

# Review of Linear Spectral Mixture Analysis

Bernard Lampe, *Student Member, IEEE*

**Abstract**—Linear spectral mixture analysis (LSMA) is a ubiquitous theory applied to hyperspectral images (HSI) to determine the presence and quantity of known materials by decomposing the pixels into parts. The analysis can be performed in a supervised or unsupervised manner. In the supervised case, the model assumes the pixels are each linearly mixed from a set of known spectral signatures. In the unsupervised case, the analysis must find the number of spectral signatures and the signatures themselves as well as decompose each pixel into a linear combination of those found signatures. The methods of finding the spectral signatures are referred to as HSI unmixing. In this study, we review, implement and test five supervised LSMA algorithms to unmix hyperspectral pixel vectors into parts. The purpose of this review is to understand the differences in the algorithm approaches to solving LSMA and the positives and drawbacks of using the algorithms. The accuracy and algorithm timing are then used to numerically evaluate the algorithms.

**Index Terms**—Hyperspectral, Linear Spectral Mixture Analysis (LSMA), Active Set Methods, Geometric Method, Orthogonal Subspace Projection, Fully-Constrained-Least-Squares (FCLS), Modified FCLS

## I. INTRODUCTION

LINEAR spectral mixture analysis (LSMA) takes advantage of the linear mixture model (LMM) and assumes that all the hyperspectral pixels can be represented as a constrained linear combination of vectors [1]. LSMA postulated that any HSI image of  $L$  bands is a linear combinations of a finite set of  $p$  pure signatures  $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_p$  where  $\mathbf{t}$  is a  $L \times 1$  vector. These pure signatures are quantized and measured and the resulting spectral signatures of dimension  $L \times 1$  are  $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_p$ . The linear mixture model then further states that each HSI pixel vector  $\mathbf{r}_i$  for  $i = 1$  to  $N$  image pixels is a linear combination of the spectral signatures as denoted in equation 1. This model arranges the spectral signatures into a matrix  $\mathbf{M}$  and the goal of supervised LSMA is to find the  $p$  constrained abundance fractions in  $\alpha_i$ . In order to have a model that is consistent with real hyperspectral images, the abundance fractions are constrained to sum-to-one (ASC) and to be non-negative (ANC). The model also includes an error term  $\mathbf{n}$  to account for model error, measurement error, noise and background.

There are two critical questions which must be resolved when using the linear mixture model for LSMA [2]. The first is how to estimate the number of spectral signatures will be required  $p$ . This parameter must be chosen with care to ensure that the model is sufficient and efficient. The second is how to determine the constrained abundance fractions. The abundance fractions are constrained to sum-to-one (ASC) and to be non-negative (ANC). In supervised LSMA, we are given the  $\mathbf{M}$  matrix and  $p$ . In this study, we review five supervised algorithms. The unsupervised unmixing problem is harder and

we use non-negative matrix factorization (NMF) to determine  $\mathbf{M}$  and  $\alpha_i$  simultaneously.

An interesting point of the LSMA algorithms is the way they view the LMM when deriving the algorithms. Each model assumes the pixels are a combination of the spectral signatures, but they vary in the way they assign a pixel to those signatures. We have noted three popular interpretations of the model. The first is the spectral components based model. This model is most popular and best represented by equation 1 where the pixel  $\mathbf{r}_i$  is an affine linear combination of non-negative abundance fractions. The second interpretation of the model is based on clustering algorithms. The set of spectral signatures  $\{\mathbf{m}_i\}_{i=1}^p$  are the cluster centers and the abundance fractions are the cluster assignments. Each pixel is then given an assignment value representing the clusters contribution to the pixel. The last noted understanding of the LMM is the simplex model. In this view the spectral signatures represent the vertices of a simplex. Then the abundance fractions represent the location of pixel vector inside the simplex. Each of these understandings of the model lead to unique algorithms for LSMA which will be relevant for our review.

$$\mathbf{r}_i = \mathbf{M}\alpha_i + \mathbf{n} \quad (1)$$

## II. FULLY CONSTRAINED LINEAR SPECTRAL MIXTURE ANALYSIS RESULTS

In this section we will review three FCLS algorithms for determining the constrained abundance fractions. Each algorithm will compute the abundance fractions given the spectral signatures  $\mathbf{M}$ . Each algorithm incorporates the sum-to-one constraint (ASC) and the non-negativity constraint (ANC) in different ways.

### A. Active Set Method

The active set method developed by Heinz and Chang can be used to solve the FCLS as an iterative algorithm which is initialized by the unconstrained least squares solution  $\alpha_i^{LS} = (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{M}^T \mathbf{r}_i$  [3]. The optimization utilizes the non-negative least squares constrained (NCLS) algorithm to refine the initial solution. The iteration begins by making two sets which track the passive and active components of the optimization. The unconstrained least squares solution can have negative values which violate the model. Those dimensions of the model become active while the remaining are passive. The active dimensions are used to perform a gradient descent until the negative components become non-negative. In order to impose the sum-to-one constraint (ASC), the spectral signature matrix and pixel are augmented to include a unitary dimension as in equation 2. This dimension enforces the ASC approximately by down weighting the other

dimension values by using  $\delta = 1/\max(m_{ij})$ . The vectors  $\tilde{\mathbf{r}}$  and the matrix  $\tilde{\mathbf{M}}$  take the place of  $\mathbf{r}$  and  $\mathbf{M}$  during the optimization.

$$\tilde{\mathbf{r}} = \begin{bmatrix} \delta \mathbf{r} \\ 1 \end{bmatrix}, \tilde{\mathbf{M}} = \begin{bmatrix} \delta \mathbf{M} \\ \mathbf{1}^T \end{bmatrix} \quad (2)$$

### B. Geometric Method

The geometric method for solving FCLS was put forth by Wang, Liu and Wang [4]. This method is a close form solution which takes advantage of Cramer's rule from calculus. This rule states that the solution of a critically determined linear system can be solved one component at a time using the ratio of determinants. The FCLS algorithm can take advantage of this by noting the determinant is the volume of the simplex formed by the vertices of the spectral signatures  $\{\mathbf{m}_i\}_{i=1}^p$ .

The algorithm proceeds by introducing the unitary dimension just as the active set method did in order to impose the ASC as in equation 3. Then the computation of the volume for each simplex is then performed to compute the abundance fractions as in equation 4.

$$\tilde{\mathbf{r}} = \begin{bmatrix} 1 \\ \mathbf{r} \end{bmatrix}, \tilde{\mathbf{M}} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ \mathbf{m}_1 & \mathbf{m}_2 & \dots & \mathbf{m}_p \end{bmatrix} \quad (3)$$

The determinant of the matrix  $\tilde{\mathbf{M}}$  is not defined because it is not square. Some algorithms perform PCA to project to a lower dimension to compute the projected volume. However, this algorithm computes the volume using equation 4.

$$\text{Vol}(\tilde{\mathbf{m}}_1, \tilde{\mathbf{m}}_2, \dots, \tilde{\mathbf{m}}_p) = \frac{1}{(p-1)!} \sqrt{|\tilde{\mathbf{M}}^T \tilde{\mathbf{M}}|} \quad (4)$$

The abundance fraction components are computed as the ratio of the volumes of the augmented spectral signature matrix and the unmodified spectral signature matrix. The volumes are computed so the sign of the numerator and denominator are the same which ensures the non-negativity of the result.

$$\alpha_i = \frac{\text{Vol}(\tilde{\mathbf{m}}_1, \dots, \tilde{\mathbf{m}}_{i-1}, \tilde{\mathbf{r}}, \tilde{\mathbf{m}}_{i+1}, \dots, \tilde{\mathbf{m}}_p)}{\text{Vol}(\tilde{\mathbf{m}}_1, \tilde{\mathbf{m}}_2, \dots, \tilde{\mathbf{m}}_p)} \quad (5)$$

### C. OSP Method

The last FCLS method reviewed takes advantage of the orthogonal subspace projection (OSP) projector to suppress the signatures not of interest. The projector is computed by removing the spectral signature of interest as in equation 6. Then computing the projector to suppress the other signatures as in equation 7. These  $p$  projectors are used to estimate the contribution of each signature. The magnitudes of the projections are the abundance fractions.

$$\tilde{\mathbf{M}} = [\mathbf{m}_1, \dots, \mathbf{m}_{i-1}, \mathbf{m}_{i+1}, \dots, \mathbf{m}_p] \quad (6)$$

$$\delta_{\mathbf{m}_i} = \mathbf{P}_U^\perp = \mathbf{I} - \tilde{\mathbf{M}} (\tilde{\mathbf{M}}^T \tilde{\mathbf{M}})^{-1} \tilde{\mathbf{M}}^T \quad (7)$$

## III. MODIFIED FULLY CONSTRAINED LINEAR SPECTRAL MIXTURE ANALYSIS

The modified FCLS (MFCLS) algorithm developed by Wong and Chang, makes use of two insightful constraints that allows for the Lagrange multiplier solution to be implemented [5]. The MFCLS approach takes advantage of the ASC but adds another constraint using the absolute value, abundance sum-to-one constraint (AASC) as in equation 8. These constraints results in an algorithm which is solvable with minimal iteration via Lagrange multiplier. There are two version of the algorithm we will review.

$$\sum_{i=1}^p \alpha_i = 1 \text{ and } \sum_{i=1}^p |\alpha_i| = 1 \quad (8)$$

### A. Lagrange Multiplier

The first version of the algorithm is below. Note that the algorithm is initialized by the  $\alpha^{SCLS}$  which is the sum-to-one constrained abundance solution presented by Heinz and Chang [3]. Also, note the operator  $\text{sign}(\alpha)$  is the vector consisting of the sign of each entry such as  $\text{sign}(\alpha_i) = \alpha_i/|\alpha_i|$ .

- 1) Initial condition is  $\alpha^{MFCLS} = \alpha^{SCLS}$
- 2) Compute  $\lambda_1$  and  $\lambda_2$  using constraints  $\sum_{i=1}^p \alpha_i = 1$  and  $\sum_{i=1}^p |\alpha_i| = 1$
- 3)  $\alpha^{MFCLS} = \alpha^{SCLS} - (\mathbf{M}^T \mathbf{M})^{-1} [\lambda_1 \mathbf{1} + \lambda_2 \text{sign}(\alpha^{SCLS})]$
- 4) Iterate if any negative components left

The second step of the MFCLS algorithm is to compute  $\lambda_1, \lambda_2$ . These are computed using the ASC and AASC. When forming the problem, the following matrix in equations 9 and 10 are derived. This  $2 \times 2$  linear system is readily solvable for  $\lambda_1$  and  $\lambda_2$ .

$$\begin{bmatrix} 1 - \mathbf{1}^T \alpha^{MFCLS} \\ 1 - \text{sign}(\alpha_{LS})^T \alpha^{MFCLS} \end{bmatrix} \mathbf{X} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} \quad (9)$$

$$\mathbf{X} = \begin{bmatrix} \mathbf{1}^T (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{1} & \mathbf{1}^T (\mathbf{M}^T \mathbf{M})^{-1} \text{sign}(\alpha_{LS}) \\ \text{sign}(\alpha_{LS})^T (\mathbf{M}^T \mathbf{M})^{-1} \mathbf{1} & \text{sign}(\alpha_{LS})^T (\mathbf{M}^T \mathbf{M})^{-1} \text{sign}(\alpha_{LS}) \end{bmatrix} \quad (10)$$

### B. Iterative Algorithm

The iterative algorithm for MFCLS put forth by Wong and Chang is similar to the active set method [5]. The algorithm chooses an active set and performs gradient descent on the dimensions chosen. However, the steering matrix computed during gradient descent is augmented to include the ASC and AASC constraints. This results in few iterations of the algorithm when compared to active set FCLS.

$$\mathbf{s} = \begin{bmatrix} \delta \mathbf{r} \\ 1 \end{bmatrix}, \mathbf{N} = \begin{bmatrix} \delta \mathbf{M} \\ \mathbf{1}^T \end{bmatrix} \quad (11)$$

$$\arg \min_{\hat{\mathbf{M}}, \hat{\mathbf{A}}} \left\{ \|\mathbf{X} - \hat{\mathbf{M}} \hat{\mathbf{A}}\|^2 \right\} \quad (12)$$

#### IV. HYPERSPECTRAL DATA

We used the simulated data to test each of the LSMA and NMF algorithms. We did experiments on the TI2 and TE2 data sets described in Dr. Chang's book [6]. The data sets each take spectral signatures extracted from the Cuprite data set for five different materials. The full image with the extracted data set spatial locations are in figure 1. The spectral signatures are plotted in figure 2. The five material types are 1Alunite, Buddingtonite, Calcite, Kaolinite and Muscovite. Each spectral signature was added to the image in an intentional programmatic way to demonstrate algorithm performance. Each material is given a primary row and the first column is a  $4 \times 4$  pure pixel panel, the second column is a  $2 \times 2$  pure pixel panel, the third column is a  $2 \times 2$  pixel panel where each pixel is a 50% mixture of the primary material for the row and the other four materials. The fourth column is a  $1 \times 1$  sub-pixel signature which consists of 50% signature and 50% background. The fifth column is also a subpixel with 25% material and 75% background. Each image has a 20 db SNR. The TI image (target implant) is designed by removing the background pixel and inserting the pure signature pixel. The TE image (target embedded) is designed by super-imposing the pure signature and the background as a linear combination. Therefore, for a TI2 pixel, the panel pixels are described as  $\mathbf{r}_i = \mathbf{m}_i + \mathbf{n}$  and the TE2 panel pixels are described as  $\mathbf{r}_i = \mathbf{m}_i + \mathbf{b} + \mathbf{n}$  where  $\mathbf{b}$  is a background signature.

#### V. EXPERIMENT RESULTS

The numeric experimental results are shown in Table I and Table II. The accuracy of the result was computed as the MSE of the abundance fractions of the panel pixels only. The background pixels were ignored because we do not have ground truth for them. The abundance fractions for the first two columns of panels should be 1 for all panel pixels. For the third and fourth columns the abundance should be 0.5 and the fifth column should be 0.25. You can see from Table I that the accuracy for the TI2 image is much better overall than the TE2. The accuracy of TI2 is much higher because the spectral signatures implanted in the image were not corrupted by background. Another way of seeing why this is the case is that the pixels of TI2 are within the simplex volume of the spectral signatures used in FCLS in  $\mathbf{M}$ . Within the simplex volume the matrix  $\mathbf{M}^T \mathbf{M}$  used for gradient descent is convex. However outside the simplex volume the problem may not be convex and could diverge when doing gradient descent. This occurs in the TE2 case where the pixels are added to a background pixel which is not accounted for in the spectral signature matrix  $\mathbf{M}$ .

Another observation is that the active set and modified FCLS algorithms worked best overall. These algorithms are initialized with the  $\alpha^{LS}$  and the  $\alpha^{SCLS}$  solutions which are close to the optimum. Then the final solutions are computed using the gradient via gradient descent or Lagrange multiplier.

We evaluated the algorithm run times and found that the geometric algorithm is fastest overall while the iterative MFCLS is comparably slow.

TABLE I: Algorithm accuracy in MSE, a. Active Set FCLS b. Geometric FCLS c. OSP FCLS d. MFCLS Lagrange e. MFCLS Iterative

	A.	B.	C.	D.	E.
TI2	0.003489	0.00394	0.01672	0.003536	0.003538
TE2	0.709771	0.13506	0.02976	5.677302	143345.3

TABLE II: Algorithm timing in seconds, a. Active Set FCLS b. Geometric FCLS c. OSP FCLS d. MFCLS Lagrange e. MFCLS Iterative

3.263822	1.762295	1.880125	2.453860	9.055446
----------	----------	----------	----------	----------

The abundance fractions for all FCLS algorithms on TI2 data are displayed in figures 4 to 8. The results are much more accurate when compared to the TE2 results in figures 9 through 13. However, in the TE2 case the OSP algorithm did much better than all other algorithms on TE2. This is because the algorithm suppresses the other spectral signatures without modifying the primary signature.

The results for the direct NMF algorithm are in figures 14 to 17. While this algorithm was initialized with the spectral signatures matrix  $\mathbf{M}$ , the results are not impressive. The algorithm quickly ends up in a bad state. We see the same effect in figures 18 to 21 for the sparse NMF algorithm.

#### VI. CONCLUSIONS

The FCLS works well on the TI2 data set because the spectral signature set accounts for the panel pixels in the image. The TE2 image results are poor because the background added to the panel pixels is not accounted for in the spectral signature matrix. The active set and modified FCLS algorithms were most accurate overall on the TI2 data set. The OSP was the best performing algorithm on the TE2 data set. Overall, the geometric algorithm was the fastest and almost as accurate as the other FCLS algorithms.

The alternating least squares and multiplicative updates methods for the non-negative matrix factorization results were poor even when enforcing sum-to-one using the augmented mixture matrix with unity dimension. Adding sparsity and incoherence to the algorithm helped and was necessary for the unsupervised method.

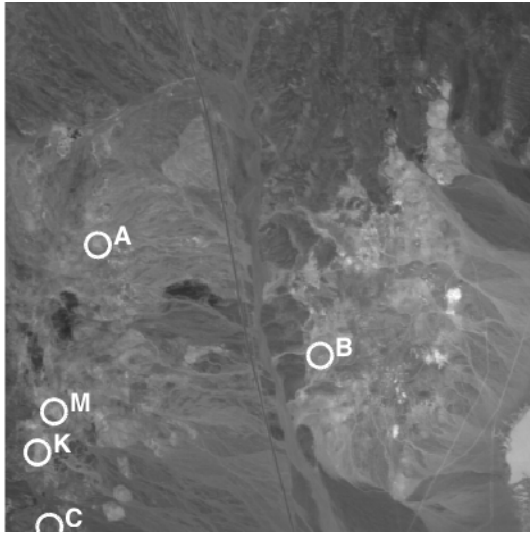


Fig. 1: Cuprite Data, Band 80, Marking Materials

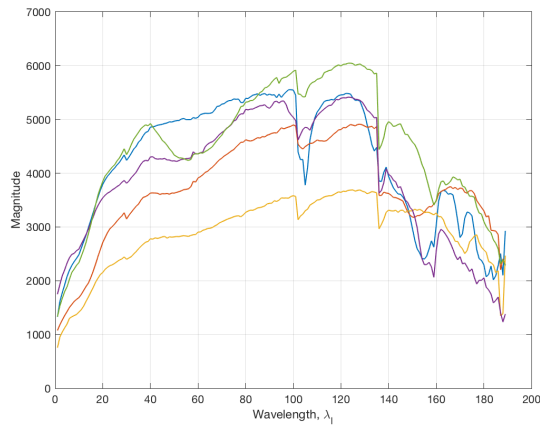


Fig. 2: Plotted reflectance for A, B, C, K, and M materials

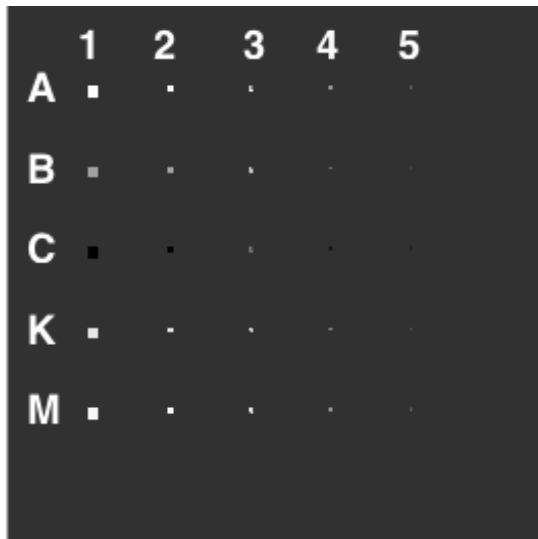


Fig. 3: Synthetically generated data pure pixels in the Cuprite image TE2.

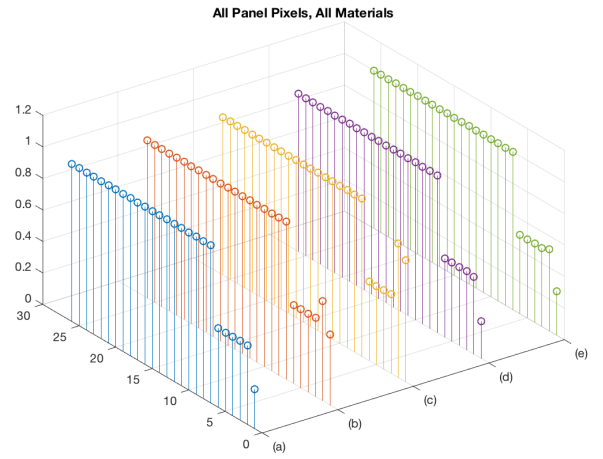


Fig. 4: FCLS, Active Set Method for TI2

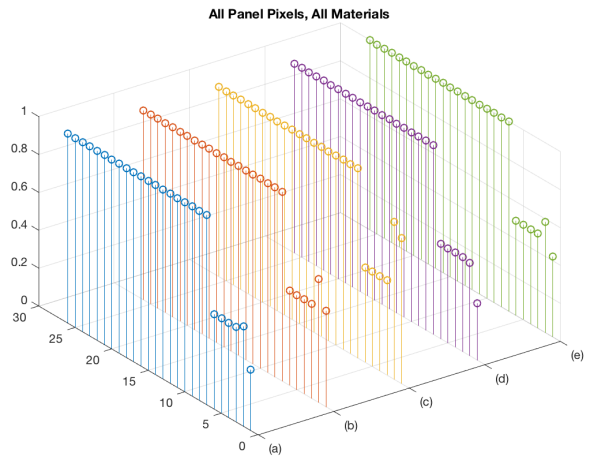


Fig. 5: FCLS, Geometric Volume Method for TI2

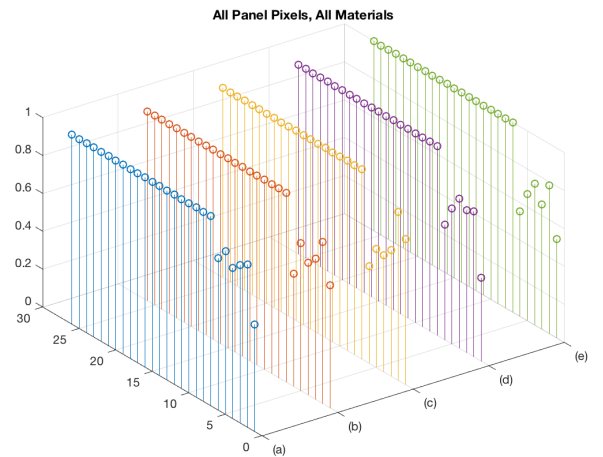


Fig. 6: FCLS, OSP Method for TI2

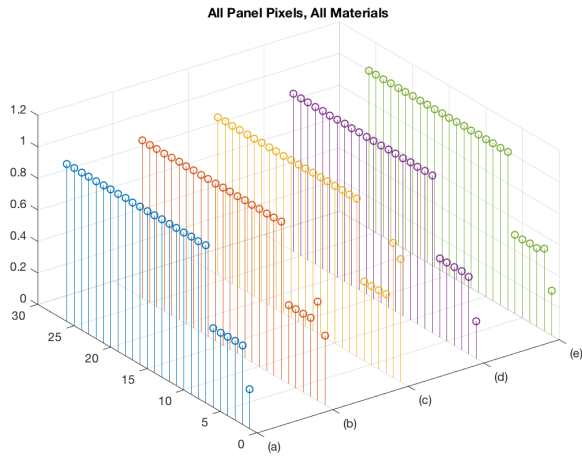


Fig. 7: MFCLS, Lagrange Multiplier Method for TI2

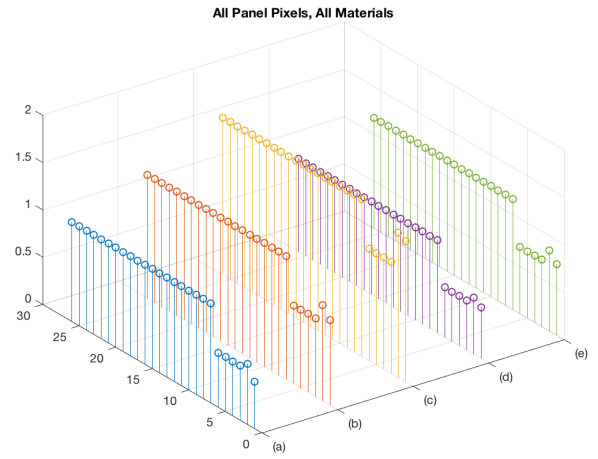


Fig. 10: FCLS, Geometric Volume Method for TE2

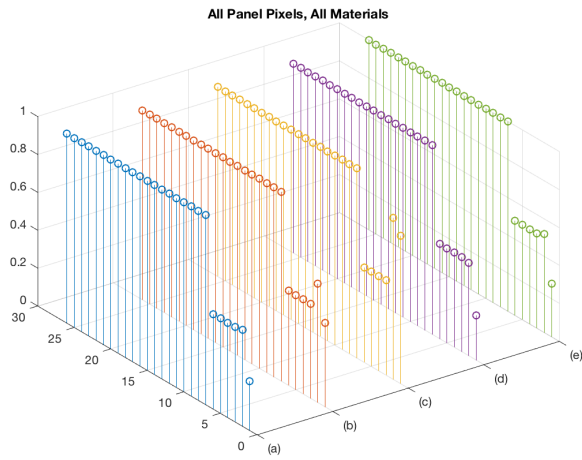


Fig. 8: MFCLS, Augmented Active Set Method for TI2

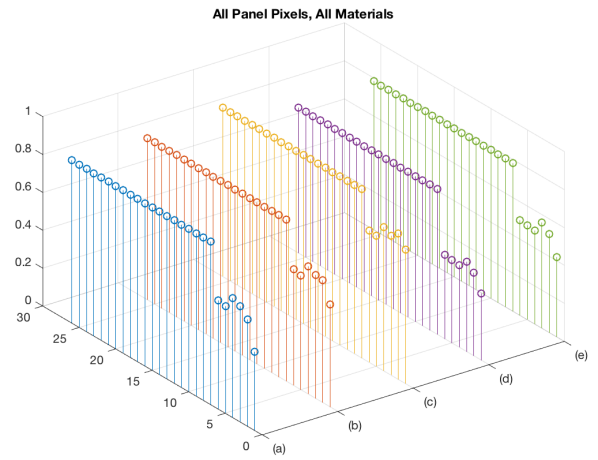


Fig. 11: FCLS, OSP Method for TE2

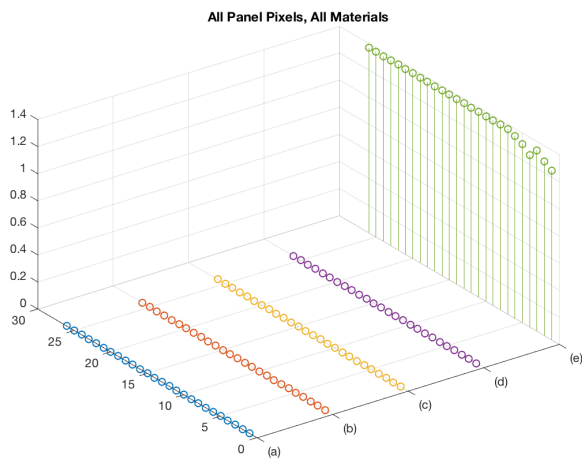


Fig. 9: FCLS, Active Set Method for TE2

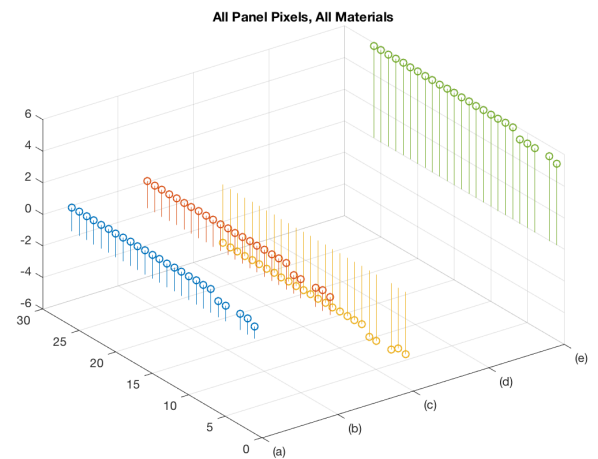


Fig. 12: MFCLS, Lagrange Multiplier Method for TE2

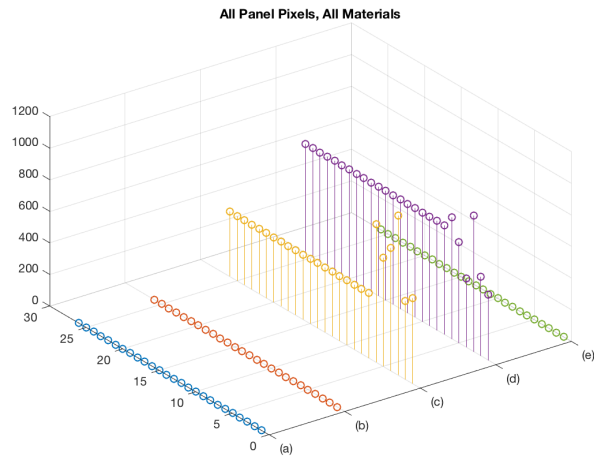


Fig. 13: MFCLS, Augmented Active Set Method for TE2

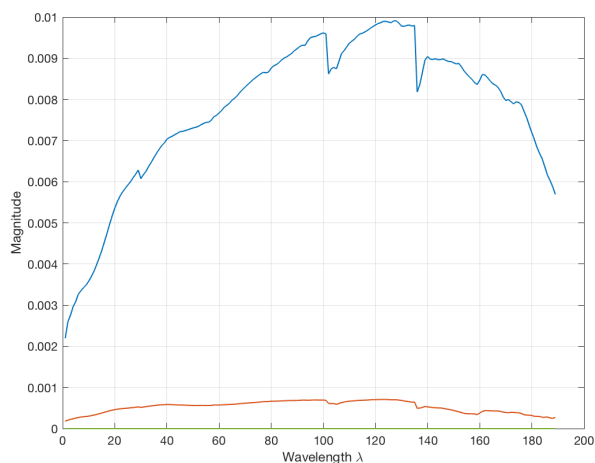


Fig 14. Direct NMF TI2 Endmembers

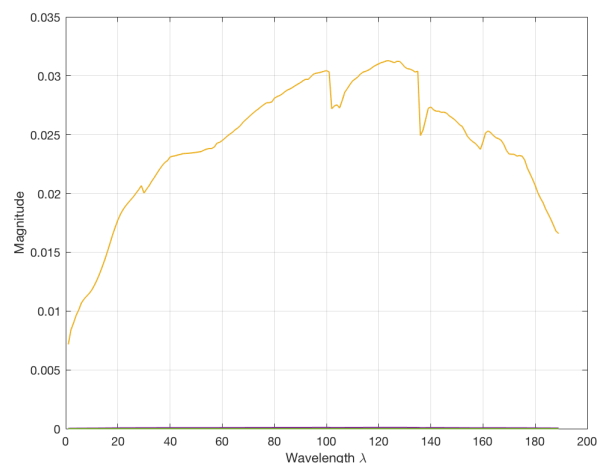


Fig 15. Direct NMF TE2 Endmembers

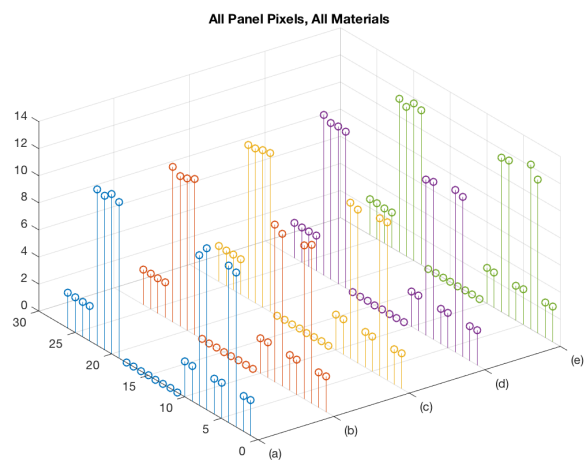


Fig 16. Direct NMF TI2 Abundance Fractions

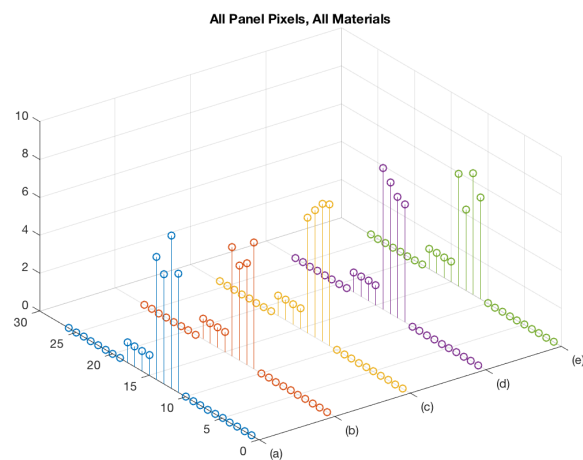


Fig 17. Direct NMF TE2 Abundance Fractions

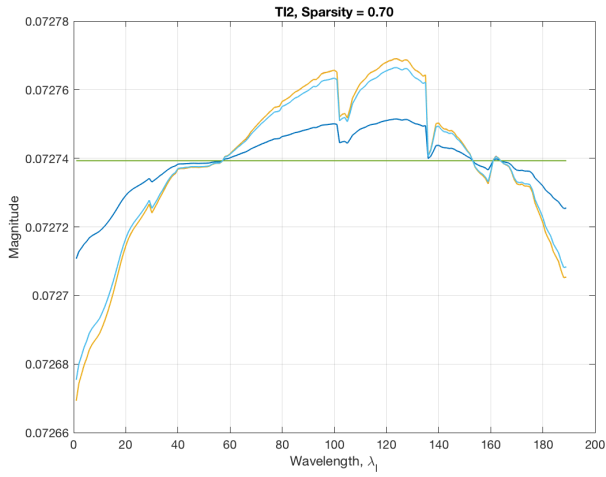


Fig. 18 Endmembers for Sparse NMF TI2 sparsity = 0.70

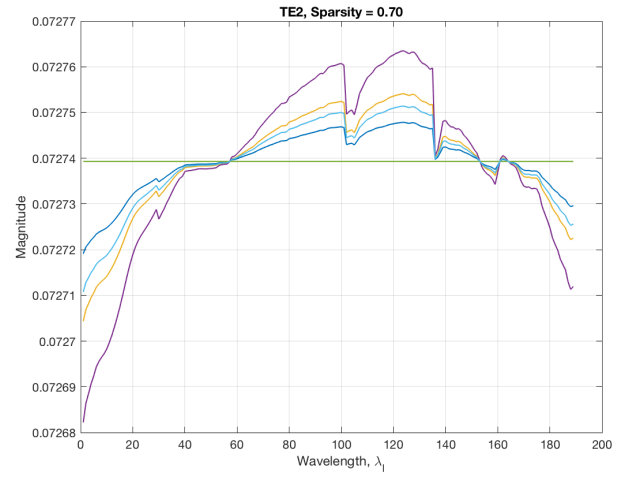


Fig 19. Endmembers for Sparse NMF TE2 sparsity = 0.70

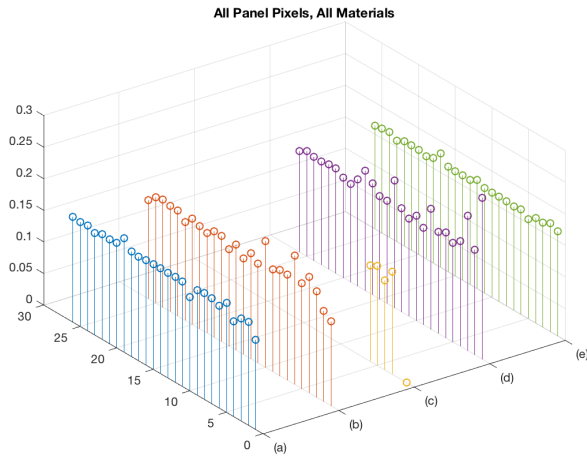


Fig 20. Abundance Fractions for TI2 and sparsity = 0.70

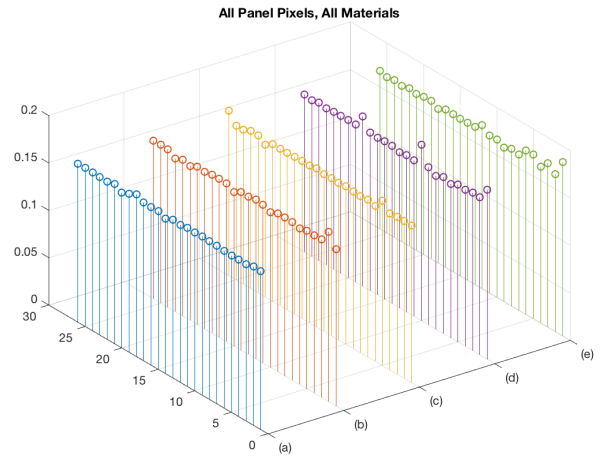


Fig 21. Abundance Fractions for TE2 and sparsity = 0.70



## REFERENCES

- [1] C.-I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. Plenum Publishing Co., 2003.
- [2] —, *Real-Time Progressive Hyperspectral Image Processing*, 1st ed. Springer Publishing Company, Incorporated, 2016.
- [3] D. C. Heinz and C. Chang, “Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery,” *IEEE Trans. Geoscience and Remote Sensing*, vol. 39, no. 3, pp. 529–545, 2001.
- [4] L. Wang, D. Liu, and Q. Wang, “Geometric method of fully constrained least squares linear spectral mixture analysis,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 6, June 2013.
- [5] E. Wong and C.-I. Chang, “Modified fully abundance-constrained spectral unmixing,” *Proc. SPIE 8539, High-Performance Computing in Remote Sensing II*, 2012.
- [6] C.-I. Chang, *Hyperspectral Data Processing*. John Wiley & Sons, Inc., 2013.
- [7] D. D. Lee and H. S. Seung, “Algorithms for non-negative matrix factorization,” in *Advances in Neural Information Processing Systems 13*, T. K. Leen, T. G. Dietterich, and V. Tresp, Eds. MIT Press, 2001, pp. 556–562.
- [8] H. S. S. Daniel D. Lee, “Learning the parts of objects by non-negative matrix factorization,” *Nature*, Vol. 401, No. 6755, pp. pp. 788–791, October 1999.
- [9] H. Q. Lidan Miao, “Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 3, 2007.
- [10] O. Mandula, *nmfpack*, 2011, <https://github.com/aludnam/MATLAB/tree/master/nmfpack>.
- [11] P. Hoyer, “Non-negative sparse coding,” *Neural Networks for Signal Processing*, vol. 12, pp. 557 – 565, February 2002.
- [12] —, “Non-negative matrix factorization with sparseness constraints,” *Machine Learning Research*, vol. 5, pp. 1457 – 1469, 2004.
- [13] S. Jia and Y. Qian, “Constrained nonnegative matrix factorization for hyperspectral unmixing,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 1, January 2009.

PLACE  
PHOTO  
HERE

**Bernard Lampe** (S’09) became a IEEE Student Member (S) in 2009 and received his bachelors of computer science degree from The University of Michigan in Ann Arbor, Michigan, USA in 2009. He received his masters in electrical engineering from the University of Maryland Baltimore County (UMBC) in 2016. His interests are computer vision, image processing and compressive sensing. He is currently pursuing his doctorate in electorol engineering at UMBC.