to the discovery and analysis of nuclear proliferation signatures. The program was focused on mitigating LWIR image background clutter to ease analyst burden and enable:

- faster more accurate analysis of large volumes of high clutter data
- greater detection sensitivity of nuclear proliferation signatures (primarily released gasses) in LWIR hyperspectral images via clutter reduction
- quantified confidence estimates of the signature materials detected.

This program targeted:

- investigation of fundamental limits and logical modifications of the more traditional statistical discovery and analysis tools applied to hyperspectral imaging and other disciplines
- development and testing of new software incorporating advanced mathematical tools and physics based analysis
- development of an integrated tool set that optimizes and improves hyper-spectral analysis beyond the capabilities included the HIP (LANL), GTK (Sandia), HEAT (NGA) and ENVI (ITT - commercial) codes, and
- demonstrating the new code on relevant hyperspectral data sets from various campaigns

This report concentrates on the latter stages of the program to showcase the analysis techniques and the resulting improvements. To avoid classification the details of the releases targeted and the specific sensor systems used have not been included.

To put the program in perspective it is helpful to consider the elements that go into the extraction and meaningful quantification of data from an LWIR hyperspectral data cube. The intent of the report is not to discuss the technology or difficulties behind the actual data collection and what can be done to improve that process, but rather what needs to transpire after the data is collected to maximize information extraction. Figure 1 illustrates the abbreviated processes that could and should be considered in the processing and analysis of a data cube for maximum benefit.

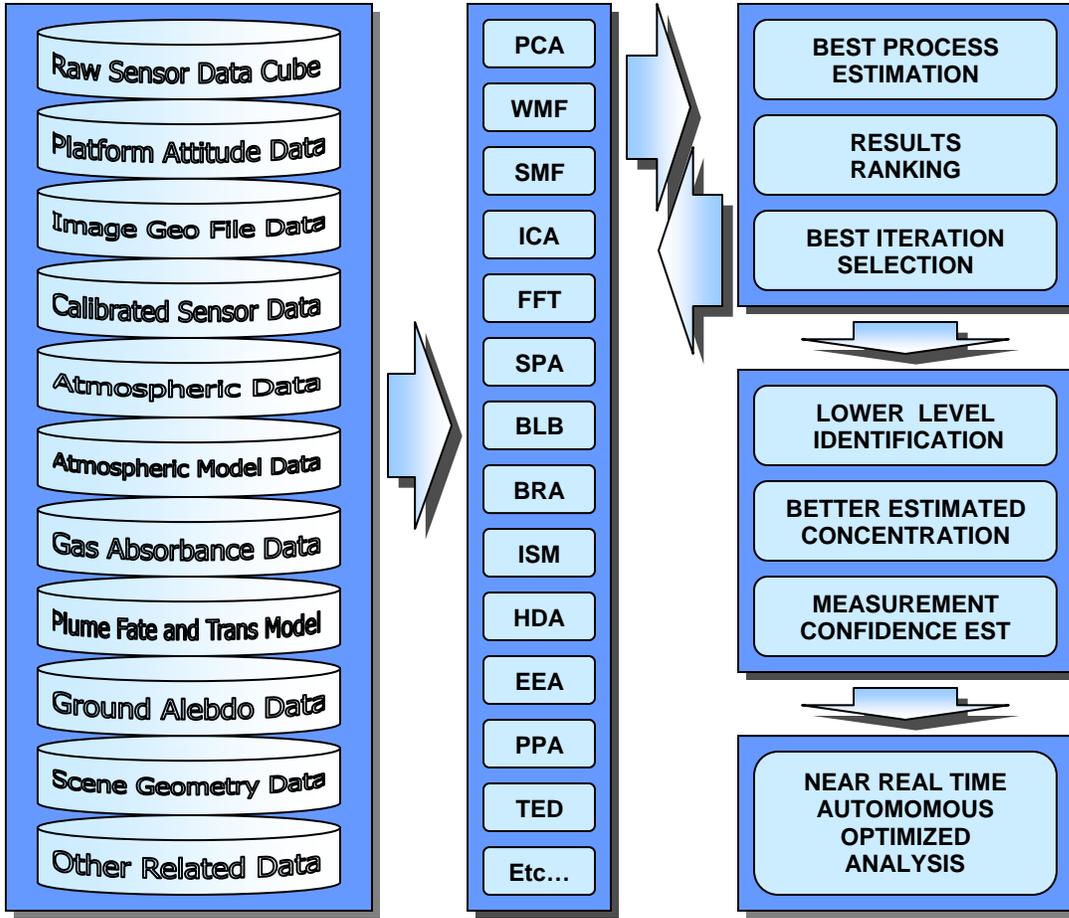


Figure 1. *Ideal processing of a hyperspectral data cube.*

The left hand block in the figure represents some of the data that must be available at the time of processing in addition to the data cube to fully exploit the LWIR data cube. For example, it is helpful to have the sensor attitude and altitude for assigning geospatial coordinates and viewing path information to each pixel. This can be especially important for geo-registration of features in line scan sensors where atmospheric buffeting causes pixel overlap and voids when compared to a simultaneously captured focal plane image. It is also critical for multi image utilization for noise reduction and realization of the advantages provided by staring modalities. In addition the sensor must be fully calibrated for absolute radiance as well as drift characteristics and compensated for collection optics corrections. Atmospheric path information is critical to high accuracy characterization as is the data on the materials being sought. In addition information about the background and the plume characteristics is required to make good quantitative estimates of effluent volumes. In actual practice, even on well designed experiments, little of this data (ground truth) is actually available to the analyst, making his job all the more difficult. Fortunately there are methods available for making some of the needed corrections to the data cube, but these established processes will not be discussed in detail here.

The second column represents some of the analysis techniques that have been investigated to various degrees in this project. These are

- PCA - Principal Components Analysis
- WMF – Whitened Matched Filters
- SMF - Spectral Match Filter / K-Means Clusters
- ICA - Independent Components Analysis
- FFT – Fast Fourier Transforms or 3D-FFT in three dimensions
- SPA - Super Pixel Analysis
- BLB – Blob Analysis to find regions of interest
- BRA – Bayesian Regression Analysis
- ISM – Iso-Mapping / manifold geometry analysis
- HDA – Hyper Diagram Analysis for multidimensional constructs
- EEA – Entropy Evolution Analysis – using Shannon entropy
- PPA - Projection Pursuit Analysis
- TED - Temperature – Emissivity Deconvolution
- Etc. – the list goes on

The results of our investigations showed that no single technique yields superior results in all cases for both sensitivity and specificity over various types of background clutter. In fact combinations of techniques and modifications of the more traditional processes often yield superior results with less or minimally added computational overhead if properly applied. The results of these investigations are illustrated in this final report. The results presented in this report were obtained using manual batch processing of a variety of data cubes and analysis techniques as they was insufficient resource available to build a fully automated analysis program to yield the best results on its own, a goal still worthy of pursuit in the future.

The third column of the figure points the way toward an automated process of optimization which was planned for the following year of the project but was unrealized in the early termination of the project except via a manually entered batch file process during the final year. By comparing the outputs of a few of the processes in the second column, ranking the results and iterating using the techniques on modified portions of the cube described later in this report, it is possible to obtain the better low level detection of effluents with better concentration estimates and higher confidence estimates. The challenge remaining is to do this in near real time, (shown to be feasible in this program with parallel processing) autonomous environment.

1.2 Approach

Since we did not have access to the instrumentation, the acquisition parameters or the calibration data in this project, we must make the assumption that we are starting with data cubes that are properly corrected to yield geo-spatially registered, spectrally corrected and calibrated radiance data. While found not to be strictly true for most real

data cubes examined, sufficient additional information can still be extracted to show the added value of the general approach.

The MIST approach uses three layers of computational analysis to provide optimized results as shown in Figure 2. The first layer is the cube analysis layer. The detection problem is approached from three different perspectives: blind source separation (looking for anomalous regions of interest in unknown backgrounds), known source separation (predefined materials in unknown backgrounds), and site knowledge source separation (substantial knowledge of the source site to mitigate what would normally be regarded as image clutter via site specific models or multiple registered images). The algorithms for each approach were evolved after considering the alternative computational approaches in the second column of Figure 1. Each approach can be used to iterate an optimal solution using the available analysis tools.

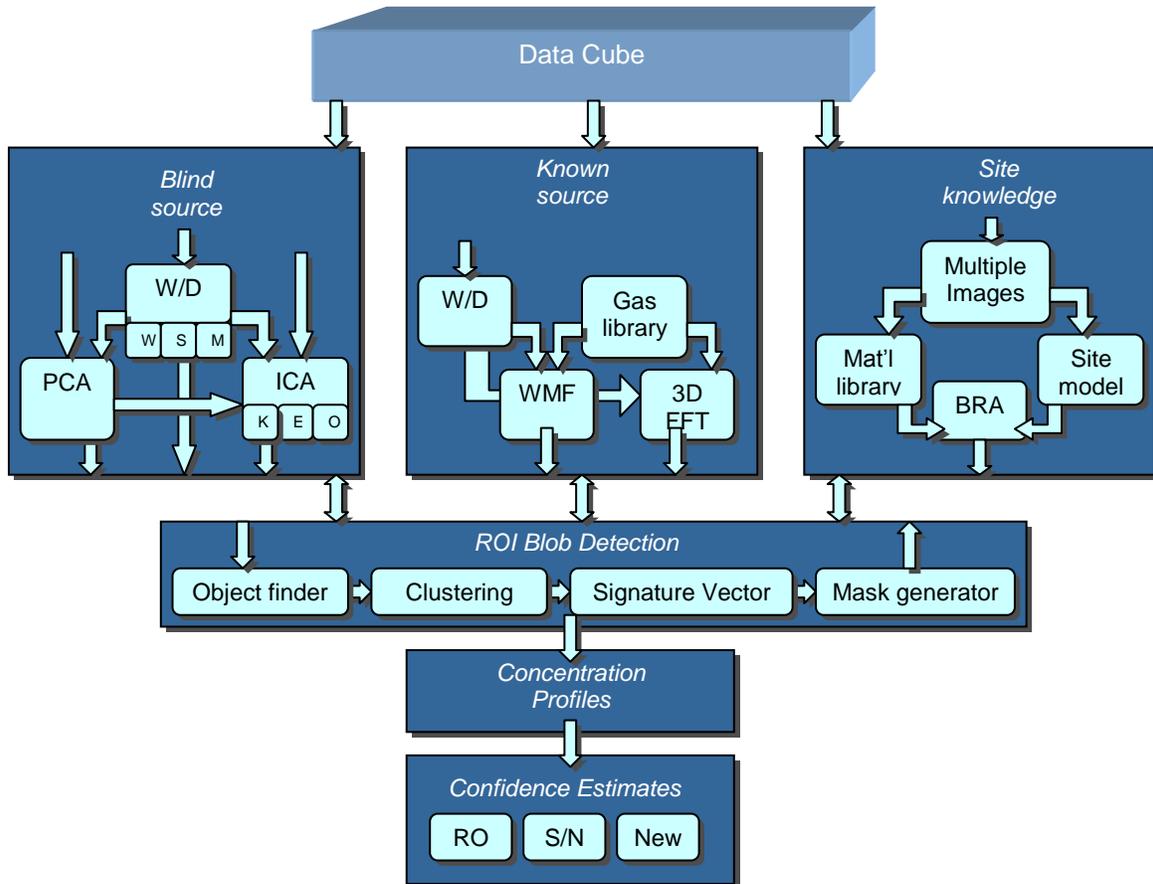


Figure 2. MIST program showing an implementation of the three different approaches with iteration for improving data analysis.

Once the data cube is assembled, one of the three analysis approaches is initially selected based on the type of ancillary information available. The details inside the boxes and the strengths and weaknesses for each computation method used in the three approaches will

be discussed later in the report. The reader is referred to the papers listed at the end of this report for additional discussion of the methods.

1.2.1 Blind source separation

The cube analysis process yields image outputs in which objects of interest (plumes in our case) are enhanced (normally unseen in spectral sampling alone). The process is illustrated graphically in Figure 3 for the blind source separation pathway.

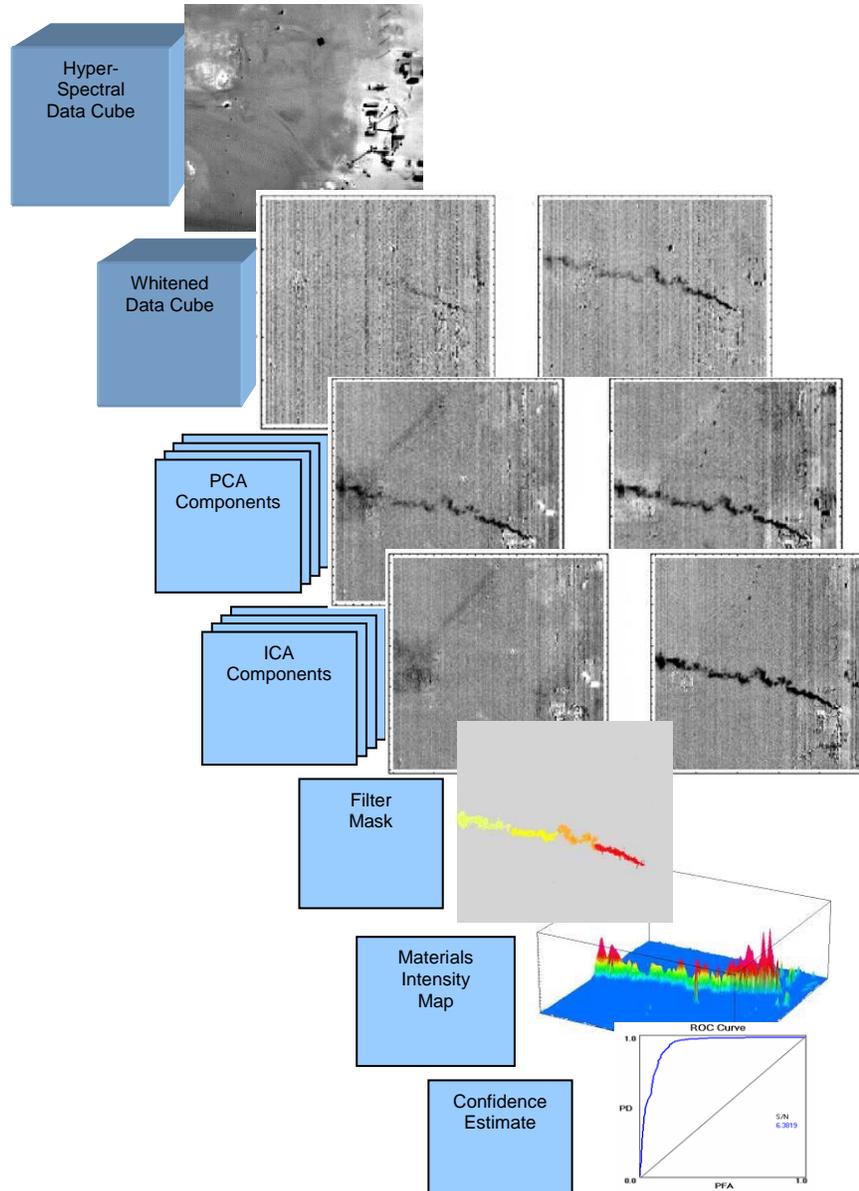


Figure 3. An illustration of the analysis process starting with the calibrated data cube and ending with a concentration profile and confidence estimate.

The process starts by importing a calibrated hyperspectral data cube. The cube illustrated in the figure depicts a gas release near the edge of a group of structures in a desert environment. A broadband image of the cube is shown at the top of the figure. No plume is readily apparent in any of the spectral bands.

The second step in the analysis process is to whiten the data cube. Whitening decorrelates and normalizes the data. It has the effect of reducing the noise by reducing the common feature similarities and thereby increasing the less dominant features. In the illustrated case the plume is one of those features. The procedure requires applying a covariance matrix calculated from either the entire image or from selected sections of the image. In the example above simple whitening over the entire cube reveals the first traces of this strong plume in several wavelength frames as shown in the second row of photographs.

The principle components analysis (PCA) using the parts of the covariance matrix that do not contain the plume further enhanced the plume contrast in some of those components. This is shown in the third row of images. It should be noted that PCA over the entire image is a common processing technique used in several analysis programs.

The fourth row of images shows further signal to noise enhancement is possible by applying independent components analysis (ICA) to those PCA components that showed the plume. Ideally only one ICA component will yield the plume. Note this ICA component yields the highest signal to noise ratio in this data cube. Since ICA is computationally intensive it can be greatly facilitated by selection of the appropriate PCA components first for inclusion in the ICA computation. PCA is much less computationally intensive; hence faster overall computation is possible as the preceding step.

Once the highest definition feature is extracted, a mask of the plume is created. The mask can be simple threshold based with manual exclusion of pixels obviously not in the plume or it can be automatic probability based on plume like characteristics. The blob detection algorithm was established to do the latter, but is still in its exploratory phases at this writing. These regions of interest (ROI) result in a ROI mask. With this mask the data can again be iterated through the analysis layer improving the mask and refining the signature probabilities. Once the final mask is established, the concentration profiles can be determined and confidence estimates prepared.

In the absence of another predictive mechanism, the mask provides a “truth” area for a concentration estimate on gas absorption levels under the mask and a confidence estimate based on residuals inside and outside of the mask area. There are many ways of looking at the confidence estimate that were investigated. No one method proved to be universally acceptable. The ROC curve shown in the illustration is one method that is useful in many circumstances. The ROC curve generation program also provides a contrast to noise ratio often used as the signal to noise ratio for flat images and data cube projections. Other measures of signal to noise are also helpful in other situations.

It should be noted in the illustration above that depending on the gas and the concentration it may not be necessary to go through all the steps cited. Simple whitening to ICA may be sufficient to yield a reasonable ROI in scenes with high concentration or high absorption coefficient gasses.

1.2.2 Known source separation

When the source is known (target gas is known), then it is more expedient to first employ the method shown in middle box of Figure 1. Here a materials library (gas library or a single gas spectral absorbance file in our case) is used to construct a whitened matched filter (WMF) which is convolved with either the data cube or a previously whitened data cube to yield the results.

Alternatively a Fast Fourier Transform (FFT) can be used to perform the operation using either a raw data cube or a whitened data cube for the operation. The advantage a 3-D FFT is that both spatial and spectral correlations can be performed simultaneously so that spatial and spectral filters can be used.

Both methods yield reasonable results with relatively short computational times. The methods are described in greater detail later in the report and are further explained in the reports cited at the end of the report.

1.2.3 Site knowledge and multiple view separation

When the site can be characterized in detail, then using physics based modeling it is often possible to extract even greater information from backgrounds that would otherwise be considered clutter. The approach can involve applying a variety of tools and usually some form of regression (Bayesian in our case) to find targeted features. Unfortunately this approach is often time consuming for targets of interest since additional information about the target area must be incorporated.

Examples of this type of analysis are the following:

- a) Multiple site view that allows image averaging thereby increasing the signal to noise (S/N) ratio by the square root of the number of registered images.
- b) Physical site modeling taking into account real materials characteristics, atmospheric conditions, illumination conditions, viewing angles, gas propagation models, etc..
 - In the LWIR both temperature and emissivity play a large role and when they can be deconvolved, better estimates of the region of interest and the gas concentrations can be derived.
 - When likely plume sources can be located and estimated of plume dispersions made, then better starting estimates for Bayesian regression can be made.

- When the background can be modeled in sufficient detail, better estimates for reducing background clutter can be made.

All these process require registration in the sub-pixel resolution range to be most effective. A summary review and status of what has been accomplished in the MIST program for all the above methods follows in this report.

2 Basic Statistical Methods Investigated

2.1 Principal Components Analysis

Principal Components Analysis (PCA) is a well established technique for finding objects buried in image clutter and is a common component of many commercial software packages. In our case it is used as a baseline comparison technique.

The advantage of PCA is that it is a relatively rapid calculation that allows visualization of high dimensionality data cubes in a lower dimensional representation. Technically it is an orthogonal linear transformation that rotates the data into a new coordinate system such that the largest variance lies on the first coordinate (principal component). The next largest variance lies on the second coordinate and so forth up to the maximum spectral dimension of the data cube.

The calculation is done by organizing the data cube into column vectors representing the N spectral values (column) at each M pixel (row) in the scene. This matrix X is an array of dimension $M \times N$. The empirical mean vector of the data is then calculated (dimension $M \times 1$)

$$\mu[m] = \frac{1}{M} \sum_{n=1}^M X[m, n]$$

and subtracted from the original matrix to form a scatter matrix A ,

$$A = X - \mu \cdot \mathbf{h}$$

where \mathbf{h} is a unit $1 \times N$ row vector.

Then the $N \times N$ covariance matrix is calculated by taking the outer product of the scatter matrix with itself for these real-value data.

$$K = \frac{1}{M} AA^T$$

The eigenvectors \mathbf{V} , which diagonalize the covariance matrix, are then calculated using one of several standard algorithms. The corresponding eigenvectors are reordered by decreasing eigenvalue into a transformation matrix \mathbf{W} , comprised of a selected subset of the eigenvectors with the highest eigenvalues. The resultant projection of any component is then viewed as

$$\mathbf{B} = \mathbf{W}^T \mathbf{A}.$$

Depending on the strength of the plume and the uniformity of the background clutter, the technique can reduce the dimensionality of the original cube substantially with the plume becoming visible on several of the higher eigenvalue projections. The lowest eigenvalue projections contain the bulk of the clutter.

2.2 Whitened Matched Filter

When looking for plumes with known constituents Whitened Matched Filters are a commonly used technique. The MIST program also examined this technique with a few new twists that proved useful for S/N enhancements.

Many data cubes are highly structured and contain significant statistical redundancy. Whitening linearly transforms the signal mixtures (vectors) \mathbf{A} into another set of mixtures (vectors) $\tilde{\mathbf{A}}$, each with unit variance, that are uncorrelated with each other. In the spectral domain the whitening matrix \mathbf{V} is derived from the covariance matrix \mathbf{K} using standard eigenvalue decomposition. By definition

$$\mathbf{K} = \mathbf{E} \mathbf{D} \mathbf{E}^T$$

where \mathbf{E} is the orthogonal eigenvector matrix and \mathbf{D} is the diagonal matrix of its eigenvalues, then

$$\mathbf{K}^{-1} = \mathbf{E} \mathbf{D}^{-1} \mathbf{E}^T$$

and

$$\mathbf{V} = \mathbf{E} \mathbf{D}^{-1/2} \mathbf{E}^T = \mathbf{K}^{-1/2}.$$

Applying whitening transforms the matrix \mathbf{A} into a new matrix that is both orthogonal and effectively reduced in the number of parameters from n^2 to $n(n-1)/2$. The whitened matched filter (WMF) operation is equivalent to correlating the whitened signature vector \mathbf{b} with a whitened data cube \mathbf{A} . The resulting whitened matched filter equals the whitened signature times the whitened data cube.

$$(\mathbf{K}^{-1} \mathbf{b})^T \mathbf{A} = (\mathbf{b}^T \mathbf{K}^{-1/2}) (\mathbf{K}^{-1/2} \mathbf{A})$$

In addition, the MIST program showed that the effectiveness of whitening in reducing or eliminating clutter can be increased by careful selection of areas of the image for the whitening calculation that do not contain the target of interest.¹

2.3 Independent Components Analysis

Independent Component Analysis (ICA) has been shown to be a useful tool for blind source separation of mixed signals in many applications. In the MIST program it was applied to the high clutter LWIR data cube problem.

The method finds components that are both maximally statistically independent and maximally non-Gaussian. Thus, the components share minimum mutual information or maximum joint entropy. In an ideal situation, the de-mixing would be complete and the plume would emerge in a single component with the clutter removed. However, this is rarely the case in real data.

There are many approaches to ICA. The techniques used in the MIST program uses the fast ICA² algorithm. It approximates a solution to the problem

$$\mathbf{x} = \mathbf{A}\mathbf{s}$$

where \mathbf{x} is a vector of mixed components, \mathbf{A} is a square mixing matrix, and \mathbf{s} is a vector of independent components. The problem is to find \mathbf{s} and \mathbf{A} when neither is known. This is done by assuming the components of \mathbf{s} have non-Gaussian statistics and thus sums of those components will be more Gaussian than the individual components.

The algorithm looks for an array of independent components \mathbf{S} that are as non-Gaussian as possible by constructing an un-mixing matrix \mathbf{W} of orthogonal basis vectors pointing in the directions of maximum non-Gaussianity and multiplying by an array of whitened mixture vectors \mathbf{Z} .

$$\mathbf{W}\mathbf{Z} = \mathbf{S}$$

Whitening \mathbf{Z} removes the second order correlations from the mixtures contained in the covariance matrix enabling use of only the higher order information and orthogonalizes the data. As discussed in the whitening analysis section above

$$\mathbf{Z} = \mathbf{K}^{-1/2}\mathbf{X}$$

¹ Lind MA and J Barr. *Modeling and Software Scene Simulation for Long Wavelength Infra-Red Hyper-Spectral Imaging Ground Clutter*. PNNL-SA-55971, Pacific Northwest National Laboratory, Richland, Washington.

² Hyvarinen A. 1999. "Fast and Robust Fixed-Point Algorithms for Independent Component Analysis", *IEEE Trans. On Neural Networks*, 10(3):626-634

where \mathbf{K} is the covariance matrix of the input mixtures \mathbf{A} .

That actual algorithm (a simplified Newton method) proceeds as follows:

1. Center the data for zero mean
2. Whiten the data to give \mathbf{Z}
3. Select the number of independent components that are to be estimated $m < n$ and set the loop counter to $i=1$
4. Randomly select an initial value for \mathbf{W}_i normalized to unity
5. Let $\mathbf{W}_i = E\{\mathbf{Z}g(\mathbf{W}_i^T\mathbf{Z})\} - E\{g'(\mathbf{W}_i^T\mathbf{Z})\}\mathbf{W}_i$, where E is the expectation value and g is a nonlinear function of negentropy (other non linear functions can be used for g such as Kurtosis)
6. Orthogonalize $\mathbf{W}_i = \mathbf{W}_i - \sum_{j=1}^{i-1} (\mathbf{W}_i^T\mathbf{W}_j)\mathbf{W}_j$
7. Normalize $\mathbf{W}_i/\|\mathbf{W}_i\|$
8. If \mathbf{W}_i has not converged, go back to step 5
9. Increment the loop counter and go to step 3
10. Calculate the IC components $\mathbf{S} = \mathbf{WZ}$.

Note that the initial vector is chosen at random, so the order of the components can vary from calculation to calculation. Additionally convergence speed depends on the non-linear function g that is chosen. Kurtosis, entropy and other high order functions can be used. A more detailed explanation of the underlying concepts for the algorithm can be found in the Hyvarinen reference.³

The downside of ICA is that it is computationally intensive often taking many hours to calculate the components of a large hyper-spectral data cube. On the upside, the computational time can be decrease substantially using parallel processing. This was demonstrated under the MIST program.

Three techniques were explored under the MIST program to decrease computation time while taking maximum advantage of the benefits derived from non-Gaussian clutter reduction and the reduced number of resultant component views containing plume information. One technique used parallel processing with up to 32 processors to cut processing time substantially, another used random sampling of the data cube to reduce the physical size of the data cube and thus decrease the calculation time, and the third employed PCA prior to ICA taking only those PCA components that showed evidence of the plume to reduce the size of the computational matrix and provide a single component for viewing.

³ Hyvarinen A, J Karhunen, and E Oja. 2001. *Independent Component Analysis*. John Wiley & Sons, Inc. New York.

2.3.1 Enhancements using parallel processing

Since ICA is can be coded for parallel processing, a series of test were performed to explore the computational time enhancements that were possible. In this preliminary investigation we used up to 32 processors operating in several configurations on large representative data cubes to calculate the ICA components. The results follow the trends shown in Figure 4. Clearly there is a computation time advantage with decreasing efficiency as the number of processors increases. For our algorithms the most efficient tradeoffs between compute time and power started saturating at about 16 processors where the improvement was about a factor of 8. The exact benefit depends somewhat on the size of the cube and the configuration of the processors. Adding another 16 processors can decrease computation time by another 25%, where the process is reaching highly diminished returns.

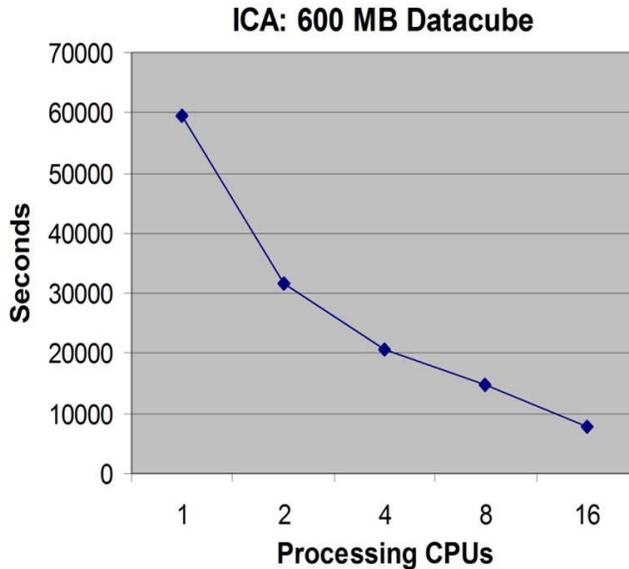


Figure 4. Compute time vs. number of processor CPU for a 600MB data cube.

2.3.2 Enhancements using random sampling

The time consuming part of the Fast ICA algorithm involves an expected value calculation over all the pixel vectors in the cube. Since a typical cube contains on the order of one to 10 million elements, a good approximation to the expected value can be made with 10-20% of the pixels, giving a corresponding reduction in computation time. The process is illustrated in Figure 5. Here the original data cube is sub-sampled to 25%, 10% , 5% and 1% of the total cube via random sampling. The plume appears in different components since the components are essentially randomly distributed unlike PCA where they can be eigenvalue ordered. Some plume appear white and some dark, but they are all visible although some what decreased in image contrast based signal to noise as the sub-sample density is reduced.

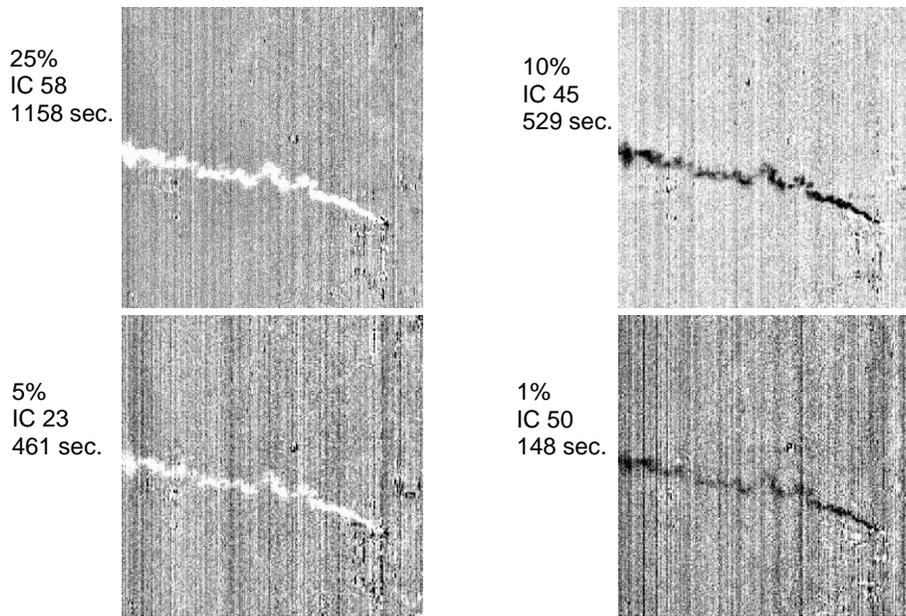


Figure 5. Sub-sampling from 25% down to 1% of the original cube still yields valuable results with a large savings in computation burden.

The general trend of the image contrast signal to noise and the threshold based ROC curves for the four images is showing in Figure 6. Clearly as the sampling frequency reduces, the S/N decreases as does the probability of detection/probability of false alarm ratio. For a complete explanation of these performance parameters the reader is referred to a related paper.⁴

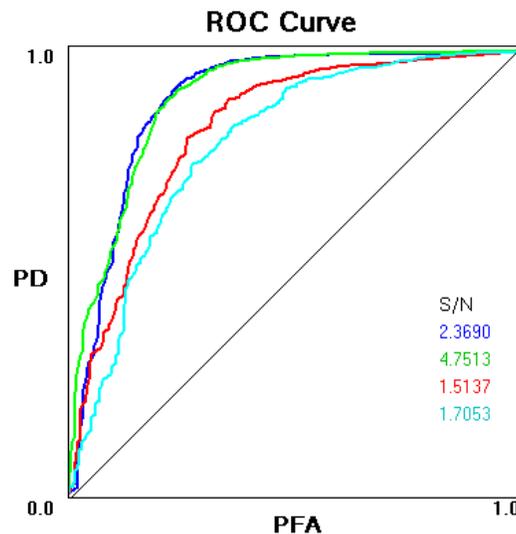


Figure 6. ROC curve and S/N values for the ICAs generated from 25%, 10%, 5% and 1% randomly sub-sampled data cubes

⁴ Michael A. Lind, *Software for Generating ROC Curves for Hyperspectral Data Cubes and Flat Images*, PNNL-17243, December 2007, Pacific Northwest National Laboratory, Richland, WA

2.3.3 Enhancements by combining with PCA

The bands of the HSI cubes are generally highly correlated resulting in a large data compression in the PCA components. The eigenvalue spectrum can be used to select an appropriate cut-off and reduce the number of components by more than 50% and often more than 75%. This reduced set of PCA components can be fed to the ICA step with sub-sampling to give a substantial improvement in overall compute time over ICA of the full data cube. The PCA components often contain multiple copies of a detected plume. Applying a subsequent ICA step can frequently combine these multiple plume copies into a single stronger plume component. The following example illustrates the process of the combined PCA-ICA processing.

To illustrate the point consider ground based desert image cubes analyzed with a combination of PCA and ICA methods. These cubes proved to be particularly good candidates for these blind detection methods. The plumes are viewed against a relatively uncluttered sky as the background which is less cluttered than most cubes collected from a nadir viewing geometry. Figure 7 shows an image sequence of the first 10 components of the PCA. The gas plume from the stack which is not visible in any of the spectral components alone is easily visible in several of the principle components.

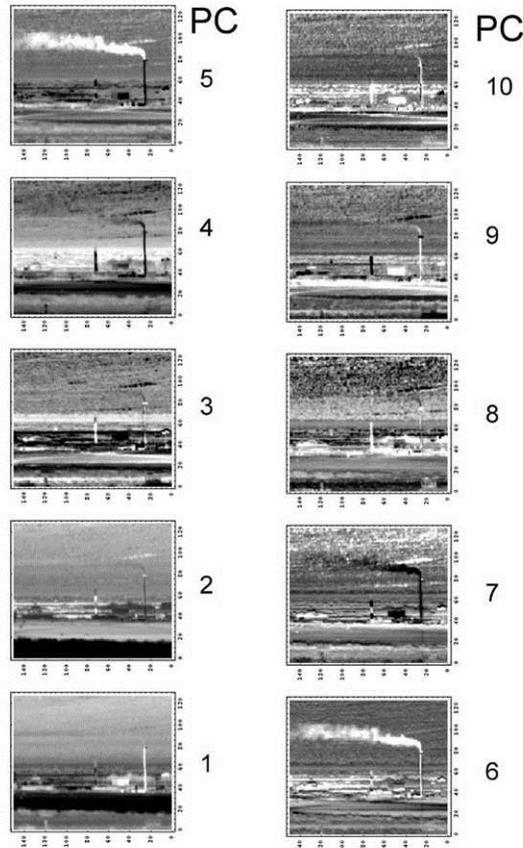


Figure 7. The first ten principle components of a ground based hyperspectral image showing the stack plume in components 3,4,5,6,7,9 and 10.

Figure 8 shows the results of ICA processing of this group of PCs. It also shows the associated intensity histograms. Here the plume is isolated as one independent component, a substantial improvement over the PCA result. ICA assumes the input data can be modeled as sums of independent components and ICA looks for components whose distributions are less Gaussian than the input mixtures. Component 7 which contains the plume is clearly non-Gaussian. Several of the other components are also visibly non-Gaussian. For the cubes with stronger plumes the ICA detection compares favorably with the Whitened Matched Filter output.

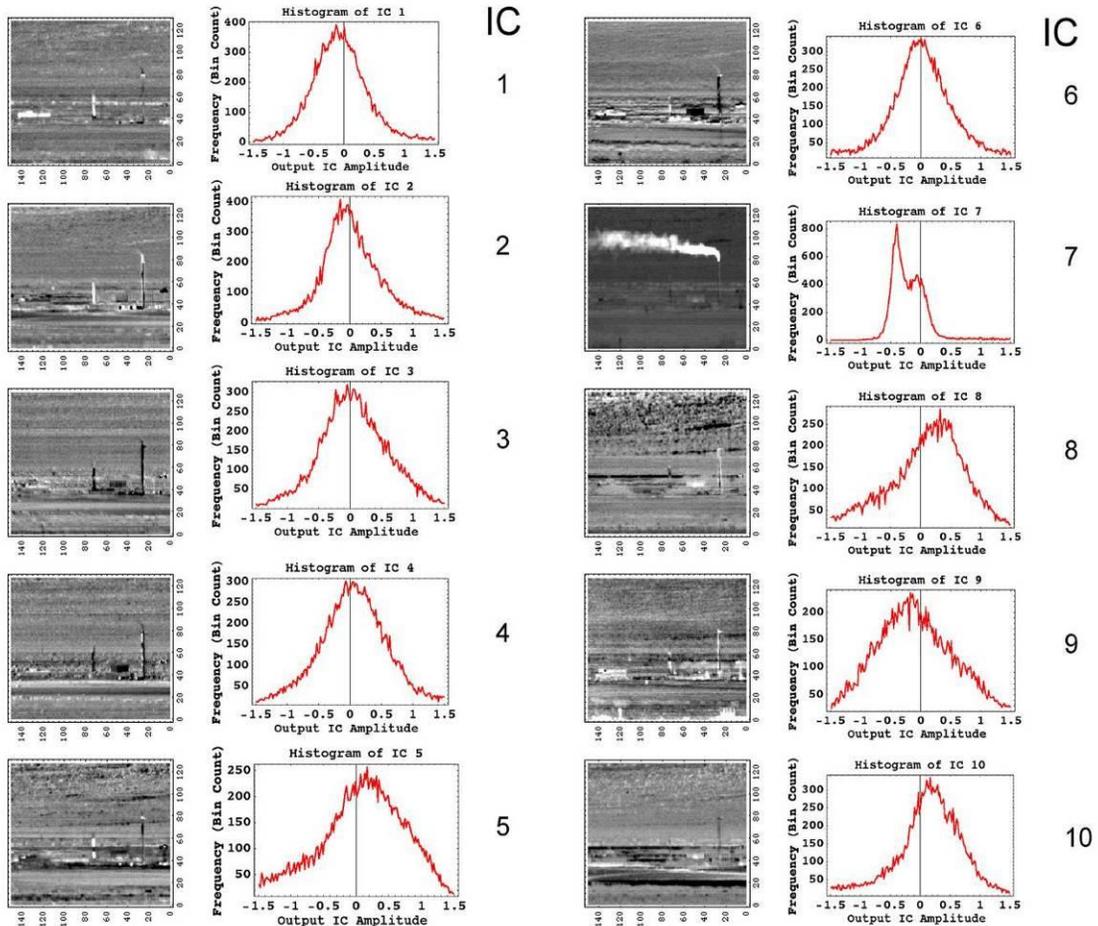


Figure 8. Independent components 1 – 10 from principle components 1 – 10 with their associated histograms.

2.4 Fast Fourier Transform

The application of whitened matched filtering to hyperspectral data cubes within the Fourier domain is another approach to the detection problem. The implementation of a Three Dimensional Fast Fourier Transform - Whitened Matched Filter (3DFFT-WMF) approach is fast method of achieving ROI identification often enhancing S/N over standard WMF techniques by using both spatial and spectral filtering. However, since

the variability in shape of gaseous plumes precludes the full use of spatial conformation in the matched filtering, the 3DFFT-WMF results often do not differ from the two other WMF methods.

Instead of a single-dimensional matched filter correlated separately along each pixel of the hyperspectral data cube (i.e., pixel by pixel), a 3D filter cube is correlated with the entire cube. This technique increases the signal to noise ratio in the filtered image by taking advantage of the fact that plume pixels are localized near one another. The Fourier domain provides opportunities for both improved processing efficiency and improved filtering

The overview diagram in Figure 9 displays the steps of applying the 3DFFT to a data cube. The inputs are a) the 3D data cube and b) the 3D filter representation of the spectrum of interest. Both inputs are 3DFFTed shifting them into frequency domain. Whitening in Fourier space is achieved by creating a smoothing function from the Fourier-transformed data cube to apply to the Fourier-transformed spectral filter cube, thus creating a Whitened 3D Matched Filter (W3DMF). Multiplying this W3DMF by the Fourier-transformed data cube is equivalent to correlation in the spatial domain. Applying the 3D inverse FFT then produces the filtered image data cube in which the first layer contains the brightened pixels that correspond to the signature of interest.

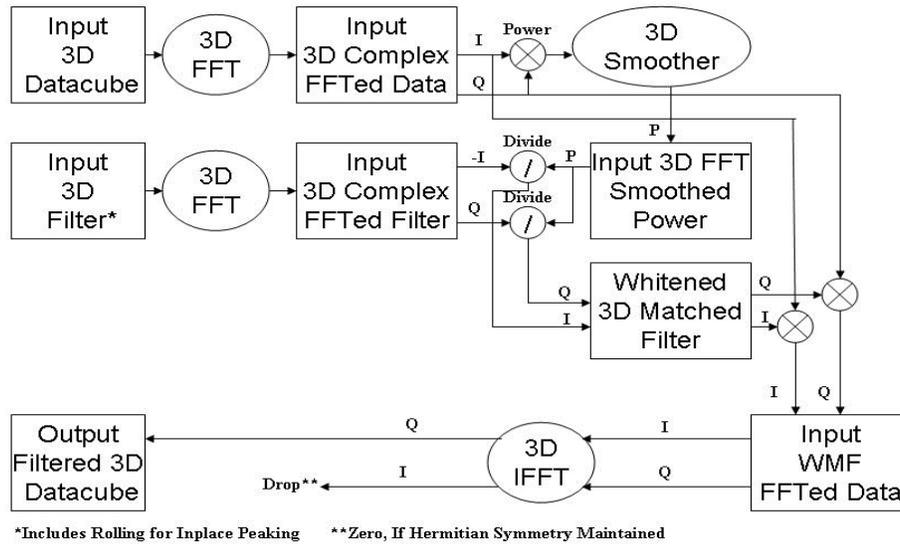


Figure 9. Overview of steps for using the 3DFFT as a Whitened 3D Matched Filter on a data cube, I is the imaginary and Q the real part of the Fourier transform.

Detailed steps in pre-processing/whitening are displayed in Figure 10. Data inputs are the 3D data cube, a chemical spectrum, and the wavelengths of the sensor. There is an option for whitening the data cube to reduce noise. Since the spectral filter and data cube-processed results are multiplied together in step 9, both the filter and the data cube may be whitened. In this process the whitening step consists of multiplying by the square root of the inverse covariance matrix (ICM) of the data cube. Pre-FFT whitening may be

accomplished either by multiplying both the data cube and filter by the square root of the ICM or by multiplying the data cube by the ICM.

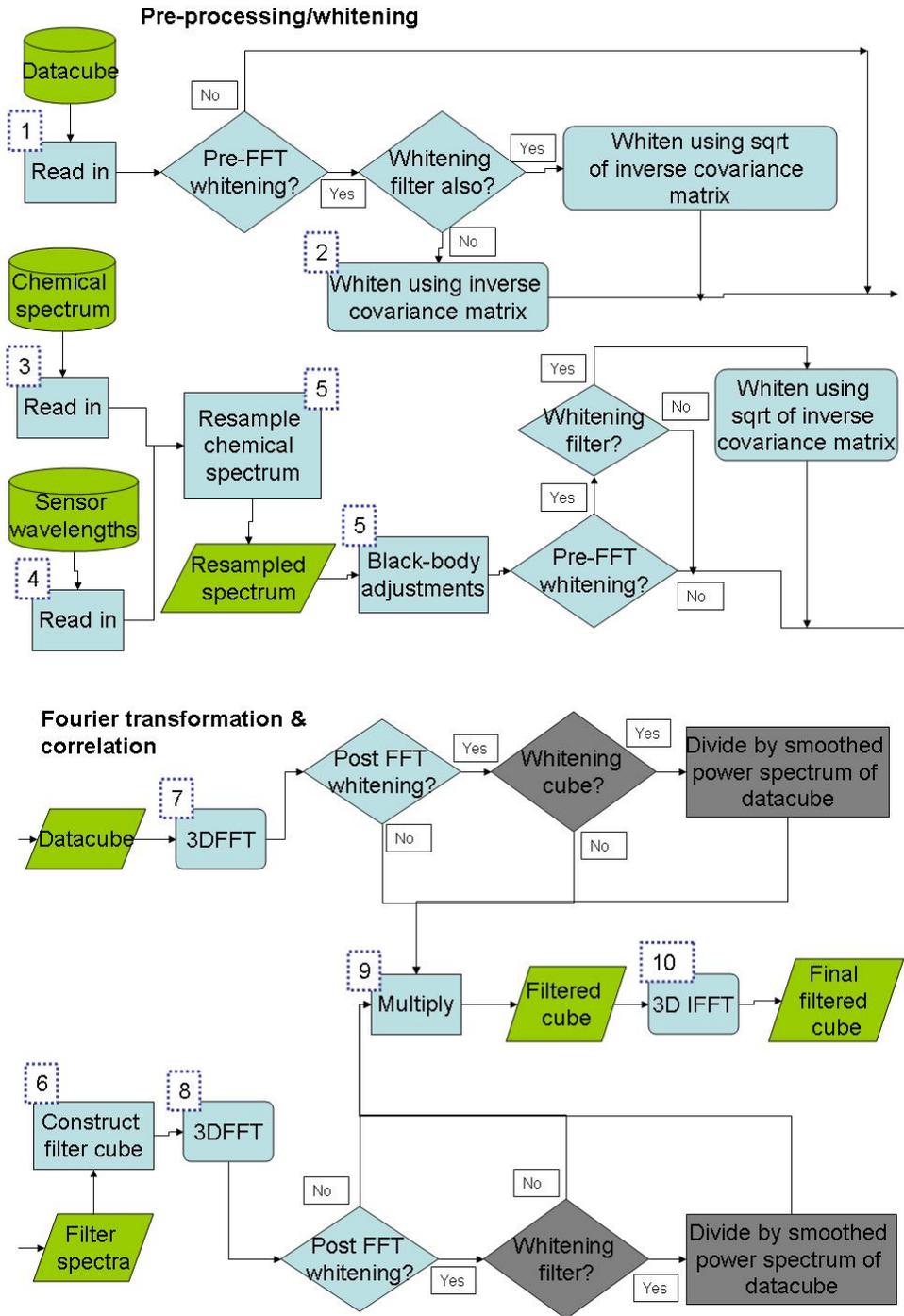


Figure 10. Block diagram of pre-processing steps for implementation of processing steps for 3DFFT-WMF.

With hyperspectral images the spectral signature must be matched to the image format of the sensor. This is achieved by sampling a finely delineated spectrum of the object of interest in a manner that matches the sensor apodization. Using environmental information to adjust for differences between sampling in a natural environment and the laboratory spectral measurements creates the real-world situation-specific signature. The resulting one-dimensional signature must be expanded by replication into a cube to match the spatial size of the hyperspectral data cube.

Both the 3D data cube and filter cube are subjected to a 3DFFT. Each of the results may be whitened by the smoothed power spectrum of the data cube. The two cubes (whitened again or not) are then multiplied together, and the resulting filtered cube subjected to an inverse 3DFFT. The resulting cube tends to contain the signature of interest in one band.

Due to the well-developed techniques to calculate FFT and inverse FFT the 3DFFT-WMF can have an analytical advantage over the standard WMFs in processing time. The details of the implementation, such as which libraries are used, have the major influence on computation time.

Including a spatial low-pass filter within the Fourier space can improve signal to noise ratios in a manner similar to super pixel averaging and therefore improve detection limit by facilitating the mitigation of high frequency clutter. The low pass improvement only occurs if the filter diameter is smaller than the plume diameter. High pass filtering can also be useful in some backgrounds, but less so routinely. An example of the spatial averaging using the FFT is shown in Figure 11. Here we show low pass spatial filters with radial dimensions between 1 and 15 pixels.

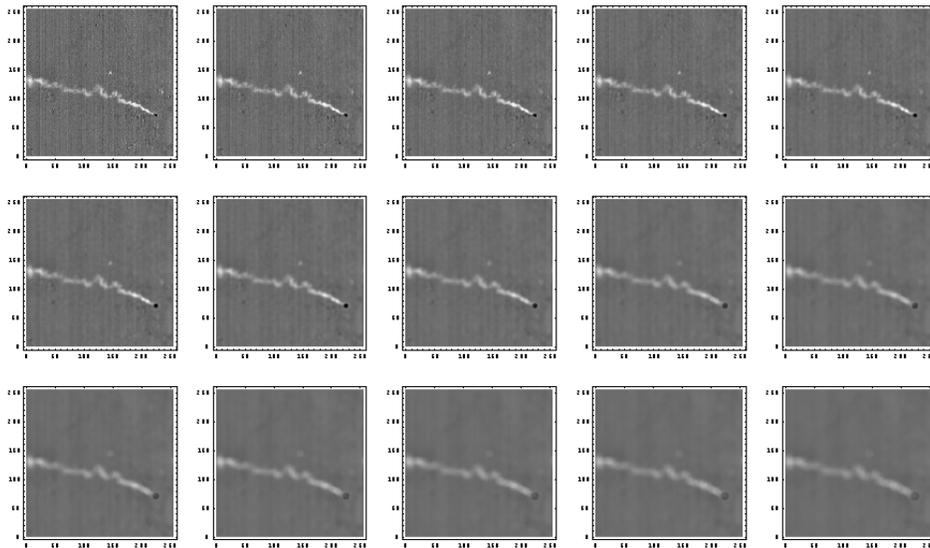


Figure 11. *The effect of a FFT low pass filter with filter averaging diameters from 1 to 15 pixels.*

The signal to noise of the plume region 1 highlighted in red in the figure varies from 14dB with the radius 1 filter to 24dB with the radius 15 filter. The S/N starts to degrade again when the filter radius reached the flume width. The results are plotted in Figure 12.

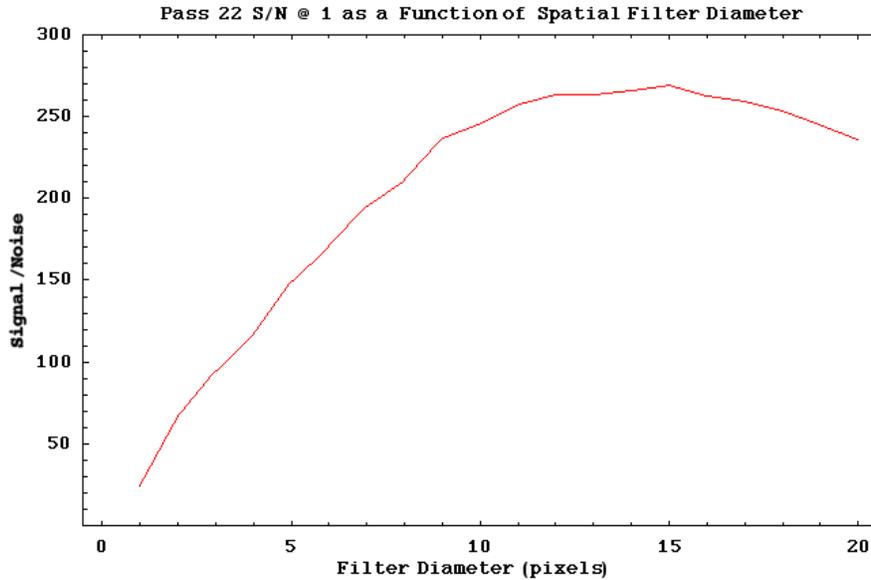


Figure 12. Signal to noise of the plume in region 1 as a function of filter diameter.

2.5 Entropy Analysis

Entropy analysis is a technique for rapid screening of data cubes that can yield impressive rapid plume location in some cases. The technique is usually used to measure the degree of randomness in the image data and hence to determine the degree of compressibility, as high entropy images are less well compressed than low entropy images. Although there are many ways to approach this characterization, we examined the “zero-order” or Shannon Entropy which is defined as

$$H(p) = -\sum_{i=1}^M p_i \log_2 p_i$$

where H is calculated as the sum of the variation between adjacent pixels across the entire image in a selected wavelength band and p_i is the relative probability that the spectral radiance level falls in the i^{th} quantized level of a finite histogram of M levels of radiance values in that wavelength band.

By calculating the entropy for selected wavelength bands and performing a simple subtraction between the target band and a nearby reference band, the plume image appears. If the plume concentration is high enough, then the reference band can be the average of all bands. The technique does not have the sensitivity of the other statistical

methods described above, but it does provide a fast method to locate high concentration plumes. An example of a strong plume in an urban region is shown in Figure 13. The plume shows dark in the image,



Figure 13. *The results of an entropy analysis of an urban scene with a high concentration gas release.*

A second example of a ground based image of a near ground release plume, again relatively high concentration although not visible in any individual spectral component, is shown in Figure 14.



Figure 14. *Ground based entropy image of near ground gas release (dark near ground region) courtesy of Harold Trease, PNNL*

2.6 Sub-sampling for Adaptive Whitening

Adaptive whitening stems from the observation that the background clutter does not contribute equally to the whitening step. Thus the use of selected areas of the data cube, preferably those most representative of the background clutter and not containing the plume, for the whitening should maximize the contrast and the signal to noise of the plume.

To illustrate the relationship between clutter content, whitening decorrelation, and plume detection sensitivity, the data cubes are divided into grid blocks whose individual contributions to plume detection are measured by calculating the signal to noise ratio of the known plume using each of the whitening blocks. The whitening blocks are ordered

by decreasing plume S/N ratios and then progressively accumulated one block at a time while recomputing the plume S/N at each step in the progression. The resulting plot of S/N ratio vs. number of whitening blocks shows the benefits of adaptive whitening.

Two clutter environments are examined, a desert environment dominated by slowly varying random clutter, and a highly organized urban area with a mixture of structures, streets, vegetation area, and parking lots. The desert plume in Figure 15 shows the relative block S/N performance (best = 1, worst = 64) and the cumulative result of combining whitening blocks. Clearly, data blocks containing the plume have a negative effect on detection performance. However, note that there is considerable variation in the off-plume blocks as well. The result indicates there was a substantial benefit to detection performance.

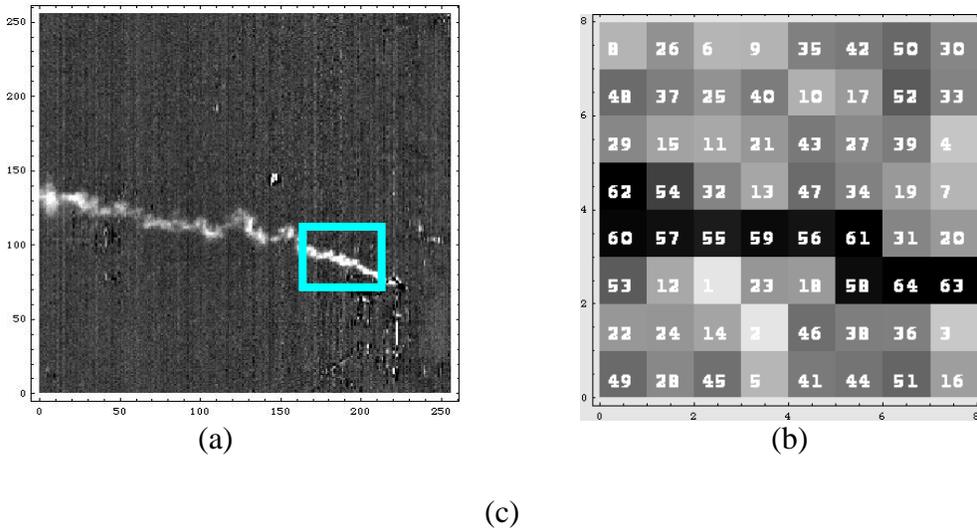
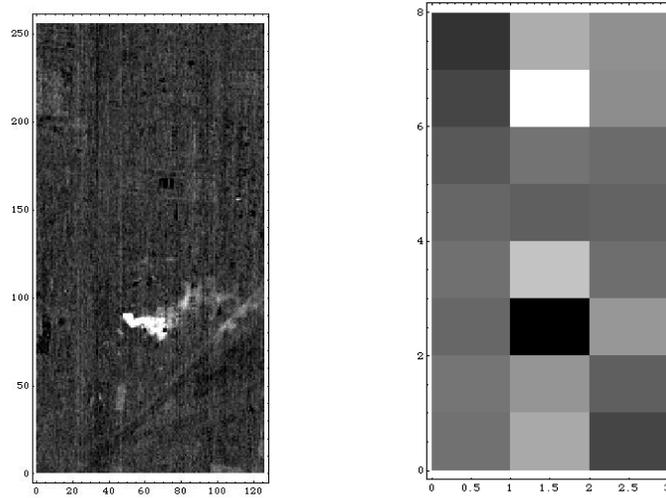


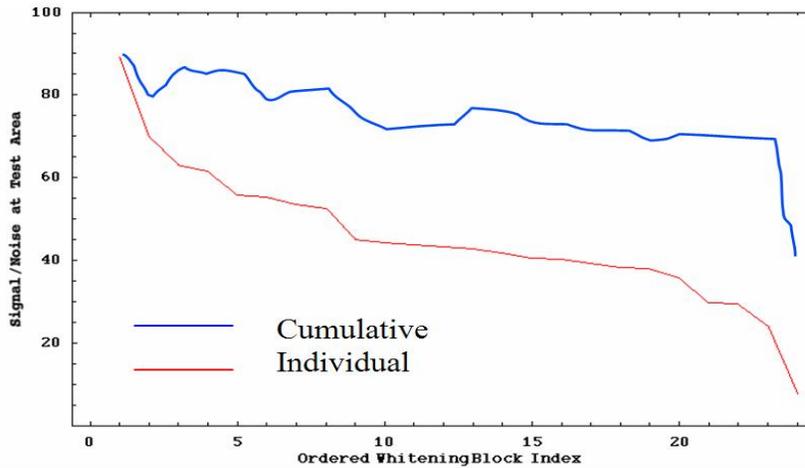
Figure 15. Illustrations of a) the signal to noise test region, b) the whitening sensitivity grid array showing the order in which the blocks contribute the most (1) to the least (64) in plume S/N, and c) the S/N performance curve derived from the individual and whitening blocks accumulated in order of sensitivity.

Figure 16 shows similar results for an urban plume. The smaller LWIR cube was partitioned a little differently, but the general results are even more dramatic. The highest S/N is obtained using a reference region similar in clutter to that under the plume, while



the lowest S/N region is contributed by including the plume area itself. Additional information on the technique can be found elsewhere.⁵

(a) (b)



(c)

Figure 16. Illustration of a plume in an urban clutter environment, a) the WMF image of the plume, b) the partitioning scheme showing the areas that contributed the highest S/N (white) to the lowest (black), and c) the S/N from individual and cumulative background whitening.

⁵ Patricia A Medvick, Lisa L Nuffer, Harlan P. Foote, Michael A. Lind, *Adaptive Whitening for Rapid Processing of Hyperspectral Images*, PNNL-16658, May 2007, Pacific Northwest National Laboratory, Richland, WA

2.7 Blob Analysis for ROI

The intent of the blob analysis work is quickly locate regions of interest (ROI) for further and more detailed analysis thus reducing the total analysis time for further spatial and spectral evaluation of a small regions of the entire data cube. The approach taken was to create software structure to segregate image features into separate components, delineate objects in components, and automatically select objects of interest. This is done by:

- Defining spatial-spectral features of blobs from multiple previously analyzed images to create an object finder. This can be a simple amplitude threshold detector for the components.
- Creating feature/signature vectors that describe the objects of interest (plumes in our case). The vectors initially can be based on available cube properties such as amplitude threshold, pixel proximity, and object continuity of similarity in multiple orders.
- Running the vectors through a self-organizing map (SOM)
- Partitioning signature objects for additional analysis and visual verification.
- Gathering information from interesting signature objects
- Repeating the process as necessary to extract the required information

The process is illustrated on an ICA analysis of a typical gas plume in Figure 17 below. The illustration shows: a) the application the object-finder to an ICA output image yielding a crude map of the plume and also highlights significant amounts of noise. The separate blobs are colored by intensity while the disconnected pixels are colored in black, b) a scatter plot of signature vectors derived from the blobs delineated by the object-finder and c) the blobs whose signature vectors were isolated in the scatter plot and are displayed as colored segments overlain on the original ICA output image.

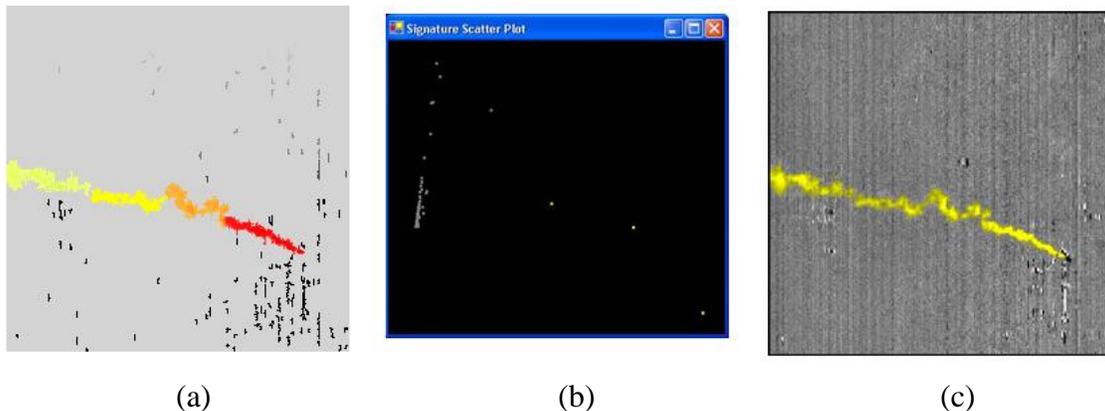


Figure 17. The blob analysis process illustrating a) the object finder output, b) the signature vector scatter plot, and c) and final blob out put overlaid on the original image.

This process has been investigated for other applications and the initial software was adapted for our use. Unfortunately time constraints forced a very limited looks at this

automated ROI identification process beyond simple threshold discrimination and continuity vector groupings which were used above.

2.8 Error Estimates & Confidence

There are many source so error inherent in hyperspectral images. They range from the instrumentation error (calibration, pixel registration, instrument temporal fluctuations, etc) to the electronic noise (detector noise, shot noise, etc.) to the image noise (thermal source, emissivity source, atmospheric variations, etc.) . Estimates are usually instrument operation, viewing propagation, and scene image specific. All can effect the final analysis confidence in one form or another. Statistically one can define confidence intervals or approximate confidence intervals as applied to single value random variables to indicate the probability over joint distributions or use Bayesian interval estimates (credible intervals) where the calculation is based on conditional observed variables includes nuisance parameters.

There are numerous metrics in the literature that can be used to compare analysis techniques and provide relative estimates of confidence. The methods include smallest distance ⁶, agreement index ⁷, noise amplification index ⁸, universal quality index ⁹, and distance/background variance ¹⁰, among others. A more complete discussion of these methods can be found elsewhere ¹¹.

The two more holistic measures of error that drove the research in this project and provided a basis of comparison for the various statistical and physics base modeling were signal to noise and Receiver Operating Characteristic (ROC) curves on the actual images recognizing that statistically there are many other estimates and confidence tests for relative error which have been used. These comparative error estimates provide the first step toward confidence estimates.

⁶ S. Singh and K. Bovis, *An Evaluation of Contrast Enhancement Techniques for Mammographic Breast Masses*, IEEE Transactions on Information Technology in Biomedicine, vol 9, no 1 (2005).

⁷ Y.M. Salman, M.A. Assal, A.M. Badawi, S.M. Alian, and M.El-M. El- Bayome, *Validation Techniques for Quantitative Brain Tumors Measurements*, Proceedings of IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, Sept. 1-4 (2005).

⁸ K. Rank, M. Lendl, and R. Unbehauen, *Estimate of image noise variance*, IEE Proc.-Vis. Image Signal Process, Vol. 146, No. 2, (1999).

⁹ Z. Wang, A. Bovik, *A Universal Image Quality Index*, IEEE Signal Processing Letters, vol 9, no 3 (2002)

¹⁰ A. Vanzo, G. Ramponi, and G.L. Sicuranza, *An image enhancement technique using polynomial filters*, Proceedings of IEEE Image Processing, Austin, TX, Nov. 13-16 (1994)

¹¹ Andrea S. Renholds, Sandy E. Thompson, and SJ Walsh. 2007. *Metric Literature Review 2007*. ., PNNL-SA-54694. Richland, WA

2.8.1 Signal to Noise

Both from the raw image perspective and the analyzed image perspective there are several methods for estimating signal to noise. Two signal to noise (S/N) methods have been used in our comparison work.

The standard measure of S/N in an image, actually a variation of the contrast to noise ratio. It requires identification of a region of interest (a plume, an area of suspected activity etc. It is represented by

$$S / N = |(\mu_{S+N} - \mu_S) / \sigma_N|.$$

Here μ_{S+N} is the mean signal plus the noise (background clutter) in the region of interest, μ_N is the mean of the noise outside of the region of interest and σ_N is the standard deviation of the noise outside the region of interest. In images where the region of interest is small in relation to the entire image, the standard deviation of the noise is often calculated over the entire image. Taking the absolute value maintains a positive value whether the region of interest is dark or light. Note that using this method the signal-to-noise ratios are relatively small since the images have not been contrast enhanced or whitened in any way.

A second method of calculating signal to noise is to apply

$$S / N = 20 \text{Log} \left[\frac{\mu_{S+N} - \mu_N}{\sigma_N} \right]$$

where the plume is delineated by contiguous pixels that exceed a threshold value of $t = \mu_N + 2.3 \sigma_N$. While most commonly used to calculate electronic S/N, it can be useful in images as well. Here the values tend to be larger and remain positive assuming the area of interest is brighter than the background.

2.8.2 ROC curves

The Receiver Operating Characteristics (ROC) Curve has been widely used in range of applications from structure response to the medical efficacy of drugs since its original implementation in providing a quantitative measure of radio signal transmission characteristics. The curve is a plot of the probability of detection versus the probability of false alarm or the probability of true event being registered versus the probability of an event that registers as true, but is not really true. More importantly in the field of hyperspectral imaging it can be interpreted as a comparative measure of sensitivity vs. (one minus specificity) for a particular measurement and detection strategy. Typical curves are shown in Figure 6 above.

The ROC curve analysis relies critically on the selection of a “truth” region in the image

which can be generated in any number of ways depending upon the data characteristics. If ground truth data is available it can provide the most reliable method. If it is not (the usual case), then simple amplitude threshold or image intensity threshold criteria can be applied using either pre-selected or statistically based methods (e.g. 2σ above the average background). It can also be done by manual selection of areas of potential interest and iterating on that selection for the best results. For uniform results the data images are normalized by simple minimum to maximum scaling to remove the offset bias.

Probability histograms are then generated from the selected data and the generated truth mask and plotted against the normalized amplitude. This is done by counting the number of true data points and the number of false data points that fall into specific amplitude bins and plotting as a function of relative amplitude. From the probability distributions, the integrated probability curves are generated by summing all the truth and then the false counts below a moving threshold and plotting the integrated probability vs. the relative threshold amplitude. The ROC curves are generated by plotting the integrated truth values against the integrated false values.

Interpretation of ROC curves is straightforward in many cases. In general, the greater the area above the 45 degree diagonal, the better the detection method should be. ROC curves that lie on the 45 degree line are neutral, equally useful and useless representing 50/50 probabilities. Those curves that fall below the line are generally less than useful. Comparative results may be straight forward or difficult to interpret.

Bottom line is that we found no consistent, non subjective way to compare results exactly, but the two methods above represent reasonable methods in many cases.

2.9 Bayesian Approaches

Some effort was expended in the program to determine the extent to which multiple registered images could improve plume detection and gas concentration estimates. The approach taken was to extend existing Bayesian single pixel regression models to incorporate both spatial and temporal information from the multiple images. Considered were stochastic models for the spatial/temporal structure of the background clutter and for models the propagating gas plumes recognizing the registered background clutter will be relatively static with respect to the spatial variations and the plumes will not. This assumes that the thermal background radiance remains relatively constant while the plume has some drift and shape evolution. For longer time period it might be necessary to perform a temperature/emissivity separation to account for variation in the background radiance which can be quite significant¹².

Unfortunately the real data cubes that were available for comparison could not be registered well enough in the time available to conduct a conclusive quantitative

¹² Patrick G. Heasler, Michael G. Foley, and Sandy E. Thompson, 2007. *Consequences of Mixed Pixels on Temperature Emissivity Separation*, PNNL-16330, Pacific Northwest National Laboratory, Richland, WA.

investigation, so some simplifying assumptions were made. Presented here is a brief summary of the basic methodology and the reader is referred to an earlier report¹³ for a detailed explanation.

The general approach using Bayesian Regression is to supply information on the clutter parameters in the form of a prior distribution. It is assumed that neighboring pixels are similar to one another and the spatial parameters can be used to predict the clutter parameters better than with a single image. The trick is then to look for spatial/temporal variations between neighboring pixels to locate and characterize the plume. The general regression model takes the form at the pixel location x at time t

$$L(x, t) = F(C(x, t), \beta(x, t)) + e$$

where F is the physics based radiation transfer model, C is a vector of plume gas concentrations, β is a ground clutter vector and e is the random error term. In practice, this is a bit more complex and x will be a weighted average of similar nearest neighbor pixels depending on the model and the prior distributions available used.

One can build a standard ground radiance model in the form of

$$L_g = F(\beta)$$

where L_g is the radiance from the ground and F is the transformation that expresses L_g in terms of the independent normal variables in terms of prior β which describes the variability in the ground radiance. If the background radiance changes as a function of time then it is necessary to perform a similar operation on the temperature/emissivity deconvolution to refine the prior for β .

From these basics an iterative algorithm similar to a Hidden Markov Modeling Algorithm is used to fit the regression using the more complicated priors. The steps are a) to obtain initial estimates of C and β from a fit of the regression model using a single prior derived from the whole image, then b) estimate parameters at a pixel using the conditional prior with neighborhood values obtained in a), then repeat the process until the estimated converge.

One can develop a stochastic model for the background clutter by assuming the spatial structure for β is Gaussian and stationary over time and further assuming the low frequency components are both inconsequential compared to the higher frequency components in the pixel neighborhoods and linked to the basic ground radiance model for clutter estimation. This process effectively smoothes the clutter into distinct regions of

¹³ Patrick Heasler, Sandy Thompson, Mike Foley, *MIST Task 1 Report: Bayesian Algorithms for Multiple Image Plume Estimation*, PNNL-17160, December 2007, Pacific Northwest National Laboratory, Richland, WA

stochastic similarity which logically segments the image without forcing it to be stationary (relatively independent of time) over its entirety.

With minimal ancillary data such as the wind vector one can then construct simple triangular neighborhoods for estimation of plume origins assuming a gas mixture and calculate a Bayesian model average using the regional neighborhood parameters for the prior, BY assuming the wind direction and applying a Markov random field one can enhance convergence. Details and examples can be found in reference 7. The effect is to enhance the probability of detection (POD) of weak plumes assuming they are emitted from the same point in successive images using the Bayesian prior.

Figure 18 a comparative example of the POD as a function of plume length for a weak plume in a 10,000 pixel image where the probability of false detection has been set to 1% and the signal to noise ration has been set to 1. The scenarios are represented no spatial information (independent pixel prior) is used similar to the classical statistical analysis based on the whole scene, b) spatial information is included and c) spatial and temporal information is included.

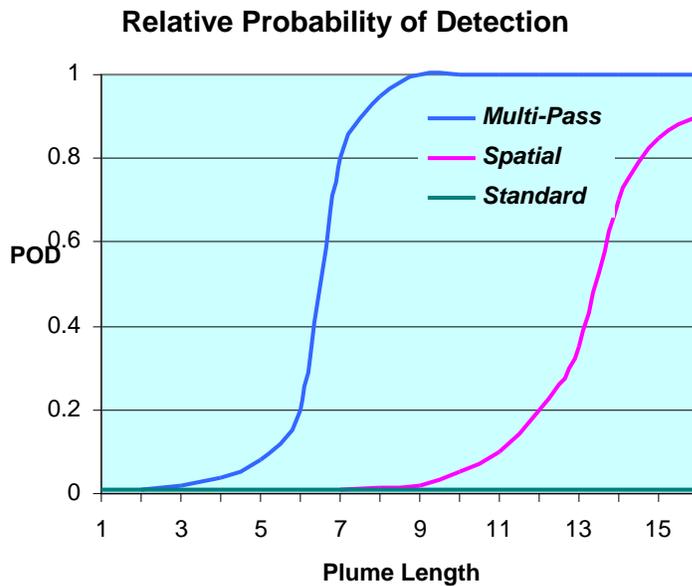


Figure 18. Relative probability of detection verses plume length for a) standard BMA, b) spatial BMA, and c) spatial and temporal BMA.

3 Physics Based Models Investigated

Another potential method of reducing clutter is to compensate for a major portion of it by careful modeling of the scene and subtracting its effects. This procedure could provide a first estimate for clutter subtraction or provide prior for the Bayesian Regression models discussed above or just provide cleaner images for other statistical processing applications. This disadvantage of this approach is of course the extensive time it takes

to construct a detailed enough site model from aerial photographs and ancillary data (e.g. high resolution LIDAR, etc) or detailed site drawing. Still it could be useful for frequently revisited sites. As a trial application one industrial site for which we were able to obtain extensive background data was selected for detailed evaluation of this hypothesis.

3.1 Model construction

Figure 19 shows a high resolution aerial photograph of the site selected for demonstration of the concept. This chemical manufacturing site was the subject of hyperspectral campaign several years ago. Figure 20 shows the final physical model that was constructed this portion of the site. The model was constructed via a combination of established automated commercial sector software and a fair amount of hand stitching of the nodes. Additional tasks included assigning representative materials to the discrete features. Note that the perspectives in the figures not the same, but the roadways are easily identified features that can assist the reader in orienting the views.



Figure 19. *Aerial view of the industrial chemical production site that was modeled.*

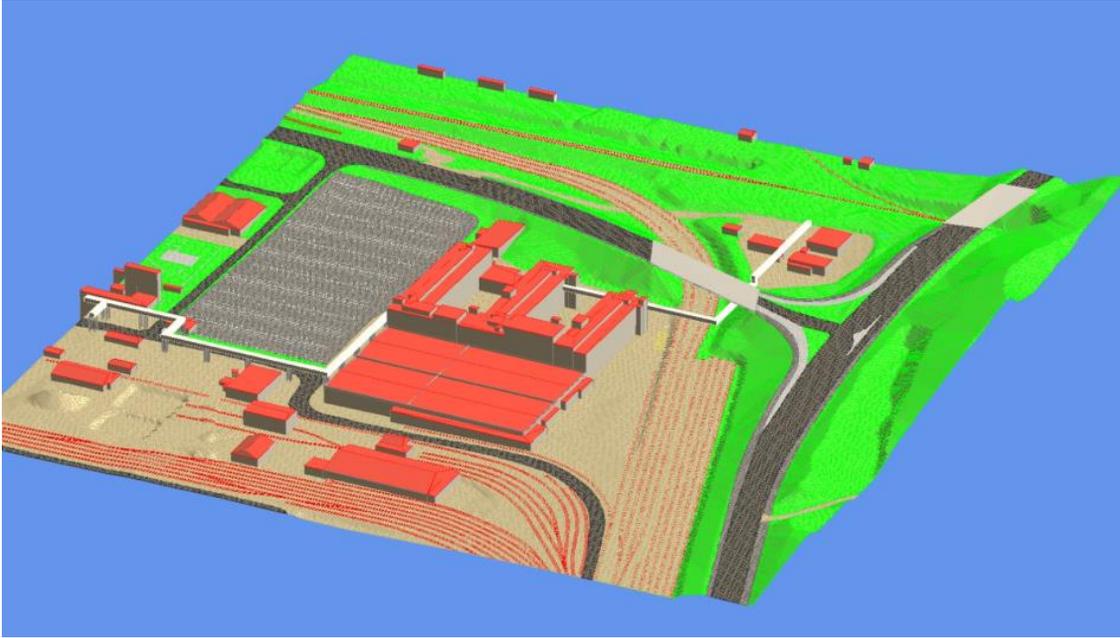


Figure 20. *The site model that was constructed.*

The physical model was then imported into MuSES, a commercial software package from ThermoAnalytics, Inc, of Calumet, MI. The software performs an energy balance using convection, conduction, and radiation taking into account solar loads, sky and earth emission and weather effects. From this data a radiance model is constructed which is time and weather dependent. One such model frame is shown in Figure 21.

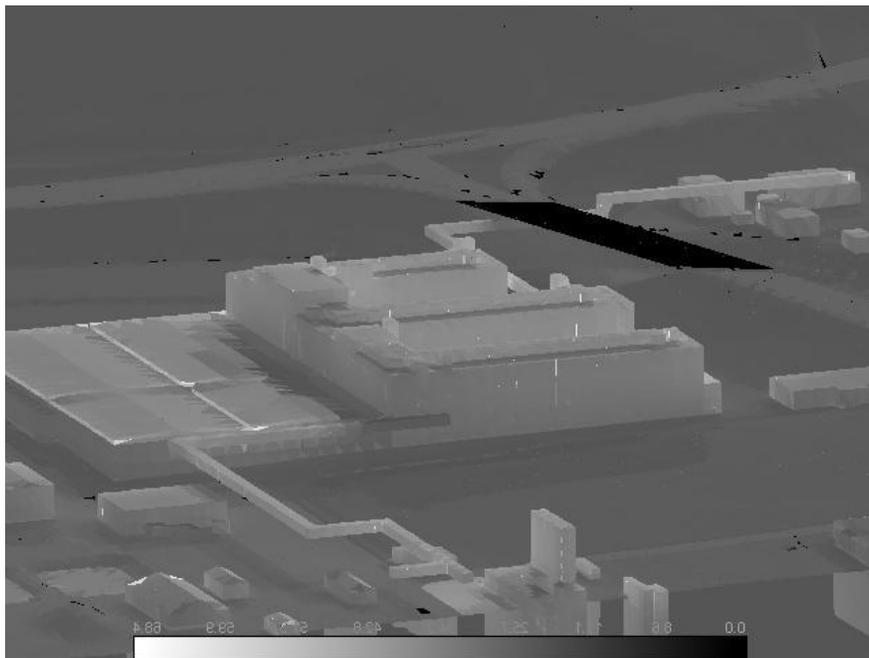


Figure 21. *The broadband radiance from the MuSES model of the site.*

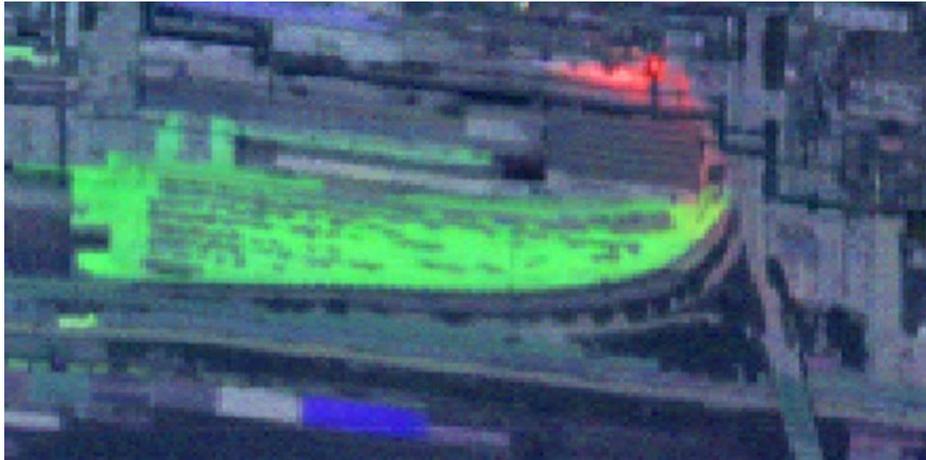


Figure 22. *Composite RGB image of the hyperspectral data from a portion of the site showing a plume in red.*

With the model in hand it should have been a straight forward exercise to map the actual hyperspectral data onto the model for use in either a first blush covariance matrix or as the prior in a Bayesian regression calculation, but it proved more difficult than hoped with the hyperspectral data in hand. Figure 22 is a composite RGB image of the actual analyzed hyperspectral data of a portion of the site modeled which shows a real chemical plume in red. The green area is residual materials left along the rail spurs. Unfortunately there were insufficient resources available to complete the required pixel to pixel image registration for the highly distorted hyperspectral images to effectively illustrate the value of the hypothesis and show analysis advantage. The precision registration issue continues to be a significant hindrance to this approach as discussed in the next section.

3.2 Registration

As mentioned above one of the key element in using physics based modeling to assist in the elimination of background clutter is the registration of the actual sensor image with the model. The motion (ground, atmospheric, platform, etc.) and distortion (thermal, optical, spectral, etc.) of the sensor image, particularly in push broom implementations, often makes this a difficult task, especially with low altitude mobile sensing platforms. The task usually involves either picking an entity in one image and determining its corresponding location on another or picking an object and determining its location. PNNL worked on creating an easy to use interface to the second approach¹⁴ based on

¹⁴ MG Foley, RJ Carter, HP Foote, LH Gerhardstein, SA Schulze, LR Stoops, SE Thompson, *PNNL Precision Registration System*, PNNL 15376, September 2005, Pacific Northwest National Laboratory, Richland, WA

standard photogrammetric principles that determine the relative geometries and orientation of modeled objects and sensors. Here the mapping is done using a selected set of parameters such as interior orientation (IOR-horizontal and vertical size of image, field of view of the imager) and the exterior orientation (XOR-three spatial coordinated, three angular orientation values). This is easily accomplished with static and stable images and platforms by using lines and points to automatically extract feature lines for matching. However the process becomes much more complicated with time evolved images and platforms such as those illustrated above where the platform is buffeted by atmospheric turbulence and the flight path of the platform is subject to dynamic pitch and yawl where the lines become disjointed and overlapping curves and the point scattered.

We believe the problem is still manageable, but beyond the scope of this immediate project. Because of the nature and magnitude of the required development work, the scope was shifted to NGA for funding.

4 Example Results

Among the last sets of data examined were the FY07 tests performed at NTS. Here an LWIR hyperspectral sensor was ground based. Various gases pertinent to nuclear weapons production were simultaneously released in varying concentrations against the background of an industrial-like setting. Detailed results are available but too voluminous for inclusion here, however the images below give an indication of the observed results.

Figure 23 shows the analysis results for five gasses comparing the 3D-FFT approach without spatial filtering to the more standard WMF approach using the entire image for whitening in both cases. Only three gases were released at this time period. The relative release rates for the gases are indicated by the bar graphs on the right. The photograph on the left is an observer's view of the actual scene. The ellipses highlight the highest signal-to-noise level for the observed gas plumes which are either dark or light against the background depending on the temperature contrast. The short diagonal region on the lower left portion of the image is not a plume, but rather a building roof structure feature that is bright in many of the analysis images. In this case the 3D-FFT approach proved to be slightly better than the WMF approach. The dispersion pattern of each gas is slightly different as expected. Each gas is individually separated out of the mixture.

Figure 24 is a similar construct showing a different gas and concentration mixture. Here the WMF approach has the advantage over the 3D-FFT approach without spatial filtering in two of the three detections. Again the gases are easily separated.

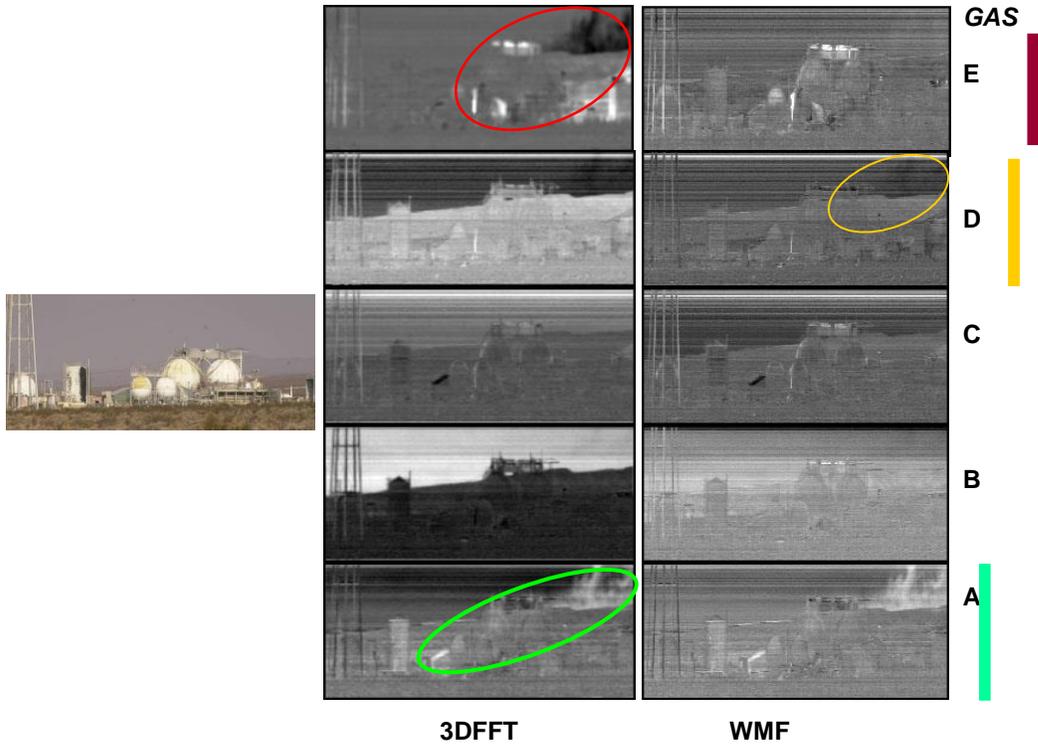


Figure 23. Ground based LWIR HSI analysis images for three simultaneous gases using 3D-FFT (no spatial filtering) and the more traditional WMF analysis.

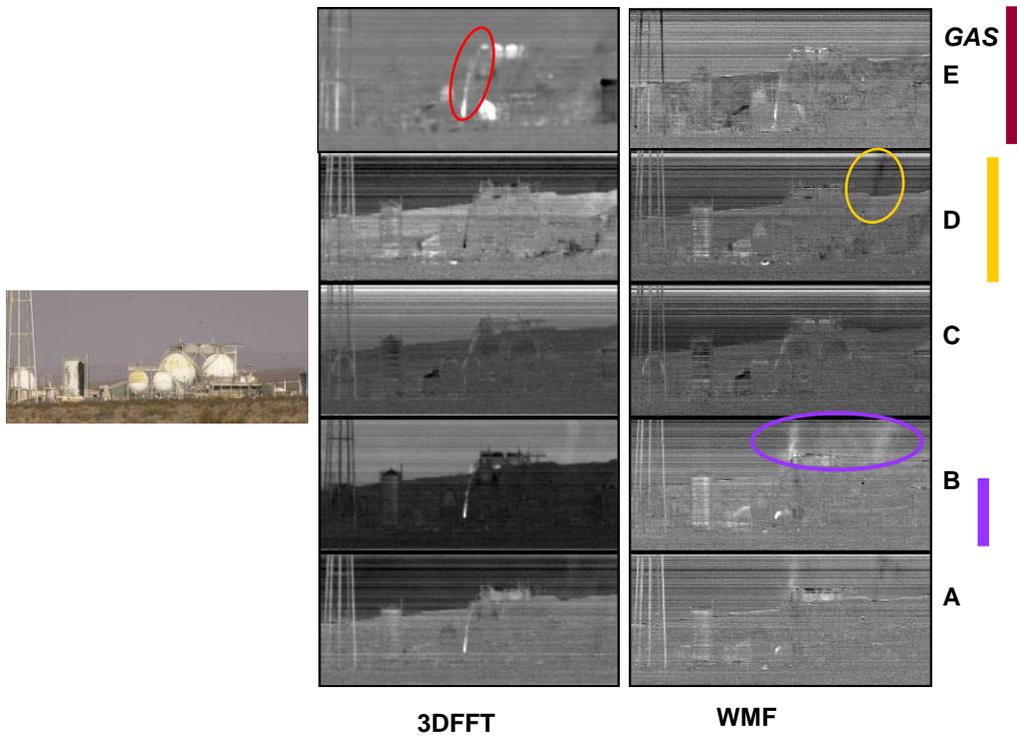


Figure 24. Ground based LWIR HSI analysis images for three simultaneous gases using 3D-FFT (no spatial filtering) and the more traditional WMF analysis.

There are additional analysis benefits to be gained by appropriate selection of specific regions as illustrated in Figure 25. Two different gas releases (upper grouping and lower grouping) are represented with the same scene background as shown in Figure 23 and Figure 24. The red outline overlaid on the eight individual analysis images in each grouping is area that was used for covariance matrix whitening. Each area has a different effect on eliminating background clutter and highlighting the plume. Clearly areas outside of the plume and most representative of the clutter under the plume have the largest effect. The highest signal to noise for the plume are highlighted by the yellow ellipses.

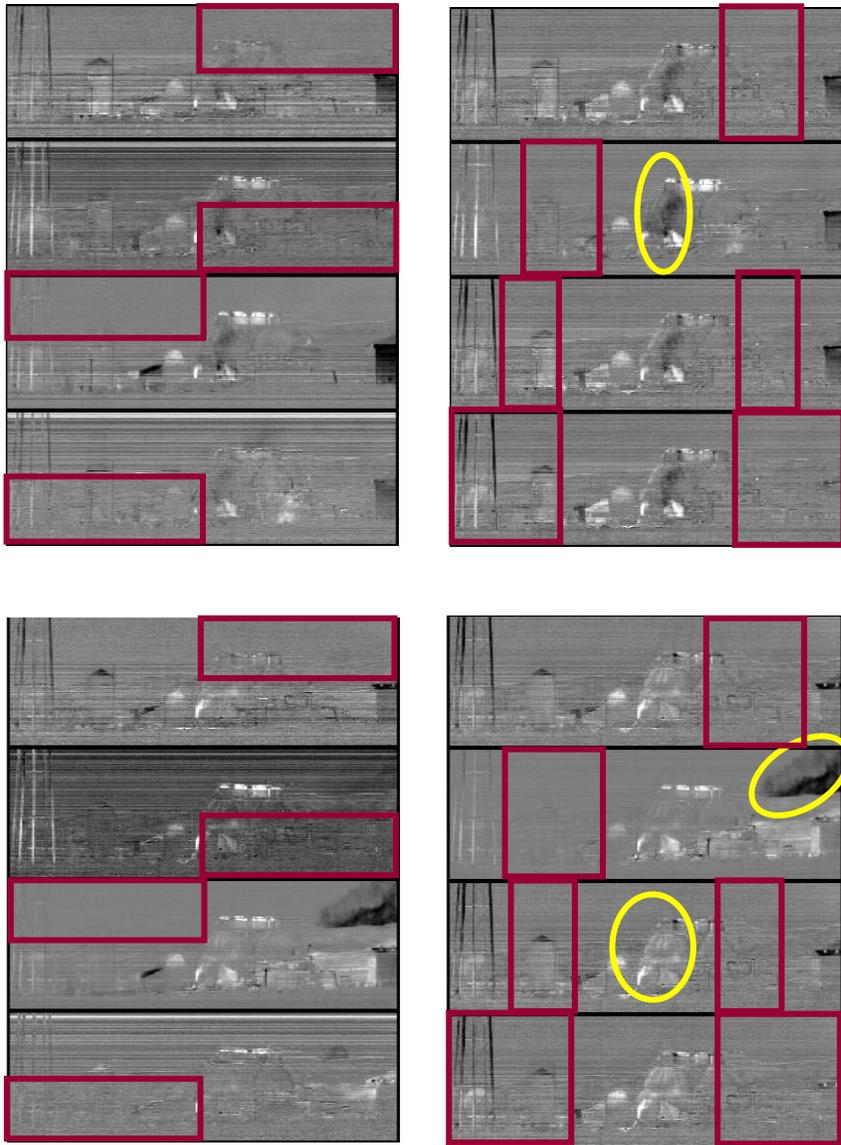


Figure 25. Whitened images of two gas releases (upper and lower groupings) using the regions highlighted in red as the whitening area for clutter mitigation.

5 Summary and Recommendations

The MIST program made extensive explorations of both improving old techniques and exploring new methods for analysis of gas plumes from LWIR hyperspectral sensors. The main contributions to date have been the recognition that significant signal to noise enhancement via background clutter mitigation can be achieved by using selective pieces of the background for whitening the data in an iterative approach. This is true for both aerial and ground based sensor data.

Since the program was terminated a year prior to plan several of the final integration pieces were left incomplete. An easy to use comparative and iterative interface for applying all the tools for an optimized analysis result which was scheduled for the final year of the project was not completed. The manual results showed the clear benefit of the approach. Precision registration of images (first on a pixel and then on a sub-pixel level) with physics based models of the scene to help mitigate clutter was initiated under the program started but not achieved due to poor image quality. A more robust automatic registration process or short exposure, 3-D focal plane data cubes are required to explore the technique. An evaluation of the value of the approach is therefore still pending and should be completed in the future, but the effort required will be substantial either on the hardware or software side. Lack of precision registration has also hampered the integration of multiple images to capitalize on statistical advantage of time evolution and multiple images of the same scene.

While the MIST program represents a step forward, the fusion of other measurement modalities should also be explored. More orthogonal vectors added to the data enhance the statistics even further allowing better determinations of gas types, concentrations, volumes, release rates and measurement confidence.

6 Papers, Publications, Presentations & Software

The following is list of publications, presentations and software that are a product of the program.

Michael A. Lind, *Software for Generating ROC Curves for Hyperspectral Data Cubes and Flat Images*, PNNL-17243, December 2007, Pacific Northwest National Laboratory, Richland, WA

Michael A Lind, Harlan Foote, Patrick Mackey, Patricia Medvick, and Lisa Nuffer, *Using the frequency Domain for Plume Detection: Whitened 3D_FFT compared to WMF*, presented at SAEF/SMUG#12 November 14-16,2007, Chantilly, VA PNNL-SA-57957

Patricia Medvick, Harland Foote, Lisa Nuffer and Michael Lind, *Whitening Strategies for Enhanced Chemical Detection*, presented at SAEF/SMUG#12 November 14-16,2007 Chantilly, VA PNNL-SA-57956

Patrick Heasler, Sandy Thompson, Mike Foley, *MIST Task 1 Report: Bayesian Algorithms for Multiple Image Plume Estimation*, PNNL-17160, December 2007, Pacific Northwest National Laboratory, Richland, WA

Patricia A, Medvick, Michael A. Lind, Patrick S. Mackey, Lisa L, Nuffer, Harlan P. Foote, *3D-FFT for Signature Detection in LWIR Images*, PNNL 17105, November 2007, Pacific Northwest National Laboratory, Richland, WA

Michael A Lind, Harlan Foote, Patricia A. Medvick, *Comparative Analysis of Gas Plumes Detection in Simulated LWIR Hyperspectral Data Cubes Using Variations of PCA, WFM and Entropy*, PNNL-SA-57144, August 2007, Pacific Northwest National Laboratory, Richland, WA

Michael A Lind, Harlan P. Foote, Patricia A. Medvick, *Comparative Analysis of Gas Plumes Using PCA, WFM, ICA, and Entropy on Artificial LWIR Hyperspectral Data Cubes*, PNNL-SA-56072, June 2007, Pacific Northwest National Laboratory, Richland, WA

Patricia A Medvick, Lisa L Nuffer, Harlan P. Foote, Michael A. Lind, *Adaptive Whitening for Rapid Processing of Hyperspectral Images*, PNNL-16658, May 2007, Pacific Northwest National Laboratory, Richland, WA

Michael A. Lind, Jonathan L. Barr, *Modeling and Software Scene Simulation for Long Wavelength Infra-Red Hyper-Spectral Dada Cubes*, PNNL-17213, 2007, Pacific Northwest National Laboratory, Richland, WA

Michael A. Lind and Jonathan L Barr. 2007, *Modeling and Software Scene Simulation for Long Wavelength Infra-Red Hyper-Spectral Imaging Ground Clutter* , PNNL-SA-55971 Pacific Northwest National Laboratory, Richland, WA.

Patrick G. Heasler, Michael G. Foley, and Sandy E. Thompson, 2007. *Consequences of Mixed Pixels on Temperature Emissivity Separation* , PNNL-16330, Pacific Northwest National Laboratory, Richland, WA.

Andrea S. Renholds, Sandy E. Thompson, and SJ Walsh. 2007. *Metric Literature Review 2007*. ,. PNNL-SA-54694. Richland, WA

Patrick Heasler, Christian Posse, Jeff Hylden and Kevin Anderson, *Nonlinear Bayesian Algorithms of Gas Plume Detection and Estimation form Hyperspectral Thermal Image Data*, *Sensors* 7(6) 905-920, 2006

Thompson SE, and MG Foley. 2006. *Improvements in Spectral Processing Science*. Presented by Sandra E Thompson (Invited Speaker) at NA-22 meeting, New York City, NY on December 12, 2006. PNNL-SA-53118

Andrea Renholds, *Error Propagation through MIST Algorithms*, April, 2006

Renholds AS, and JC Solinsky. 2005. *Physical Model Overview and Nonlinear Bayesian Regression for MIST* . PNNL-15421, Pacific Northwest National Laboratory, Richland, WA.

MG Foley, RJ Carter, HP Foote, LH Gerhardstein, SA Schulze, LR Stoops, SE Thompson, *PNNL Precision Registration System*, PNNL 15376, September 2005, Pacific Northwest National Laboratory, Richland, WA

JC Solinsky, HP Foote, PA Medvick, and LL Nuffer, *Sample Location Using Independent Component Analysis (ICA)*, TIE2005 Conference presentation, Sandia National Lab, (4/13/05)

JC Solinsky, *Detection and Estimation Using Models of Real World "Object" Measurements*, PNNL 14847, November 2004

JC Solinsky, HP Foote, *Detection Metrics in Image Analysis*, PNNL-14795, August 2004, Pacific Northwest National Laboratory, Richland, WA

Solinsky JC, MG Foley, HP Foote, LH Gerhardstein, LR Stoops, SK Wurstner, and AF Schenk. 2003. *Assisted, 2-1/2D Registration of Imagery to Ground-Based Objects for Deriving Measurements from Multi-Spectral Imagery*. PNNL-SA-38780, Pacific Northwest National Laboratory, Richland, WA

Solinsky JC, MG Foley, HP Foote, LH Gerhardstein, LR Stoops, and SK Wurstner. 2003, , PNNL-SA- *Assisted 2-D Registration of Imagery to Ground –Based Objects for*

Deriving Measurements for Multi-Spectral Imagry PNNL-SA-38564, Pacific Northwest National Laboratory, Richland, WA. Michael Lind 2006

SOFTWARE

A general description of the software developed under the program is referenced in the publications above. Software is available upon request either as an executables or as source codes.

Spectral Binning v4. This program quickly estimates the effects of spectral binning errors and calibration errors in LWIR detectors taking into account the gas being selected and the intervening atmosphere to establish base maximum accuracy numbers. Michael Lind 2005

LWIR FILE Conversion Utility v4. This program converts material, atmospheric, albedo, and other high resolution long wave infra red spectra into detector apodized data files. Multiple apodization schemes, detector profiles and I/O file formats are accommodated. Michael Lind 2006

LWIR Scene Generator v14. This easy to use program generates simulated random background NADIR LWIR data cubes taking into account the full range of atmospheric conditions, plume gasses and temperatures, ground albedos and temperatures. A variety of plume shapes, concentration profiles, background shapes, background clutter statistics and detector profiles are possible. These artificial data cubes are realistic representations of a variety of ground clutter situations with full ground truth. They are used for baseline comparisons of various analysis techniques. The program includes analysis routines using FFT, Entropy, Covariance Matrix Eigenvalues. Michael Lind 2007.

IROC Curve Generator v8. This program generates receiver operating characteristic (ROC) curves and calculates signal to noise from flat images and data cubes. It calculates the image statistics and has provisions for both automatic and manual determination of the “truth” region in the images. Michael Lind 2007.

pca. This program calculates principal components and has been combined with ***runICA.*** Lisa Nuffer 2007

runICA. This programs performs whitening (with and without region of interest masking for plume enhancement) and independent component analysis on plain text hyperspectral data files and has an option to calculate the eigenvalues of the covariance matrix. It uses the CLAPACK and BLAS math libraries Lisa Nuffer, Harlan Foote and Pat Medvick 2007

wmfilter. The original program developed in 2004 is a command line and supervision driven C code that displays bitmaps of matched filter results. The new version finished in 2007 fits into ***TestEngine1*** and runs without supervision. The LAPACK library is used extensively. Larry Gerhartstein, Patrick Mackey and Pat Medvick.

bsa. This program, still in the early development process is used for blob identification and analysis via attribute vectors to identify regions of interest. Larry Gerhardstein, Patricia Medvick 2006.

3DFFT-MF. This program applies a whitened matched filter in the Fourier domain. The program makes extensive use of the FFTW library which requires open source distribution. The matched filter is constructed from a specified atmospheric compensated and apodized gas spectra and applied to each pixel of the data cube. Spatial filters that average over a specified observation diameter can also be applied as well as contrast adjustment. Color can be added using a blackbody color map and the output saved as a JPEG. Patrick Mackey and Harlan Foote 2007.

TestEngine1. This program still in its early phases is the glue that binds all the above elements together either via batch processing files or a GUI interface allowing comparison calculations and controlling output display formats. Patrick Mackey and Patricia Medvick 2007.