

Chemical Detection Using Hyperspectral Imaging: Vapor Sensing in the Atmosphere

Russell E. Warren

EO-Stat Inc. of North Carolina
10010 Vail Drive
Chapel Hill, NC 27517 USA

Stanley J. Osher

Department of Mathematics
University of California
Los Angeles, CA 90095 USA

ABSTRACT

We have developed an optimization theoretic approach for using hyperspectral (HS) data with the vapor materials' absorption spectra to identify and locate chemical materials in the atmosphere against arbitrary clutter-backgrounds. The method makes no prior modeling assumptions about the existence of pure pixels or the statistics of the clutter and sensor noise. It also requires neither prior background measurements taken without the vapor absorption nor the use of libraries of background spectra. There are two main components to the method: (1) a hyperspectral unmixing algorithm based on the Alternating Direction Method of Multipliers that is used over local spatial subsets of the imaging to resolve the HS data into a set of linearly independent spectral and spatial components; and (2) the fitting of those unmixing spectra to a set of candidate absorption spectral templates for chemical detection, identification, and location. The algorithm is illustrated using HS data collected by a Telops Hyper-Cam passive sensor during the release of vapor materials at Dugway Proving Ground, UT.

1 Introduction

Hyperspectral (HS) imaging is a well-established technology having many commercial and military applications. One of those uses the HS spectral information to detect and locate materials by their known spectral signatures. In general, this is a highly nontrivial task due to the many radiance sources from the clutter background and atmosphere that can easily mask and distort the often weak radiance from the chemical materials of interest. Traditional approaches for separating the radiance components in the data involve either (1) unmixing using the "pure-pixel" or "endmember" assumption that the target materials can be found in a few individual pixels without interference from other radiance sources; or (2) the use of likelihood ratio testing in a statistical approach.

The pure pixel or endmember algorithms are more successful on data for which the pure pixel assumption is approximately valid. If such pixels exist and can be identified, least-squares methods can be used to locate the image spatial regions associated with each endmember. Standard simplex-based algorithms such as N-findr [1] and VCA [2] are useful if endmember pixels exist under high SNR conditions. We note, however, that these algorithms may not work well because of the nonexistence of such pixels due to spatial resolution- or sensor-noise limitations. This was already

recognized by the likely originator of the term "endmember," Schowengerdt [3]: "... endmembers only exist as a conceptual convenience and as idealizations in real images." Also, the concept of spectrally pure pixels may not have meaning in the LWIR spectral region where our data reside.

At the other extreme are unmixing approaches, notably those based on compressive sensing methods, which attempt to fit the data to a sparse set of spectra taken from large libraries. These methods require additional prior information about the data that may not exist in general. Other methods for sparse basis fitting are well known in HS processing. We mention two sources: Dadon et al. [4] and Adler-Golden and Conforti [5] that use sparse fitting, but in the context of endmember bases. In contrast, we will show in Section 3 that the ADMM unmixing algorithm constructs a sparse basis of spectra directly from the data at each local spatial region.

Targets that are subpixel or mixed pixel are well known to be quite challenging compared to the traditional endmember processing. These targets are either small in relation to the pixel size or partly concealed by vegetation or other coverings. For this class of targets, statistical methods are used such as the Adaptive Matched Detector (AMD) [6], and the Adaptive Subspace Detector (ASD) that apply Generalized Likelihood Ratio Testing (GLRT) [7] using background sample covariance matrices to whiten the data followed by subspace projection onto the target spectral subspace.

Both classes of methods make optimistic assumptions about the data--either the existence and detectability of pure pixels--or that the data statistics are well approximated by multivariate normal probability densities. The latter method assumes that the clutter correlation structure can be estimated from the data, assuming spatial homogeneity and the ability to mask the target spectral presence. These are often difficult assumptions to justify in practice.

We have developed an optimization theoretic method for identifying and locating materials in unknown clutter backgrounds using their chemical spectra that makes no prior modeling assumptions about the presence of pure (endmember) pixels or the statistics of the clutter and sensor noise. Also, in contrast to any other method we know of, separate estimates of the clutter-background are not required. The generality of the approach makes it potentially applicable to a wide variety of military and civilian scenarios that are significantly widened by the elimination of the requirement to obtain a clean background data scan. An earlier paper [8] described the optimization approach in the context of chemical sensing on spectrally unknown surfaces. Here we focus on finding vapor materials in the atmosphere within arbitrary thermal backgrounds.

There are two main components to our method: (1) a hyperspectral unmixing algorithm [9] based on the Alternating Direction Method of Multipliers (ADMM) [10] that is used over local subsets of the imaging to resolve the HS data into a set of linearly independent spectral and spatial components; and (2) the fitting of those unmixing spectra to a set of candidate template spectra for chemical detection, identification, and, when combined with the spatial information, location. The fitted spectra can be used for material identification by either a very simple threshold test on the fit quality for each template, or, better, as input to a state-of-the-art classifier such as the support vector machine. We have followed the first choice here because of the limited data that were

available; the earlier paper on chemical sensing on surfaces used an SVM trained of HS test data with chemical-agent simulants.

In Section 2 we derive the ADMM unmixing algorithm for vapor sensing in the atmosphere. It is a simplified version of the algorithm described in [8,9]. Although the unmixing is an essential component of our chemical sensing approach, the optical interaction between the chemical reflectance spectral structure and the other radiance components requires an additional material-dependent fitting step to convert the unmixing spectral vectors into estimates of the spectral reflectance of a given material. Section 3 describes the fitting method used for this step. In Section 4 we illustrate the algorithm on HS data collected by a Telops Hyper-Cam sensor during the releases of the refrigerant R-134a and acetic acid at Dugway Proving Ground (DPG) in UT. Section 5 summarizes the algorithm and its application to vapor sensing.

2 ADMM Hyperspectral Unmixing Algorithm for Vapor Sensing

In this section we discuss the method we have developed for the unmixing task of resolving hyperspectral data cubes into independent components as functions of wavelength band (spectrum) and spatial structure (the analogue of abundances). The algorithm derived here is quite straightforward to code using the suggested parameters in Section 4.

We formulate the HS unmixing problem in terms of the spectral and spatial arrays ρ and C as the biconvex multiplicative model $G = \rho C + n$, where G represents the HS data cube as $G \in R^{M \times N}$ with M the number of spectral bands (wavelengths) and $N = N_1 N_2$ is the total number of pixels after stacking the N_1 rows and N_2 columns into a single vector. We model $\rho \in R_+^{M \times L}$ and $C \in R_+^{L \times N}$ with L the number of assumed independent components in the cube. n is an additive noise term taken to be zero-mean, independent, and identically distributed with bounded variance. We then have the constrained problem

$$\min_{\rho, C} \frac{1}{2} \|G - \rho C\|_F^2 \quad \text{such that } \rho \geq 0, \|\rho_l\| = 1, 1 \leq l \leq L, C \geq 0. \quad (1)$$

Because we have only a biconvex optimization problem (convex in either ρ or C given the other, but not jointly convex), we cannot expect to find a global minimizing solution but only approximately optimal solutions that are numerically well-behaved under different initializations of ρ . Fortunately, the ADMM formalism can accommodate this biconvex problem, and produce good unmixing results with the addition of appropriate constraints. ADMM implementation for the unmixing problem leads to an alternating algorithm that, for a given estimate of ρ , updates the estimate of C and vice versa until convergence.

Constraints are needed to produce physically meaningful results. They are chosen to be positivity on the spatial and spectral estimates, and a unit-vector constraint on the spectrum of each material. The latter replaces the spatially convex constraints

used for abundances in the usual endmember approaches, and seem more natural physically. The constraints are enforced on the parameter arrays through augmented Lagrangians with Bregman splitting. For the splitting we have used the SOC (Splitting Orthogonal Constraint) method of Lai and Osher [11] that introduces auxiliary variables to split the positivity constraints from the estimation of concentration and the spectra. Improved results are often obtained by introducing an l_1 regularization of the spectral estimates through total variation (TV) also implemented through Bregman splitting and soft thresholding following Goldstein and Osher [12]. The ADMM algorithm in [8,9] includes this TV regularization. For simplicity and computational efficiency, we do not use TV regularization on the spectral estimates in the algorithm derived here. The resulting algorithm is considerably simpler, easier to implement, and faster than the one described in [8,9].

In addition to the need for constraints, we have observed that much better results are obtained by applying the unmixing and spectral fitting to local subsets of the total data cube. This is particularly true when the materials of interest are present in relatively small regions of the total imagery. The subsets are conveniently chosen to be squares typically having about 30 pixels on a side. This choice allows the processing to adapt to the local structure of the image while supplying enough spatial information to generate good unmixing results. Besides providing better spectral estimates, the use of local processing can be computationally advantageous over a single global fit since the processing steps are identical for each local region, and can therefore be implemented by parallel processing methods.

For the HS unmixing problem we introduce two new sets of parameters: one set $r = \rho$ for the spectra, and a set $e = C$ for the spatial concentration array. These parameters are iteratively updated independently on each data block of the image cube, initialized either randomly (1st data block) or using the estimates at convergence from the prior block. Letting ρ^k, r^k, C^k, e^k denote the parameter estimates at iteration k , we have the problem

$$\min_{\rho, r, C, e} J(C, \rho) \text{ such that } \rho = r, r \geq 0, \|r\| = 1, C = e, e \geq 0, \quad (2)$$

with $J(C, \rho) = \frac{1}{2} \|G - \rho C\|_F^2$, where the subscript F denotes the Frobenius (Hilbert-Schmidt) matrix norm.

The augmented Lagrangian for this problem is then

$$\begin{aligned} L(C, e, \rho, r, q) = & J(C, \rho) + \text{tr} \langle p, e - C \rangle + \frac{\lambda_C}{2} \|e - C\|_F^2 + I_+(e) + I_+(r) + \\ & \text{tr} \langle q, r - \rho \rangle + \frac{\lambda_\rho}{2} \|r - \rho\|_F^2, \end{aligned} \quad (3)$$

with dual Lagrange multiplier parameters p and q that enforce the equality constraints. The Lagrange multipliers are updated within the iterative ADMM framework by the method of multipliers developed independently by Hestenes [13] and Powell [14] for

equality constraints, and later generalized to inequality constraints by Rockafellar [15]. The third and seventh terms on the right represent the quadratic penalty terms that augment the classical Lagrangian. Their addition promotes the constraint enforcements in (2), adds numerical stability to the solution, and provides a systematic method for locally optimizing both the multipliers and $\{\rho, C\}$ in parallel. By virtue of the additional parameters e and r , we have an unconstrained optimization over C and ρ with the remaining constraints easily enforced on e and r analytically. The parameters λ_C and λ_ρ , are termed penalty parameters. The fourth and fifth terms represent indicator functions defined as $I_+(e) = 0$, for $e \geq 0$, $= \infty$, for $e < 0$.

From the structure of (3) we see that the total problem at a given iteration k breaks into two saddle-point subproblems: finding (C^k, e^k, p^k) given ρ^{k-1} , et al., and finding (ρ^k, r^k, q^k) given C^k , et al. For the first subproblem we have

$$(C^k, e^k, p^k) = \arg \max_{\rho} \min_{C, e \geq 0} L(C, e, \rho, \rho^{k-1}, r^{k-1}, q^{k-1}). \quad (4)$$

From

$$\nabla_C L = -\rho^T (G - \rho C) - p + \lambda_C (C - e) \Big|_{\rho=\rho^{k-1}, p=p^{k-1}, e=e^{k-1}} = 0, \quad (5)$$

we find

$$C^k = (\rho^{(k-1)T} \rho^{k-1} + \lambda_C I)^{-1} (\rho^{(k-1)T} G + p^{k-1} + \lambda_C e^{k-1}). \quad (6)$$

Similarly, $\nabla_e L = p^{k-1} - \lambda_C (C^k - e) = 0$ and the positivity constraint give the projection onto the positive halfspace

$$e^k = [C^k - p^{k-1} / \lambda_C]^+ \equiv \max(C^k - p^{k-1} / \lambda_C, 0). \quad (7)$$

From the general theory of augmented Lagrangians [16] we get

$$p^k = p^{k-1} - \lambda_C (C^k - e^k). \quad (8)$$

The second subproblem assumes the form

$$(\rho^k, r^k, q^k) = \arg \max_q \min_{\substack{\rho, r \geq 0 \\ \|r\|=1, 1 \leq i \leq L}} L(C^k, e^k, \rho^k, \rho, r, q). \quad (9)$$

Differentiating L with respect to ρ gives

$$\nabla_{\rho} L = -(\mathbf{G} - \rho \mathbf{C}) \mathbf{C}^T - \mathbf{q} + \lambda_{\rho} (\rho - r) \Big|_{\mathbf{C}=\mathbf{C}^k, \mathbf{q}=\mathbf{q}^{k-1}, r=r^{k-1}, \rho^{k-1}} = 0, \quad (10)$$

and, after collecting terms, the equation for ρ^k

$$\rho^k = (\mathbf{G} \mathbf{C}^{kT} + \mathbf{q}^{k-1} + \lambda_{\rho} r^{k-1}) (\mathbf{C}^k \mathbf{C}^{kT} + \lambda_{\rho} \mathbf{I}_L)^{-1}. \quad (11)$$

From $\nabla_r L = \mathbf{q}^{k-1} - \lambda_{\rho} (\rho^k - r) = 0$ and the positive orthant sphere constraint we get

$$u \equiv \rho^k - \mathbf{q}^{k-1} / \lambda_{\rho}, \quad r^k = [u]^+ / \|[u]^+\|_F \quad (12)$$

The Lagrange multiplier updates for q are then

$$\mathbf{q}^k = \mathbf{q}^{k-1} - \lambda_{\rho} (\rho^k - r^k). \quad (13)$$

Both subproblem recursions for \mathbf{C} by (6)-(8) and ρ by (11)-(13) are iterated to convergence (with r , e , s , ρ , and q initialized at either 0 or the results from the previous sub-block) within an outer loop that iterates between the subproblems. The spectral estimates ρ^0 are initialized by either positive, unit-norm random vectors, computed from the VCA algorithm¹ or using the estimates from the previous sub-block. A summary of the overall algorithm is given in Box 1.

We make the following remarks about the algorithm in Box 1. The HS data cube is denoted by the three-dimensional array $[R(j, k_1, k_2)]$ where $1 \leq j \leq M$, $1 \leq k_1 \leq N_1$, and $1 \leq k_2 \leq N_2$. For the unmixing, R is partitioned into disjoint square blocks of length N_s . \mathbf{G} for each local block is formed by stacking the spatial columns into a single row vector. This operation is denoted by $\langle \mathfrak{g} \rangle$. In the version in Box 1 the spectral array ρ is constrained to a positive unit-norm set over spectral band using the SOC method of [11] applied to the splitting variables r . The function $[x]^+ \equiv \max(x, 0)$. Finally, the subscript F used in conjunction with the vector space norms of the spectral and spatial arrays denotes the Frobenius norm. The norm of the splitting variables r on the other hand is to be interpreted as the norm over the spectral components for each material independently.

¹ An alternative to VCA initialization would be random positive unit-norm vectors. VCA estimates tend to give faster convergence.

```

Input:  $N_1, N_2, M, N_s, R, \lambda_c, \lambda_\rho, \varepsilon, \rho^{(0)}, L, \text{tol}$ 
Initialize:  $C^{(0)} = e^{(0)} = p^{(0)} = 0, r^{(0)} = q^{(0)} = 0, \bar{\rho}^{(0)} = 0$ 
 $N_{s_1} = \text{floor}(N_1/N_s), N_{s_2} = \text{floor}(N_2/N_s),$ 
for  $n_1 = 1 : N_{s_1}$ 
   $i_1 = (n_1 - 1)N_s + 1 : n_1N_s$ 
  for  $n_2 = 1 : N_{s_2}$ 
     $i_2 = (n_2 - 1)N_s + 1 : n_2N_s$ 
     $G = \langle R(1 : M, i_1, i_2) \rangle$ 
     $D\rho = 1, k = 1$ 
    while  $D\rho > \varepsilon$ 
       $\delta = 1$ 
      while  $\delta > \text{tol}$ 
         $C^{(k)} = (\rho^{(k-1)T} \rho^{(k-1)} + \lambda_c I_L)^{-1} (\rho^{(k-1)T} G + p^{(k-1)} + \lambda_c e^{(k-1)})$ 
         $e^{(k)} = [C^{(k)} - p^{(k-1)} / \lambda_c]^+$ 
         $p^{(k)} = p^{(k-1)} - \lambda_c (C^{(k)} - e^{(k)})$ 
         $\delta = \|C^{(k)} - C^{(k-1)}\|_F / L$ 
      end
       $\delta = 1$ 
      while  $\delta > \text{tol}$ 
         $\rho^{(k)} = (GC^{(k)T} + q^{(k-1)} + \lambda_\rho r^{(k-1)}) (C^{(k)} C^{(k)T} + \lambda_\rho I_L)^{-1}$ 
         $r^{(k)} \equiv [ \rho^{(k)} - q^{(k-1)} / \lambda_\rho ]^+, \quad r^{(k)} = r^{(k)} / \|r^{(k)}\|$ 
         $q^{(k)} = q^{(k-1)} - \lambda_\rho (\rho^{(k)} - r^{(k)})$ 
         $\delta = \|\rho^{(k)} - \rho^{(k-1)}\|_F$ 
      end
       $\bar{\rho}^{(k)} = \rho^{(k)}, D\rho = \|\bar{\rho}^{(k)} - \bar{\rho}^{(k-1)}\|_F$ 
       $k \rightarrow k + 1$ 
       $\rho_s(n_1, n_2) = \bar{\rho}^{(k)}$ 
    end
  end
end
end

```

Box 1. Locally partitioned ADMM unmixing algorithm computational flow.

3 Fitting Unmixing Spectra to Chemical Reflectances

In our early experience with the unmixing algorithm we had hoped that the spectral estimates would prove adequate to identify the radiance components of chemical materials without additional processing. We now know that this is not possible in general. As an example, Figure 1 plots the spectral template of the refrigerant R-134a at the Telops LWIR wavelengths. This spectrum was supplied as part of a data set [17]

collected at DPG UT in 2009 using the Telops Hyper-Cam sensor during the atmospheric release of various chemical materials.

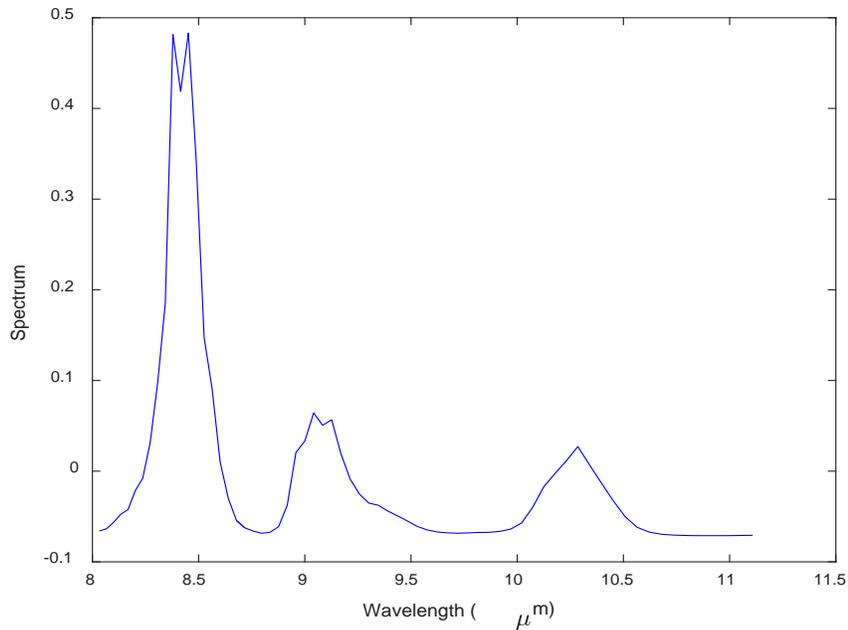


Figure 1. Spectral template of the refrigerant R-134a.

Figure 2 shows the unmixing spectra from the ADMM algorithm assuming $L = 6$ materials for each of a set of non-overlapping 32-by-32 pixel blocks of the Hyper-Cam cube. The spectra in Figure 2 were computed for the spatial block location having the highest correlation between the fitted spectra and the template (to be discussed below). Comparison of these figures indicates that some of the unmixing spectra show structure near $8.4 \mu\text{m}$ that looks something like the R-134a template in Figure 1, but certainly not similar enough for detecting the material. The unmixing spectra are evidently a mixture of the thermal clutter background and the chemical spectrum.

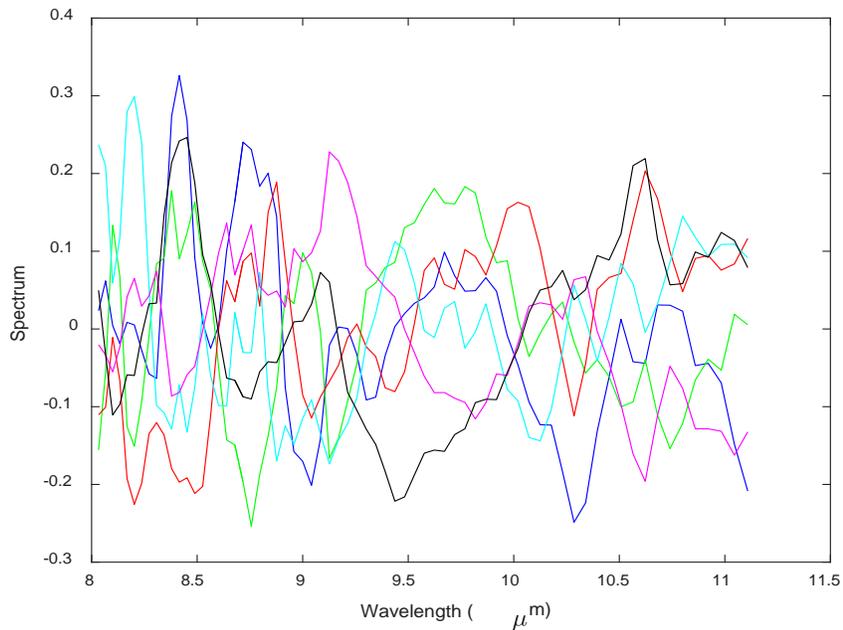


Figure 2. Spectral estimates from the ADMM algorithm on R-134a release data.

The question is then how the unmixing spectra can be used to detect the chemicals. The key is that the unmixing spectra do not represent the pure chemicals, but contain the reflectances of the chemicals as observed in the background clutter from the surface, illumination source, and possibly the atmosphere. *They are therefore best regarded as independent basis vectors that can be used to fit the templates of whatever materials may be present.* With this interpretation of the unmixing spectra, we are led to the spectral fitting algorithm below.

The basic idea is to model a given chemical reflectance spectrum as a linear combination of the spectral estimates from the unmixing algorithm (after mean subtraction and unit-vector normalization) on each data block. The assumptions we make about the spectral fitting in terms of its ability to identify and locate chemical releases are: (1) the template spectrum can be well represented as a linear combination of 6-8 unmixing spectra when the target material is present in the local subset of the data; and (2) the best linear fit to a given template will not fit that template well when the material is not present in the local block. The examples below support the validity of these assumptions.

We note that the ADMM unmixing algorithm is essential for producing a sparse set of basis vectors that spans the spectral structure of the image sub-block. Although there exists an uncountable infinity of possible basis vectors, the unmixing method used here selects a sparse set that is directly tailored to the image data. Other basis-generating methods such as principal components analysis, use empirical estimates of the spectral covariance, and the resulting eigenvectors have little relation to the actual spectral structure of the sub-block. This means that many more such vectors would be needed to produce good fits, with a decrease in the discrimination power of the method.

The fits to the linear model are computed by a standard least-squares method, and the results are compared to the input reflectance template using an inner product over wavelength-band number. For increased numerical stability, we orthonormalize the unmixing spectral estimates using the Matlab function "orth" for each data block. The algorithm pseudo-code is given in Box 2. In Box 2 ρ is the mean-subtracted M -by- L spectral array ρ from the unmixing algorithm, q is the orthogonalized array of spectral estimates (such that $q^T q = I_L$), and ρ_M is the 1-by- M row vector representing the unpolarized reflectance of a candidate material after mean-subtraction and unit-vector normalization. For multiple-candidate-material detection, ρ_M could be replaced by a K -by- M array where K is the number of materials. The regression estimate of the reflectance is denoted by $\hat{\rho}$, and S is the inner product of ρ_M and $\hat{\rho}$. Because of the unit-vector normalizations, we always have $|S| \leq 1$.

The resulting inner products are used below as a figure-of-merit to locate the simulant chemical materials in HS imagery for atmospheric vapor absorption observed against the natural thermal background clutter. In Reference [8] the regression fit estimates from the various simulant chemicals on different backgrounds were used as feature vectors to construct SVM classifiers for identifying the materials.

```

Input:  $\rho, \rho_M$ 
 $\rho = \rho - \text{mean}(\rho)$ 
 $q = \text{orth}(\rho)$ 
 $\hat{\rho} = \rho_M q q^T$ 
 $\hat{\rho} \rightarrow \hat{\rho} / \|\hat{\rho}\|$ 
 $S = \rho_M \hat{\rho}$ 
Output:  $\hat{\rho}, S, q$ 

```

Box 2. Spectral reflectance regression algorithm.

4 Application to Vapor Sensing Data

We illustrate the chemical detection method with block-processing on two vapor-release data cubes. The first [17] is the release of the refrigerant R-134a at DPG using the Telops Hyper-Cam sensor. The file name is 'dugway_released_r134a.mat.' This data set was used to motivate the need for spectral fitting in Section 3. To reduce the effect of outlier pixels, a spatial median filter of width ± 1 pixels about each pixel was applied to the input data cube. Running the block-processing combination of ADMM unmixing followed by spectral fitting for $L = 8$ materials on 32-by-32 pixel blocks gave the best template fit compared to the template itself in Figure 3. The penalty parameters used to implement the augmented Lagrangian in the unmixing were set at $\lambda_\rho = 0.1$ and $\lambda_c = 0.01$. The spectral array at sub-block (1,1) was initialized by pseudo-random unit-norm vectors. The average unmixing time per block was 0.048 s. The peak S at block (3,6) was 0.9810. The spatial structure from the total set of block fits is shown in Figure 4.

The spatial structure was computed as $\hat{C} = \rho_M q q^T R$ over HS data R within each block. There is no evident indication of the boundaries of the individual data blocks in the figure. The R-134a appears as the strong absorption feature in the plot.

Figure 5 plots the array of inner product values S . We note that aside from blocks (3,6) and (4,6), the S values are less than 0.8. The inner products could therefore be used as a crude detection method, but better results are typically achieved from a good classifier such as an SVM trained on a large set of spectral fits.

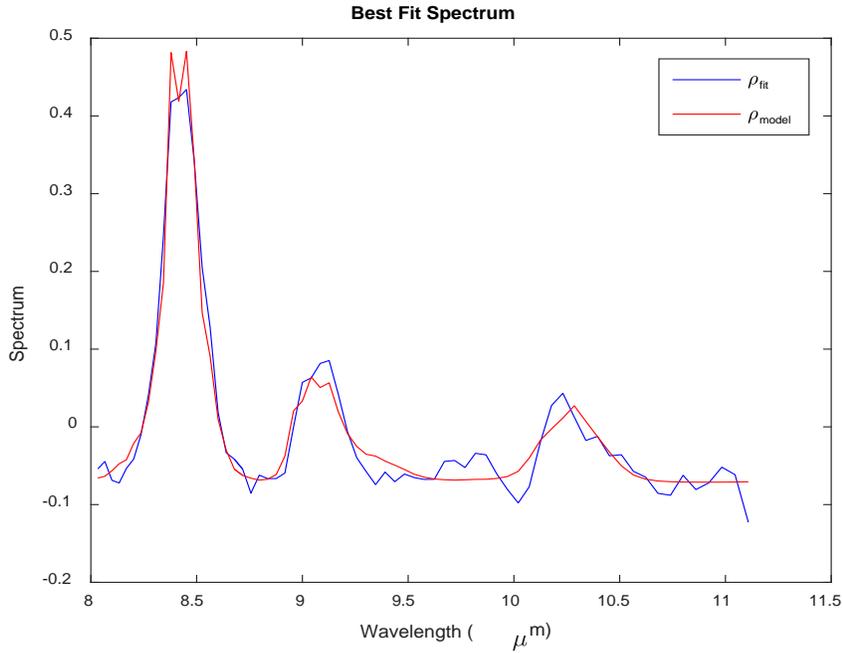


Figure 3. Template and peak-fit spectra for R134a release data.

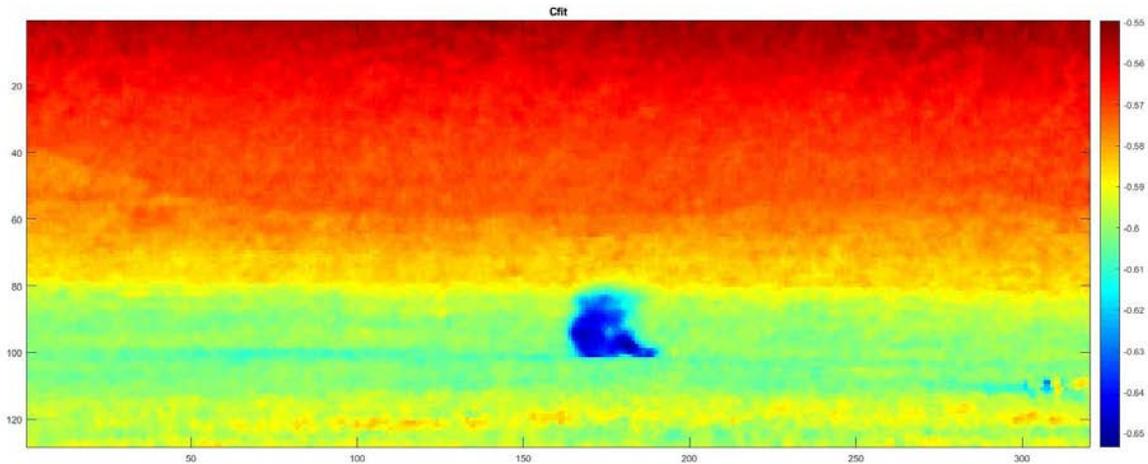


Figure 4. Spatial estimates of the R134a release data.

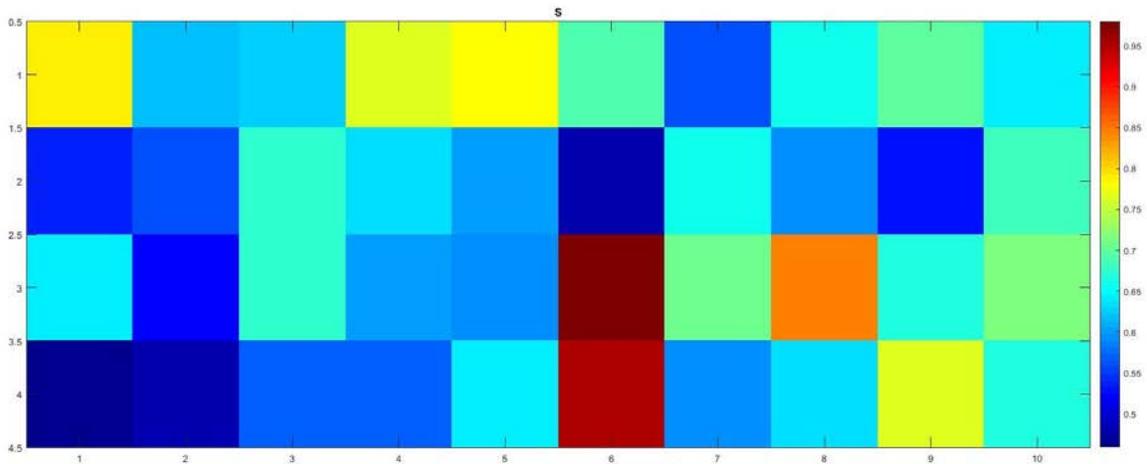


Figure 5. Inner product array S for the R134a release data.

Our second example is the open-air release of acetic acid also at DPG [18]. The file name is 'gdss_27_released_aa_blind.' Applying the unmixing for 8 materials over 30-by-30 pixel blocks using the median filter and unmixing parameters above gave the best spectral fit to the acetic acid template at block (3,18) with $S = 0.9864$. Figure 6 compares these spectra. The corresponding spatial plot is shown in Figure 7. The average unmixing time per block was 0.0619 s. We note what appear to be components close to the ground as well as in the atmosphere. The inner product array in Figure 8 shows relatively good fits for both of these apparent components.

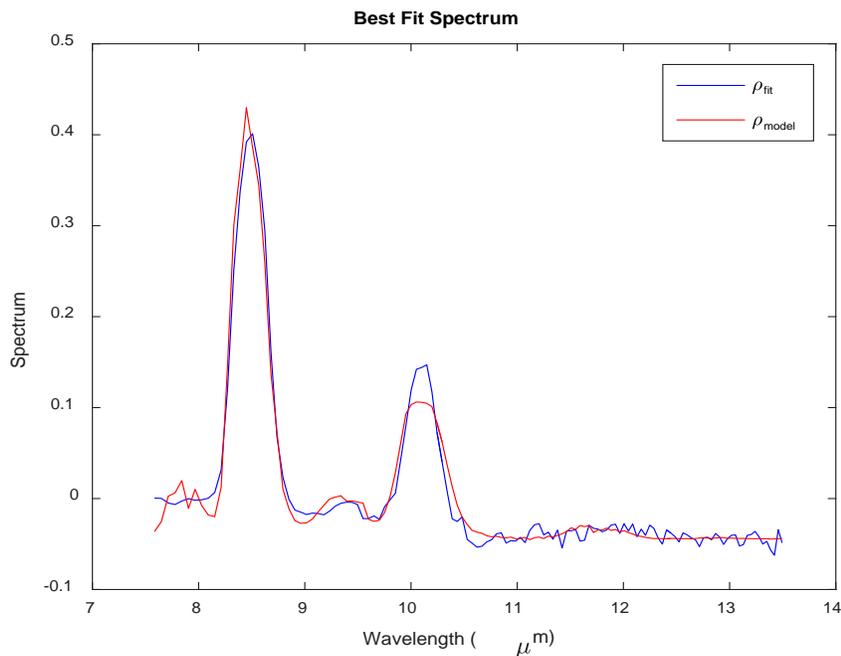


Figure 6. Template and best fit to the acetic acid release data.

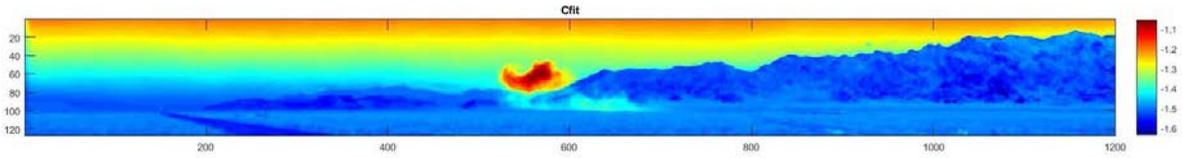


Figure 7. Spatial estimates of the acetic acid release data.

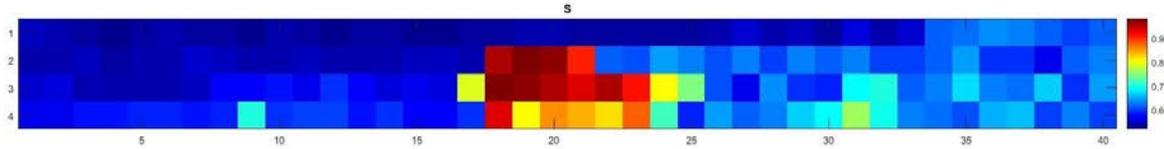


Figure 8. Inner product array S for the acetic acid release data.

5 Summary and Conclusions

We have derived and illustrated a new optimization theoretic method for using hyperspectral imaging data to identify and locate vapor materials in the atmosphere by fitting the set of spectral estimates from an unmixing algorithm to a set of candidate chemical spectra. The unmixing and spectral fitting are done locally over non-overlapping blocks of pixels. The ADMM HS unmixing algorithm, derived here, is a simplified and faster version of the method used in [9]. Neither algorithm assumes the presence of endmember pixels, and neither requires background data subtraction, or modeling assumptions such as the validity of multivariate normal statistics.

With reference to the latter class of processing techniques, the optimal (likelihood ratio) processor would attempt to estimate the spatial covariance of the scene, and spatially whiten the target-and-background combination prior to matched filtering with the chemical spectral templates. We short-circuit most of these steps by the local unmixing followed by spectral fitting and (perhaps) classification. The average unmixing time per block for the R-134a example was 0.048 s, and the average time for the acetic acid data was 0.062 s. These calculations were done with Matlab R2015b on a PC using Windows 7 with an Intel XPS8500 processor running at 3.4 GHz. In other words, we used a commercial computer, operating system, and high level software. Significant reduction in processing times could be achieved with a dedicated machine and DSP boards.

We would not claim that our HS detection algorithm will perform better than likelihood processing on data for which it is best suited, i.e., spatially homogeneous multivariate normal data with good covariance and clean background estimates. The question is, how often will such data be available? We believe that the main advantage of our optimization theoretic approach is that it can produce good results without requiring these statistical assumptions about the data.

Most of the processing time in our approach is taken up by the unmixing and fitting/classification steps (in the case of a large library of candidate materials). Both of those steps perform the same operations on all pixel blocks and library materials, and are therefore parallelizable and could be done potentially in real time. Also, the chemical templates could be provided with absolute scales, making the detection results interpretable in concentration units.

The regression algorithm that fits the candidate spectral templates to the unmixing spectra is the reverse of most hyperspectral algorithms, which fit the data to selected spectral samples from large libraries. In our approach the roles of the target function and independent fitting functions are reversed. Our basis functions are produced locally from the data via the unmixing algorithm, and the target functions are the set of candidate materials. The latter set can be quite large, but the simple regression fitting proposed here is very fast, and likely to be more efficient than the compressive sensing algorithms with their large spectral libraries, currently in vogue.

The processing requirements needed to apply the proposed chemical sensing algorithm to the vapor and other hyperspectral data applications are: (1) a good unmixing algorithm, and (2) a set of hypothesized target spectral templates. The latter could be the spectra of different materials or the same material under different atmospheric absorption conditions. The assumptions in terms of the spectral fitting are that (1) the template spectrum can be represented well as a linear combination of a small number of the locally estimated unmixing spectra when that material is present in the data, and (2) conversely, the best linear fit to a given template will not match the template well when that material is not present. We have empirically verified those assumptions on a wide variety of processing cases.

Data from training sets from a specific type of hyperspectral sensor, and collected on a representative population of target material releases (in the case of vapor sensing), can be used to train a classifier such as the SVM. We have successfully implemented and demonstrated on field data such a processor for chemical contamination sensing on spectrally unknown real-world surfaces [8] actively irradiated by a heated blackbody. In that case, proximal detection in real time from a moving vehicle could be achieved by use of an imaging spectrometer. Other applications include airborne HS data collection for natural resource and military sensing. Improved performance by increasing signal-to-noise-ratio could be achieved by use of compact, high irradiance laser sources, including the quantum cascade laser (QCL) and the broadband mixed isotope CO₂ laser.

Acknowledgments

R. Warren's work was supported by DBC Technology Corp. under contract W911-SR-14-C-0059 with the U. S. Army Edgewood Chemical Biological Center, Aberdeen Proving Ground, MD 21010.

S. Osher's work was partially supported by NSF Grant DMS-1118971.

References

1. M. E. Winter, "N-findr: an algorithm for fast autonomous spectral end-member determination in hyperspectral data," in *Imaging Spectrometry V*. 1999, vol. 3753, pp. 266-275, SPIE.
2. J. M. P. Nascimento and J. M. Bioucas Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898-910, Apr. 2005.
3. R. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing*, 2nd Ed., Academic Press, NY, 1997.
4. M. Dadon, S. Rotman, D. Blumberg, S. Adler-Golden, and P. Conforti, "Target detection in the presence of multiple subpixel targets in complex backgrounds," *Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, Los Angeles, CA, 21-24 August, 2016.
5. S. Adler-Golden and P. Conforti, "Robust hyperspectral detection with algorithm fusion," *IEEE International Geoscience and Remote Sensing Symposium*, Honolulu, HI, July 25-30, 2010.
6. C.-I Chang, *Hyperspectral Data Processing: Algorithm Design and Analysis*, John Wiley & Sons, Hoboken, NJ, 2013.
7. S. Kraut, L. Scharf, and L. McWhorter, "Adaptive subspace detector," *IEEE Trans. Signal Processing*, 49(12), 2001, 3005-3014.
8. R. Warren and D. Cohn, "Chemical detection on surfaces by hyperspectral imaging," *J. Appl. Remote Sens.* 11(1), 015013 (2017), doi: 10.1117/1.JRS.11.015013.
9. R. Warren and S. Osher, "Hyperspectral unmixing by the alternating direction method of multipliers," *Inverse Probl. Imaging*, 9(3), doi: 10.3934/ipi.2015.9.917.
10. S. Boyd, N. Parkh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, 3(1), 2010.
11. R. Lai and S. Osher, "A splitting method for orthogonally constrained problems," UCLA CAM Report 12-39, 2012.
12. T. Goldstein and S. Osher, "The split Bregman method for l_1 regularized problems," *SIAM J. on Imaging Sciences*, 2(2), 2009.
13. M. Hestenes, "Multiplier and gradient methods," *Journal of Optimization Theory and Applications*, 4 1969.

14. M. Powell, "A method for nonlinear constraints in minimization problems," in *Optimization* (ed. R. Fletcher), Academic Press, New York, 1972.
15. R. Rockafellar, "The multiplier method of Hestenes and Powell applied to convex programming," *Journal of Optimization Theory and Applications*, **12** 1973.
16. J. Nocedal and S. Wright, *Numerical Optimization*, Section 17.4, Springer, 1999.
17. Wide Area Reconnaissance (WAAR) Project sponsored by the Edgewood Chemical Biological Center with data made available through the Advanced Threat Detection Program run by DTRA and NSF. Data collected by the Applied Physics Laboratory with a Telops Hyper-Cam sensor in 2009. Information supplied by Alison Carr of APL.
18. Data set supplied by MIT/LL as part of the Gas Detection Challenge sponsored by DTRA in 2014.