# Flag-based detection of weak gas signatures in long-wave infrared hyperspectral image sequences

Timothy Marrinan[a], J. Ross Beveridge[b], Bruce Draper[b], Michael Kirby[a,b], and Chris Peterson[a]

[a]Department of Mathematics, Colorado State University, Fort Collins, CO, 80523 USA.
[b]Department of Computer Science, Colorado State University, Fort Collins, CO, 80523 USA.

## ABSTRACT

We present a flag manifold based method for detecting chemical plumes in long-wave infrared hyperspectral movies. The method encodes temporal and spatial information related to a hyperspectral pixel into a flag, or nested sequence of linear subspaces. The technique used to create the flags pushes information about the background clutter, ambient conditions, and potential chemical agents into the leading elements of the flags. Exploiting this temporal information allows for a detection algorithm that is sensitive to the presence of weak signals. This method is compared to existing techniques qualitatively on real data and quantitatively on synthetic data to show that the flag-based algorithm consistently performs better on data when the $\text{SINR}_{\text{dB}}$ is low, and beats the ACE and MF algorithms in probability of detection for low probabilities of false alarm even when the $\text{SINR}_{\text{dB}}$ is high.

**Keywords:** Grassmann manifold, SVD, Flag mean, Flag manifold, Hyperspectral Images, Gas Detection, Remote sensing.

## 1. INTRODUCTION

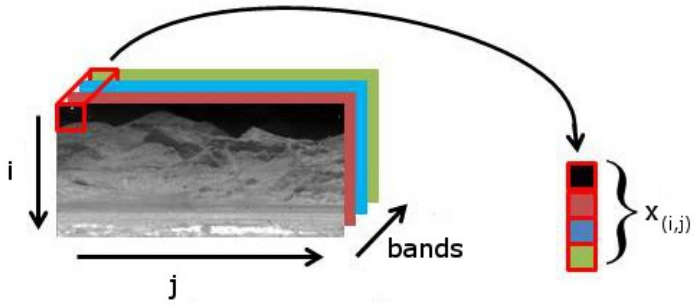
The problem of detecting chemical plumes in long-wave infrared hyperspectral images has grown increasingly popular in the last decade. Military and civilian applications are abundant as can be inferred from the variety of disciplines that research the topic.[1–5] Most algorithms for chemical plume detection, such as matched filters,[6] normalized matched filters,[7] and subspace detectors,[8] use information from a single hyperspectral image to perform their detections. When available, these methods utilize supervised estimation of the plume-free background in order to improve the accuracy of their detections. For images generated by ground-level sensors, the problem of estimating the radiance of the background clutter is exacerbated by the unequal distances between objects in the scene and the sensor itself.

One unsupervised method for estimating the plume-free background of an image is to keep the sensor in a static location and image the scene at a time known to be free of chemical agents. This method alleviates the difficulty of compensating for differences in path-length between objects and the sensor, and ensures that the background estimate will not be contaminated with the chemical being detected. Using a static sensor to repeatedly image the same scene is referred to as a 'persistent stare' application and can be useful for monitoring emissions from industrial facilities or as an early warning of dangerous chemicals near cities or sensitive government buildings.[4,5,9,10] What this method does not immediately account for are changes in lighting and temperature due to the change in time between the background image and the image being used for detection. These physical changes affect the at-sensor measurements, and need to be taken into account by any technique that wishes to make use of this type of background estimate.

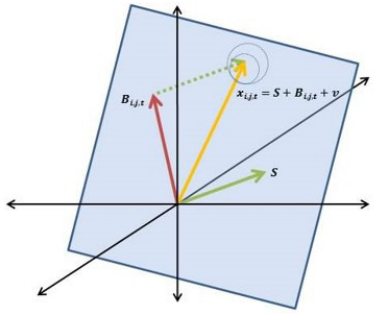
In this paper, we propose a method for detecting chemical plumes that utilizes temporal information to estimate the plume-free background of a hyperspectral image. With this information, the technique creates a flag, or nested sequence of linear subspaces, for each hyperspectral pixel in an image that contains elements which span the background clutter, the changes in illumination and temperature, and the signature of the chemical agent of interest. This flag is created using a technique referred to as the flag mean for subspaces of differing dimensions.[11] By measuring the similarity between the laboratory signature for a chemical and the representative

Further author information: (Send correspondence to T. Marrinan) E-mail: marrinan@math.colostate.edu

Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXII
edited by Miguel Velez-Reyes, David W. Messinger, Proc. of SPIE Vol. 9840, 98401N
 · doi: 10.1117/12.2224117




(a) Each pixel in a long-wave infrared hyperspectral image corresponds to a vector whose elements are the response of that pixel to different wavelengths of infrared light.




(b) The linear model for the observed spectrum of pixel $x_{i,j,t}$, most simply consists of the sum of the gas spectrum ($S$), the background of that pixel ($B_{i,j,t}$) and Gaussian noise ($\nu$).

Figure 1: Pixels of hyperspectral images correspond to vectors that can be decomposed into distinct endmembers.

flags, we determine a scalar statistic that predicts the presence of the chemical at each pixel in an image. This method of *including* background information in the model for each pixel, rather than projecting away from it, provides a sensitive detector that appears to provide improved detection accuracy for weak or optically thin chemical plumes.

There are numerous thorough and well written introductions to hyperspectral image processing available for reference.[1,12] Thus we review only the details relevant for the technique at hand. The most widely used spectral model is the linear mixing model.[13] This model assumes that the observed spectrum of a pixel consists of a linear combination of distinct spectral endmembers. In other words, pixel $(i,j)$ is represented by a vector $x_{i,j} \in \mathbb{R}^b$ where $b$ is the number of spectral bands captured by the imaging device. Then we can write $x_{i,j} = S\alpha + B_{i,j} + \nu$ where $S$ is the spectrum of the gas, $\alpha$ is the scaling factor representing the amount of gas, $B$ is the spectrum of the background at that pixel location, and $\nu$ is white noise. Alternatively, either $S$ or $B_{i,j}$ can be a matrix whose columns correspond to endmembers. In this case, the matrices would be multiplied by a unit length vector that provides a convex combination of these endmembers as well. Figure 1 shows an illustration depicting the correspondence between a pixel in the scene and a linear combination of spectral endmembers.

Two of the detection algorithms in the literature that assume a linear mixing model are the adaptive cosine (or coherence) estimator (ACE)[8] and the matched filter (MF).[14] The ACE algorithm has become particularly popular in practice because of it broad applicability, simple implementation, and speed.[15] Assume that $s, x, \hat{\mu} \in \mathbb{R}^b$ such that $s$ is the spectrum of a target signature, $x$ is the spectrum of a test pixel, $\hat{\mu}$ is the sample mean of the pixels, and $\hat{\Gamma} \in \mathbb{R}^{d\times d}$ is the sample covariance of the mean subtracted pixels. The ACE detection statistic is then computed as

$$D_{\text{ACE}}(x) = \frac{(s-\hat{\mu})^T\hat{\Gamma}^{-1}(x-\hat{\mu})}{\left((s-\hat{\mu})^T\hat{\Gamma}^{-1}(s-\hat{\mu})\right)^{\frac{1}{2}}\left((x-\hat{\mu})^T\hat{\Gamma}^{-1}(x-\hat{\mu})\right)^{\frac{1}{2}}}, \tag{1}$$

which takes a value of 1 if the whitened target spectrum and the whitened pixel are collinear, and a value of 0 if they are orthogonal.

Alternatively, the matched filter is computed as

$$D_{\text{MF}}(x) = \frac{(s-\hat{\mu})^T\hat{\Gamma}^{-1}(x-\hat{\mu})}{(s-\hat{\mu})^T\hat{\Gamma}^{-1}(s-\hat{\mu})}. \tag{2}$$

Thresholding these statistics allows us to compute a binary detection mask on a hyperspectral image. However, to determine the possible effectiveness of this thresholding we report empirical receiver-operator-characteristic

(ROC) curves and area under the curve (AUC) scores. Of particular interest is the front end of the ROC curves, where the false positive rate is low.

## 2. MATHEMATICAL BACKGROUND

For the flag-based detection algorithm, we view the observed spectrum of a pixel as a sample from a subspace of $\mathbb{R}^b$ with dimension greater than 1. This could be a plane if the target spectrum and background spectrum are uniquely determined by their materials, but since this is rarely the case[12,16] they may be higher dimensional. Identifying pixels with the subspaces from which they are samples, allows the data to be treated as a collection of points on a Grassmann manifold. The Grassmann manifold $\mathrm{Gr}(k, \mathcal{V})$ is a manifold whose points parametrize the subspaces of dimension $k$ inside the vector space $\mathcal{V}$. In this document, we will assume that $\mathcal{V}$ is a finite dimensional real vector space and thus we can identify $\mathcal{V}$ through its dimension, $n$. For the remainder of this paper, we denote by $\mathrm{Gr}(k, n)$ the Grassmann manifold of $k$ dimensional subspaces of $\mathbb{R}^n$, $GL(k)$ denotes the general linear group of invertible $k \times k$ matrices and $O(k)$ denotes the orthogonal group of $k \times k$ orthogonal matrices.

Let $\mathbb{R}^{n\times k}$ denote the vector space of $n \times k$ matrices with real entries and let $(\mathbb{R}^{n\times k})^\circ$ denote the open submanifold of full rank $n \times k$ matrices. For each $Y \in (\mathbb{R}^{n\times k})^\circ$, let $[Y]$ denote the column space of $Y$. There is a natural surjective map $\phi : (\mathbb{R}^{n\times k})^\circ \to Gr(k, n)$ given by $\phi(Y) = [Y]$ (with $[Y]$ identified with its corresponding point on $Gr(k, n)$). It is clear that $\phi(X) = \phi(Y)$ if and only if there exists an $A \in GL(k)$ such that $XA = Y$. Thus a point $q$ on $Gr(k, n)$ corresponds to a $k$-dimensional subspace $V_q$ of $\mathbb{R}^n$ and can be represented by any element of a $GL(k)$ orbit of a full rank $n \times k$ matrix, $Y$, whose column space $[Y]$ is equal to $V_q$. For the purposes of computation, we utilize a representative with orthonormal columns (note that if $Y$ is a representative with orthonormal columns and if $B \in O(k)$ then $YB$ will be another representative with orthonormal columns). Since a matrix representative, with orthonormal columns, for a point on a Grassmann manifold is only unique up to right multiplication by an orthogonal matrix, it is important that the output of any algorithm is invariant to such a multiplication.

Let $d : \mathrm{Gr}(k, n) \times \mathrm{Gr}(k, n) \to \mathbb{R}$ be a metric. The metric, $d$, is said to be *orthogonally invariant* if for every $[X], [Y] \in \mathrm{Gr}(k, n)$ and every $A \in O(n)$, $d([X], [Y]) = d([AX], [AY])$. The two metrics of interest in this paper are the geodesic distance based on arc length (the canonical metric on the Grassmann manifold), and the projection Frobenius norm. As shown by Edelman *et al.*, the geodesic distance based on arc length can also be computed as the $\ell_2$-norm of the vector of principal angles between $[X]$ and $[Y]$, that is $d([X], [Y]) = \|\Theta\|_2$.[17] The projection Frobenius norm arises from the identification of points in $\mathrm{Gr}(k, n)$ with $n \times n$ projection matrices of rank $k$ and taking the the Frobenius norm of the difference between these representations. It can be equivalently and more efficiently computed as the $\ell_2$-norm of the vector of the sine of the principal angles, $d_{pF}([X], [Y]) = \|\sin\Theta\|_2$.[17] In both cases, if $X$ (resp. $Y$) are orthonormal matrix representatives for $[X]$ (resp. $[Y]$) then the cosines of the principal angles between $[X]$ and $[Y]$ are the singular values of $X^TY$.[18] Note that if $A, B \in O(k)$, then the singular values of $X^TY$ are the same as the singular values of $(XA)^T(YB)$.

### 2.1 The flag mean

The subspaces created from the span of hyperspectral pixels may often be subspaces of differing dimensions lying in a fixed ambient space. This set of subspaces lives naturally on a collection of Grassmann manifolds rather than on a single Grassmann manifold. Suppose that $[X] \in \mathrm{Gr}(k_1, n)$ and $[Y] \in \mathrm{Gr}(k_2, n)$ for $k_1 < k_2$. As illustrated in Bjork and Golub's foundational paper, there will be $k_1$ principal angles between $[X]$ and $[Y]$[18] and we define $d_{pF}([X], [Y])$ as the $\ell_2$-norm of the vector of the sines of the $k_1$ principal angles between $[X]$ and $[Y]$. Note that $d_{pF}$ is no longer a metric due to the possibility of $d_{pF}([X], [Y]) = 0$ while $[X] \neq [Y]$ (for instance, if $[X]$ is a proper subspace of $[Y]$).

Let $\boldsymbol{q} = [q_1, q_2, \ldots, q_M]$ be an ordered set of integers such that $q_1 < q_2 < \ldots < q_M$. A flag, $\boldsymbol{F}$, in $\mathbb{R}^n$ of type $\boldsymbol{q}$ is a nested sequence of subspaces

$$\boldsymbol{F} = \{0\} \subset [S_1] \subset [S_2] \subset \cdots \subset [S_M] \subset \mathbb{R}^n \tag{3}$$

such that $\dim([S_i]) = q_i$. The flag manifold, $\mathrm{FL}(n;\boldsymbol{q})$, is then the manifold whose points correspond to all flags of type $\boldsymbol{q}$. Let $D = \{[X_i]\}_{i=1}^N$ be a finite collection of subspaces of $\mathbb{R}^n$. Compute 1-dimensional subspaces $[u^{(1)}], \ldots, [u^{(r)}]$ as the sequential optimizers of

$$\begin{aligned} &\underset{[u^{(j)}]\in Gr(1,n)}{\arg\min} \quad \sum_{[X_i]\in D} d_{pF}([u^{(j)}],[X_i])^2 \\ &\text{subject to} \qquad [u^{(j)}] \perp [u^{(l)}] \qquad \text{for } l<j \end{aligned} \tag{4}$$

where $r$ is the dimension of $\mathrm{span}\{[X_1],[X_2],\cdots,[X_N]\}$. The flag mean of $D$ was then defined by Draper *et al.*[19] to be the increasing sequential span of these 1-dimensional subspaces,

$$\boldsymbol{\mu_{pF}}(D) := \left[u^{(1)}\right] \subset \left[u^{(1)}|u^{(2)}\right] \subset \ldots \subset \left[u^{(1)}|\ldots|u^{(r)}\right]. \tag{5}$$

As detailed in Marrinan *et al.*,[20] this flag mean averages subspaces in a way that behaves like a median despite being the result of an $\ell_2$ minimization problem. Thus the flag mean pushes information shared between subspaces the the front of the flag. It is this property of the flag mean that enables us to improve chemical plume detections.

## 3. FLAG-BASED CHEMICAL PLUME DETECTION

Given a hyperspectral movie with pixelwise correspondence between frames, the flag-based detection algorithm preprocesses the data by creating a subspace from the span of a hyperspectral pixel at a time known to be free from gas, and the associated pixel from the frame under test, that is $[X_{i,j,t}] = \mathrm{span}\{x_{i,j,0}, x_{i,j,t}\}$. If a pixel contains the target spectrum in one frame and not the other the subspace will be 2-dimensional. Thus $[X_{i,j,t}] \in \mathrm{Gr}(2,b)$ or $\mathrm{Gr}(1,b)$, where $b$ is the number of spectral bands in the movie. Three horizontally adjacent subspaces (pixels) are then averaged using the flag mean, pushing mutual information to the front and creating a new representative

$$[\overline{X}_{i,j,t}] = \text{the 3-dimensional element of } \boldsymbol{\mu_{pF}}\left([X_{i,j-1,t}],[X_{i,j,t}],[X_{i,j+1,t}]\right)$$

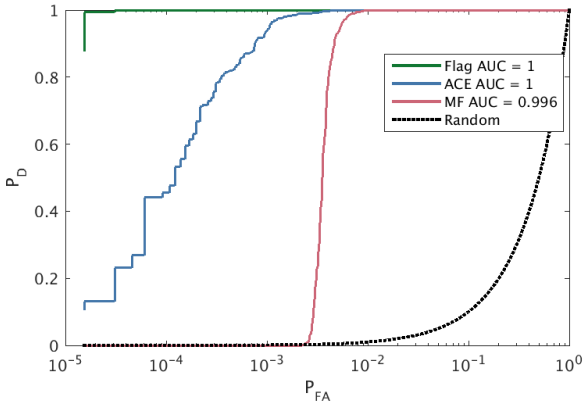
for the pixel. If all of the subspaces being averaged are 2-dimensional, the largest element of the flag mean could potentially be a 6-dimensional subspace. If the three subspaces being averaged came from pixels with homogeneous background, the 1-dimensional element of the flag will represent the background spectrum, thus in practice this dimension is discarded. However, if two of the three subspaces being averaged contain the target spectrum, the median-like property of the flag mean will push this into the second or third element of the flag mean depending on the magnitude of change in ambient conditions. The detection statistic for the flag based detection technique is then

$$D_{\text{flag}}(x_{i,j,t}) = 1 - \frac{2d([\overline{X}_{i,j,t}],[s])}{\pi}.$$

## 4. DATA SET DESCRIPTION

The long-wave infrared data for which we compute detections is a 4-dimensional array (hyperspectral movie) from the Fabry - Pérot Interferometer Sensor Data Set of size 256 rows $\times$256 columns $\times$20 bands $\times$561 frames created by the Naval Research lab. The spectrometer used to collect this data is an imaging spectroradiometer that operates efficiently in the 811 micron range. An explosive burst was used to launch and disperse the simulant Triethyl Phosphate (TEP) near frame 111 of the movie. The hyperspectral images from this movie will be used to demonstrate the effectiveness of the detection algorithms on plume-free images into which target signatures have been synthetically added, and on images after the release of the simulant to use in demonstrate practical scenarios.

(a) Randomly generated synthetic binary detection mask.




(b) ROC curves for TEP added to the hyperspectral image with $\mathrm{SINR}_{\mathrm{dB}} = 5$.

Figure 2: Detection accuracy for TEP inserted into a plume-free hyperspectral image.

## 5. QUANTITATIVE RESULTS ON SYNTHETIC DATA

The statistic computed by the flag-based detection algorithm reduces to a monotonic function of the principal angle separating the target signature from a test subspace, however, the process used to generate that representative subspace encodes a stronger signal if a neighbor of that pixels also contains the target. Hence this algorithm improves detections for signals with very small signal-to-interference-plus-noise ratios (SINR). This property will first be demonstrated on hyperspectral images where target signatures have been synthetically added to a plume-free scene at different SINRs. This artificial data will provide ground truth by which to benchmark the three algorithms quantitatively. The SINR is computed as the ratio of the signal power to the total power of the background plus white noise. Specifically, artificial plume pixels were created as $x_{i,j} = S\alpha + B_{i,j} + \nu$ where $S$ is the signature for TEP, $\alpha$ is a constant that is scaled to achieve the desired SINR, $B_{i,j}$ is the background spectrum at pixel $(i, j)$, and $\nu$ is white noise with variance $\sigma^2 = 0.2$ and covariance $\Gamma$ as described below. Thus

$$\mathrm{SINR} = \frac{\frac{1}{b}\sum_{k=1}^{b}(S(k)\alpha)^2}{\frac{1}{b}\sum_{k=1}^{b}(B_{i,j}(k))^2 + \sigma^2}$$

is the computed ratio and is typically reported in decibels as $\mathrm{SINR}_{\mathrm{dB}} = 10\log_{10}(\mathrm{SINR})$.

It is well known that adjacent bands in a hyperspectral image are highly correlated, and while each wavelength in the wavelength set of the Fabry - Pérot data have been has been selected to maximize detection sensitivity of TEP, this observation is still true. Thus, we follow the lead of Sakla *et al.*[21] and utilize a first-order Markov-based model, defined as $\nu \sim N_b[0, \Gamma]$ to generate the additive white noise. The covariance matrix $\Gamma$ is defined as $\Gamma = \sigma^2\boldsymbol{R}$ where $\boldsymbol{R}$ is the Toeplitz correlation matrix defined according to the first-order Markov model,[22]

$$\boldsymbol{R} = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^3 & \rho^{b-1} \\ \rho & 1 & \rho & \rho^2 & \cdots & \rho^{b-2} \\ \rho^2 & \rho & 1 & \rho & \cdots & \rho^{b-3} \\ \vdots & & & & \ddots & \vdots \\ \rho^{b-2} & \cdots & \rho^2 & \rho & 1 & \rho \\ \rho^{b-1} & \cdots & \rho^3 & \rho^2 & \rho & 1 \end{bmatrix} \tag{6}$$

and $\sigma^2$ is a fixed variance. In the experiments the variance was fixed at $\sigma^2 = 0.2$. The method for estimating $\rho$ follows that of Sakla *et al.*.[21]

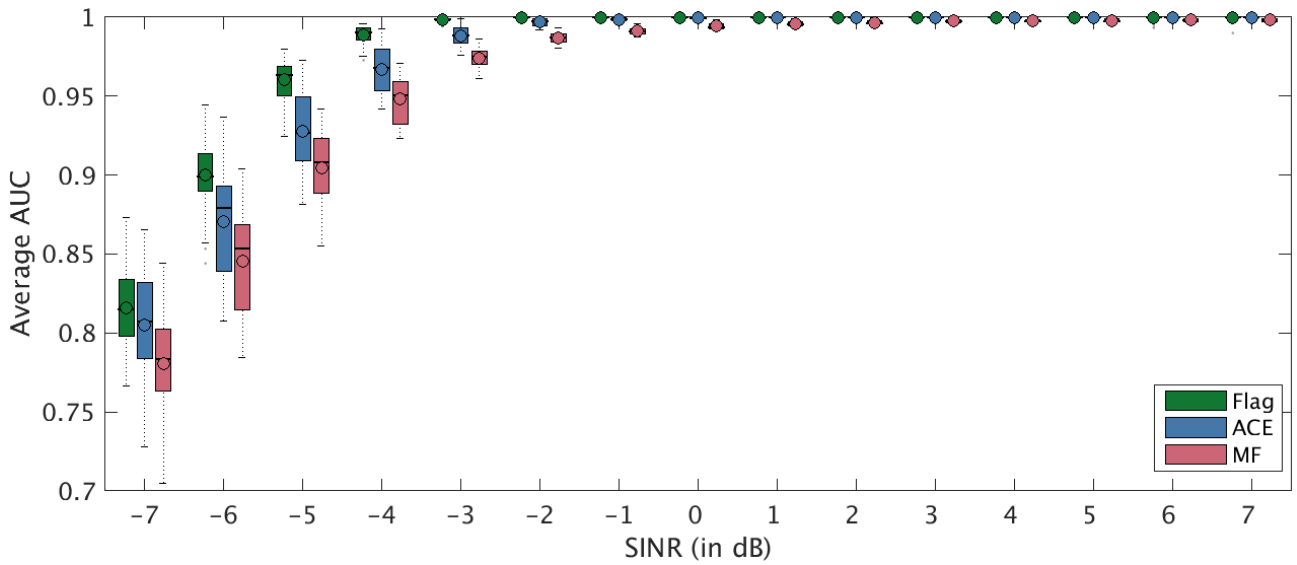


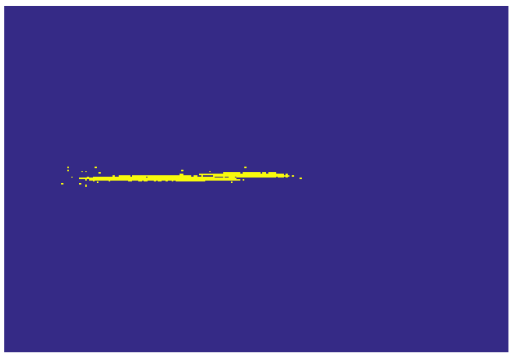
Figure 3: Box plots representing the AUC scores for detections of synthetically added TEP. Rectangular plumes were randomly generated 100 times for each reported SINR level.

Using the above method, and scaling $\alpha$ to generate the appropriate $\text{SINR}_{\text{dB}}$, TEP was added to a randomly generated rectangle accounting for 1% of the pixels in a plume-free hyperspectral image. An example of one such binary detection mask is shown in Figure 2a. White noise with the same variance and covariance was added to all pixels of the image not included in the plume. To compute detections as fairly as possible, the plume mask was used by the MF and ACE algorithms so that calculated background statistics would be as accurate as possible. The flag-based algorithm used a plume-free version of the same image with different additive white noise to create the initial 2-dimensional subspaces to be averaged. The experiment was repeated 100 times at each $\text{SINR}_{\text{dB}}$ level to generate the box plot of AUC scores in Figure 3. As can be seen, the biggest advantage of the flag-based detection algorithm over the ACE and MF methods comes at very low $\text{SINR}_{\text{dB}}$ levels. However, even when the $\text{SINR}_{\text{dB}}$ level is positive and all methods do well, the flag-based detection has a probability of detection for a low probability of false alarm as can be seen in Figure 2b

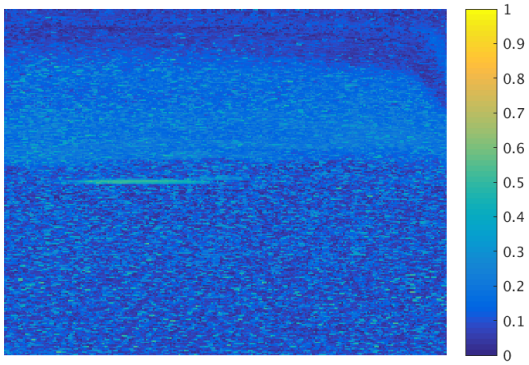
## 6. QUALITATIVE RESULTS

Other researchers have used temporal information to segment hyperspectral videos as a way to detect and track chemical plumes.[23,24] These methods do not predict whether or not a chemical agent is present, but rather cluster related hyperspectral pixels so that contiguous regions can all be identified as containing a chemical or not with another method. In an attempt to provide some measure of the success of our detections on real data, we have employed a less sophisticated supervised clustering technique to find the pixels associated with the gas plume in frame 150 of the Fabry - Pérot data, i.e. after the simulant was released. This clustering will then be used as an approximate plume mask for computing ROC curves on our various detections in that frame.

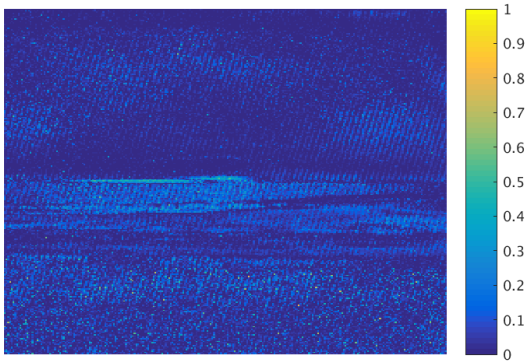
To generate the approximate plume mask, first, the temporal singular value decomposition of each pixel in the hyperspectral movie was compute. That is to say a basis was found for the space spanned by all of the spectra of a single pixel through time. The background of each pixel was estimated as the span of the first three dimensions of its basis. Each pixel was then projected into the orthogonal complement of this basis to remove most of the background information. The resulting background-removed hyperspectral pixels were clustering using $k$-means. The cluster membership along with an visualization of the spectral mean of the image was manually inspected to determine which clusters contained the plume, and those pixels were used to create the approximate plume mask. This process was heavily supervised, and while replicable, was not automated. An example of the plume mask generated for frame 150 can be seen in Figure 4a.



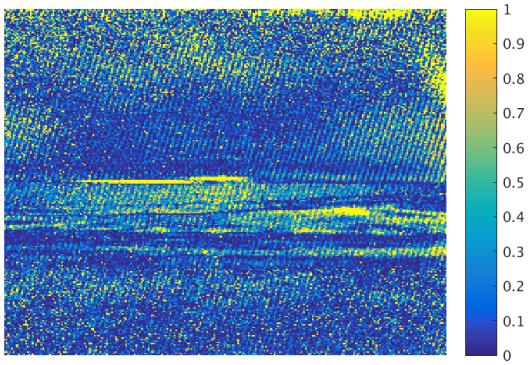
(a) Approximate plume mask computed by clustering background-removed hyperspectral pixels.




(b) Detection map for flag-based algorithm.




(c) Detection map for ACE algorithm.




(d) Detection map for matched filter algorithm

Figure 4: Detection scores associated with frame 150 of the Fabry - Pérot data. Higher scores indicate a greater likelihood of TEP being present in a given pixel.

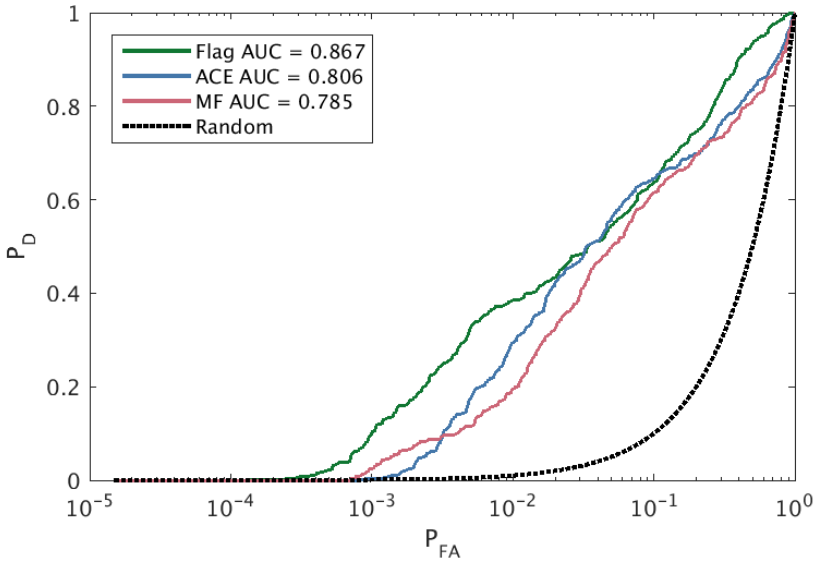


Figure 5: ROC curves generated from detections on frame 150 of the data using approximate plume as truth.

Detections were performed on images using the three algorithms. For the flag-based algorithm, the first scene of the hyperspectral movie which is free of simulant, was used as the plume-free frame to build the initial 2-dimensional subspaces for each pixel. For the MF and ACE algorithms, the approximate plume mask was used to achieve a better estimate of the background statistics. Figure 4 shows the detections maps for each algorithm on frame 150 of the data, and Figure 5 shows the associated ROC curves computed from the approximate plume mask.

In Figure 4d, the scores from matched filter algorithm, we see the highest detection statistics. However this does not necessarily translate to the best detection performance. Many of the high scoring detections are false alarms in the upper right sky pixels. In Figure 4c we see the detections of the ace algorithm. These are mostly low scores, but the plume is distinct along the horizon. Contrast is more important than high scores achieving a good level of detection, which is generally good for the ACE detections. There are a handful of speckle noise errors in the foreground of the image that generate a poor probability of detection for small probabilities of false alarm however. This agrees with the information in the ROC curves that suggests that the ACE and MF algorithms do not perform quite as well in the left-hand side of the ROC curves in Figure 5. Overall the AUC for the flag-based algorithm is higher than the two standard algorithms, and it performs better on the left side of the ROC curve. This can be identified in the detection images by the slightly better contrast between the plume and the speckle noise detections in the background.

From the AUC scores computed to be between 0.78 and 0.86, the results on synthetic data in Figure 3 would suggest that this image has a $\text{SINR}_{\text{dB}}$ around $-7$ however this cannot be computed exactly. The somewhat contiguous, higher detection statistics in the sky in Figure 4b may be attributable to the change in temperature between the first frame of the movie and the 150th frame as there is approximately a 6 second lag between frames during the capturing process and the video was capture in the morning.

## 7. CONCLUSION

In this manuscript we presented a novel flag manifold based method for detecting chemical plumes in long-wave infrared hyperspectral movies. The technique leverages knowledge of the radiance of the background scene, taken from a frame of a hyperspectral movie at a time known to be free of chemical agents, to improve the detection of chemical signatures in other frames of the movie.

The technique used to create the flags pushes information about the background clutter, ambient conditions, and potential chemical agents into the leading elements of the flags. The result of exploiting this temporal information by way of the flag structure is a novel algorithm for detecting gas plumes that appears to be sensitive to the presence of weak plumes.

Quantitative results on synthetic data show that the flag-based algorithm consistently performs better on data when the $\text{SINR}_{\text{dB}}$ is low, and beats the ACE and MF algorithms in probability of detection for low probabilities of false alarm even when the $\text{SINR}_{\text{dB}}$ is high. Qualitative experiments suggest that these results hold true on real data, when images of the scene are available at a time known to be free of the target signature.

## 8. ACKNOWLEDGMENTS

This material is based upon work partially supported by the National Science Foundation grants DMS-1228308 and DMS-1322508. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## REFERENCES

[1] Manolakis, D., Golowich, S., and DiPietro, R., "Long-wave infrared hyperspectral remote sensing of chemical clouds: A focus on signal processing approaches," *Signal Processing Magazine, IEEE* **31**(4), 120–141 (2014).

[2] Somsen, G. W. and Visser, T., "Liquid chromatography/infrared spectroscopy," *Encyclopedia of Analytical Chemistry* (2000).

[3] O'Donnell, E. M., Messinger, D. W., Salvaggio, C., and Schott, J. R., "Identification and detection of gaseous effluents from hyperspectral imagery using invariant algorithms," in [*Defense and Security*], 573–582, International Society for Optics and Photonics (2004).

[4] Grutter, M., Basaldud, R., Flores, E., and Harig, R., "Optical remote sensing for characterizing the spatial distribution of stack emissions," in [*Advanced Environmental Monitoring*], 107–118, Springer (2008).

[5] Tremblay, P., Savary, S., Rolland, M., Villemaire, A., Chamberland, M., Farley, V., Brault, L., Giroux, J., Allard, J.-L., Dupuis, É., et al., "Standoff gas identification and quantification from turbulent stack plumes with an imaging fourier-transform spectrometer," in [*SPIE Defense, Security, and Sensing*], 76730H–76730H, International Society for Optics and Photonics (2010).

[6] Scharf, L. L., [*Statistical signal processing*], vol. 98, Addison-Wesley Reading, MA (1991).

[7] Manolakis, D., Siracusa, C., and Shaw, G., "Hyperspectral subpixel target detection using the linear mixing model," *Geoscience and Remote Sensing, IEEE Transactions on* **39**(7), 1392–1409 (2001).

[8] Kraut, S., Scharf, L. L., and McWhorter, L. T., "Adaptive subspace detectors," *Signal Processing, IEEE Transactions on* **49**(1), 1–16 (2001).

[9] Kroutil, R., Lewis, P., Thomas, M., Curry, T., Miller, D., Combs, R., and Cummings, A., "Emergency response chemical detection using passive infrared spectroscopy," *SPIE Newsroom* (2006).

[10] Flanigan, D. F., "Hazardous cloud imaging: a new way of using passive infrared," *Applied optics* **36**(27), 7027–7036 (1997).

[11] Draper, B., Kirby, M., Marks, J., Marrinan, T., and Peterson, C., "A flag representation for finite collections of subspaces of mixed dimensions," *Linear Algebra and its Applications* **451**, 15–32 (2014).

[12] Manolakis, D., Marden, D., and Shaw, G. A., "Hyperspectral image processing for automatic target detection applications," *Lincoln Laboratory Journal* **14**(1), 79–116 (2003).

[13] Adams, J. B., Smith, M. O., and Gillespie, A. R., "Imaging spectroscopy: Interpretation based on spectral mixture analysis," *Remote geochemical analysis: Elemental and mineralogical composition* **7**, 145–166 (1993).

[14] Manolakis, D. and Shaw, G., "Detection algorithms for hyperspectral imaging applications," *Signal Processing Magazine, IEEE* **19**(1), 29–43 (2002).

[15] Manolakis, D., Pieper, M., Truslow, E., Cooley, T., Brueggeman, M., and Lipson, S., "The remarkable success of adaptive cosine estimator in hyperspectral target detection," in [*SPIE Defense, Security, and Sensing*], 874302–874302, International Society for Optics and Photonics (2013).

[16] Zare, A. and Ho, K., "Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing," *Signal Processing Magazine, IEEE* **31**(1), 95–104 (2014).

[17] Edelman, A., Arias, T., and Smith, S., "The geometry of algorithms with orthogonality constraints," *SIAM J. Matrix Analysis and Applications* **20**(2), 303–353 (1998).

[18] Björck, A. and Golub, G., "Numerical methods for computing angles between linear subspaces," *Mathematics of computation* **27**(123), 579–594 (1973).

[19] Draper, B., Kirby, M., Marks, J., Marrinan, T., and Peterson, C., "A flag representation for finite collections of subspaces of mixed dimensions," *Linear Algebra and its Applications* **451**, 15–32 (2014).

[20] Marrinan, T., Draper, B., Beveridge, J. R., Kirby, M., and Peterson, C., "Finding the subspace mean or median to fit your need," in [*CVPR*], 1082–1089, IEEE (2014).

[21] Sakla, A. A., Sakla, W. A., and Alam, M. S., "Hyperspectral target detection via discrete wavelet-based spectral fringe-adjusted joint transform correlation," *Applied optics* **50**(28), 5545–5554 (2011).

[22] Jain, A. K., [*Fundamentals of digital image processing*], Prentice-Hall, Inc. (1989).

[23] Gerhart, T., Sunu, J., Lieu, L., Merkurjev, E., Chang, J.-M., Gilles, J., and Bertozzi, A. L., "Detection and tracking of gas plumes in lwir hyperspectral video sequence data," in [*SPIE Defense, Security, and Sensing*], 87430J–87430J, International Society for Optics and Photonics (2013).

[24] Sunu, J., Chang, J.-M., and Bertozzi, A. L., "Simultaneous spectral analysis of multiple video sequence data for lwir gas plumes," in [*SPIE Conference on Defense, Security, and Sensing*], (2014).