# CONTEXT-BASED ENDMEMBER DETECTION FOR HYPERSPECTRAL IMAGERY

*Alina Zare and Paul Gader*

University of Florida
Department of Computer and Information Science and Engineering
Gainesville, FL 32611 USA

## ABSTRACT

An endmember detection algorithm that simultaneously partitions an input data set into distinct contexts, estimates endmembers, number of endmembers, and abundances for each partition is presented. In contrast to previous endmember detection algorithms based on the Convex Geometry Model, this method is capable of describing non-convex sets of hyperspectral pixels. Endmembers are found for non-convex regions by partitioning the set of pixels into convex regions using the Dirichlet Process and determining unique endmembers for each region. This novel endmember detection method naturally produces a classifier with a reject class. The algorithm can effectively identify to which context a test data point belongs and identify test pixels for which the associated context is unknown. Results are shown on AVIRIS Indian Pines Hyperspectral data. The results show the classification capability of this context-based endmember algorithm.

***Index Terms***— Hyperspectral, Endmember, Spectral Unmixing, Convex Geometry Model, Context, Dirichlet.

## 1. INTRODUCTION

The convex geometry model assumes that each pixel in a hyperspectral image can be described with a convex combination of the endmembers [1].

$$\mathbf{x}_i = \sum_{k=1}^{M} a_{ik}\mathbf{e}_k + \epsilon_i \quad i = 1, \ldots, N \tag{1}$$

where $N$ is the number of pixels, $M$ is the number of endmembers, $\epsilon_i$ is an error term, $a_{ik}$ is the abundance of endmember $k$ in pixel $i$, and $\mathbf{e}_k$ is the $k^{th}$ endmember. The abundances of this model satisfy the constraints in Equation 2,

$$a_{ik} \geq 0 \quad \forall k = 1, \ldots, M; \quad \sum_{k=1}^{M} a_{ik} = 1. \tag{2}$$

Several endmember detection algorithms are described in the literature. Previous endmember detection methods include methods that rely on the pixel purity assumption and search for endmembers within the data set [2], are based on Non-Negative Matrix Factorization [3, 4], use Independent Components Analysis [5, 6], and others. These methods generally search for a single set of endmembers to describe a hyperspectral scene. The presented algorithm partitions the input hyperspectral set into distinct contexts using the Dirichlet Process. Each context has its own set of endmembers.

Section 2 describes the Robust Sparsity Promoting Iterated Constrained Endmembers (R-SPICE) algorithm used to compute the endmembers within each convex region. Section 3 describes the Piece-Wise Convex Endmember (PCE) algorithm used to partition the data and find a unique set of endmembers for each partition.

## 2. ROBUST SPICE

Robust SPICE (R-SPICE) is a novel algorithm used to determine endmembers and the number of endmembers. R-SPICE is an extension of the Sparsity Promoting Iterated Constrained Endmembers (SPICE) algorithm [7]. SPICE simultaneously determines endmember spectra, estimates abundance values, and determines the number of endmembers. SPICE determines the number of endmembers by using a sparsity promoting Laplacian prior on the abundance values. The R-SPICE algorithm differs from SPICE by weighting each pixel based on its similarity to the other pixels in the data set. Therefore, outliers have little effect on endmember estimation. The weights on each pixel cause the results of the R-SPICE algorithm to be more stable over a wide range of parameter values as compared to SPICE. The objective function in R-SPICE is $ln\left(p\left(\mathbf{X}|\mathbf{A},\mathbf{E}\right)p\left(\mathbf{E}\right)p\left(\mathbf{A}\right)\right)$ where $p\left(\mathbf{X}|\mathbf{A},\mathbf{E}\right)$ is the likelihood of the data given the endmembers and abundances, $p\left(\mathbf{E}\right)$ is the prior on the endmembers, and $p\left(\mathbf{A}\right)$ is the sparsity promoting prior on the abundances. The data term is given by a weighted squared error term between the input data points and the estimated endmembers and abundances,

$$p\left(\mathbf{X}|\mathbf{A},\mathbf{E}\right) = \exp\left\{\frac{-1}{2}\sum_{i=1}^{N}\frac{w_i}{\nu}\left(\mathbf{x}_i - \sum_{k=1}^{M}a_{ik}\mathbf{e}_k\right)^T\left(\mathbf{x}_i - \sum_{k=1}^{M}a_{ik}\mathbf{e}_k\right)\right\} \tag{3}$$

978-1-4244-4687-2/09/$25.00 

where $w_i$ is the weight assigned to each data point and $\nu$ is used to weight the importance of the error term in comparison to the other terms in the objective function. In the current implementation of R-SPICE, $w_i$ is set to $\frac{1}{1+\frac{d_i}{\eta}}$ where $d_i$ is the distance from $\mathbf{x}_i$ to the mean of the data set and $\eta$ is the mean of the $d_i$ values. The priors on the endmembers and abundances are the same as those used in the SPICE algorithm. The prior on the endmembers is given by a term related to the volume enclosed by the endmembers,

$$p(\mathbf{E}) = \exp\left\{-\frac{1}{2\beta}\sum_{k=1}^{M-1}\sum_{l=k+1}^{M}(\mathbf{e}_k - \mathbf{e}_l)^T(\mathbf{e}_k - \mathbf{e}_l)\right\} \quad (4)$$

where $\beta$ is a parameter assigning weight to the endmember prior. The sparsity promoting abundance prior is given by,

$$p(\mathbf{A}) = \exp\left\{-\frac{1}{2}\sum_{k=1}^{M}\gamma_k\sum_{i=1}^{N}|a_{ik}|\right\} \quad (5)$$

where $\gamma_k = \frac{\Gamma}{\sum_{i=1}^{N} a_{ik}}$ using the previous iteration's abundance values.

The R-SPICE objective function is iteratively maximized. While holding endmembers constant, the abundance values are updated using quadratic programming. Next, the endmembers are updated. R-SPICE iterates between these two steps until some convergence criterion is met. To determine the number of endmembers, R-SPICE begins with a large number of endmembers and the sparsity promoting prior is drives abundance values of unneeded endmembers to zero. Therefore, endmembers are removed if their associated maximum abundance value falls below some threshold.

## 3. DETERMINING IMAGE CONTEXTS

In addition to determining endmembers and abundances, the proposed algorithm partitions the hyperspectral pixels where each partition, or context, has a set of associated endmembers. The number of contexts, and, therefore, the number of sets of endmembers, is determined with the Dirichlet Process. This method, Piece-wise Convex Endmembers (PCE), was originially used with Endmember Distribution (ED) detection [9]. This method is unsupervised. In this paper, PCE is applied with the R-SPICE algorithm so that the number of endmembers in each partition may be autonomously determined.

The algorithm uses a Gibbs sampling method to sequentially sample either an existing or a new set of endmembers for each data point. The probability of sampling an existing set of endmembers is determined by Equation 6,

$$\begin{aligned} &P(r_i = r_j \quad j \neq i|\mathbf{r}_{-i}, \mathbf{x}_i) \\ &= C\frac{n_{-i,j}}{\alpha + N - 1}\int f(\mathbf{x}_i|\mathbf{a}_i^{r_j}, \mathbf{E}^{r_j})H_{-i,r_j}(\mathbf{E}^{r_j}, \mathbf{A}^{r_j})d\mathbf{a}_i^{r_j}\mathbf{E}^{r_j} \\ &= C\frac{n_{-i,j}}{\alpha + N - 1} \\ &\quad \mathcal{N}\left(TR\mathbf{x}_i + SR\frac{\sum_{m\in r_j} w_m a_m \mathbf{E}^{r_j}}{\sum_{m\in r_j} w_m}, S + TRS\right) \end{aligned} \quad (6)$$

where $r_i$ is the indicator variable for the current data point, $\mathbf{x}_i$, $C$ is a normalization constant, $n_{-i,j}$ is the number of data points excluding $\mathbf{x}_i$ in partition $r_j$, $N$ is the total number of data points, and $\alpha$ is the innovation parameter for the Dirichlet process. Equation 6 comes directly from the data term in the R-SPICE objective function where $f(\mathbf{x}_i|\mathbf{a}_i^{r_j}, \mathbf{E}^{r_j}) = \exp\left\{\frac{-1}{2\nu}(\mathbf{x}_i - \mathbf{a}_i^{r_j}\mathbf{E})^T(\mathbf{x}_i - \mathbf{a}_i^{r_j}\mathbf{E}^{r_j})\right\}$. The likelihood value, $f$, measures the ability of a set of endmembers to represent a data point.

The distribution, $H_{-i,r_j}(\mathbf{E}^{r_j}, \mathbf{A}^{r_j})$ is the prior distribution updated based on the data points assigned to the ${r_j}^{th}$ partition,

$$\begin{aligned} &H_{-i,r_j}(\mathbf{E}^{r_j}, \mathbf{A}^{r_j}) \\ &= \prod_{m\in r_j}\exp\left\{\frac{-w_m}{2\nu}(\mathbf{x}_m - \mathbf{a}_m^{r_j}\mathbf{E}^{r_j})^T(\mathbf{x}_m - \mathbf{a}_m^{r_j}\mathbf{E}^{r_j})\right\} \end{aligned} \quad (7)$$

The integral of $f$ and $H$ results in the Normal distribution shown in Equation 6 where the matrices T, S, and R are $\frac{\nu}{\sum_{m\in r_j} w_m}\mathbf{I}$, $\nu\mathbf{I}$, and $(T+S)^{-1}$, respectively.

The probability of sampling a new partition is determined by Equation 8.

$$\begin{aligned} P(r_i \neq r_j \quad \forall j \neq i|\mathbf{r}_{-i}, \mathbf{x}_i) \\ = C\frac{\alpha}{\alpha + N - 1}\int f(\mathbf{x}_i|\mathbf{E}^*)G_0(\mathbf{E}^*)d\mathbf{E}^* \\ = \mathcal{N}\left(M, \left(\mathbf{V}_0^{-1} + \nu\mathbf{I}^{-1}\right)^{-1} + \nu\mathbf{I}\right) \end{aligned} \quad (8)$$

where $M = \mathbf{V}_0(\mathbf{V}_0 + \nu\mathbf{I})^{-1}\mathbf{x}_i + \nu\mathbf{I}(\mathbf{V}_0 + \nu\mathbf{I})^{-1}\mu_0$, and the prior distribution, $G_0$, is Gaussian where the mean, $\mu_0$, is the mean of the input pixels and the covariance, $\mathbf{V}_0$, is constant,

$$G_0 = \mathcal{N}\left(\mu_0 = \frac{1}{N}\sum_{j=1}^{N}\mathbf{x}_j, \mathbf{V}_0\right) \quad (9)$$

The prior distribution combined with $\alpha$, the innovation parameter in the Dirichlet process prior, dictates the probability of generating a new partition. The covariance matrix, $\mathbf{V}_0$, is set to approximate a uniform prior over the data set.

A partition is sampled for the data point under consideration according to the probabilities in Equations 6 and 8. After

assigning a data point to a partition, the weight, $w_i$, and the abundance vector, $\mathbf{a}_i^{r_j}$, for the data point are updated. Then, endmembers of that partition are updated. These values are updated by minimizing the product of Equations 3, 4 and 5.

This sampling process can only remove endmembers by driving abundance values to zero. However, since several data points may be added to a partition, additional endmembers may be required to adequately describe the associated pixels. Therefore, a schedule is employed so that the R-SPICE algorithm can be run on each partition. The R-SPICE algorithm determines the number of endmembers by starting with a large number of endmembers and removing the unneeded endmembers. Therefore, when initializing R-SPICE, the endmember set of each partition is enlarged with several randomly chosen endmembers. During R-SPICE, these endmembers are then refined and, if their abundance values go to zero, removed. In this way, after running R-SPICE, additional endmembers may be added to some contexts. This algorithm is summarized in the psuedo-code shown below.

PCE-RSPICE(X)

```
Initialize Partitions
for r ← 1 to R_initial partitions do
    Initialize E^r and A^r using R-SPICE
end for
for k ← 1 to number of total iterations do
    for i ← 1 to number of Gibbs sampling iterations do
        Randomly reorder data points in X
        for j ← 1 to number of data points do
            Remove x_j from its current partition
            Update the partition's E^{r_j} values
            Compute Dirichlet process partition probabilities
            for x_j using Equations 6 and 8.
            Sample a partition for x_j based on the Dirichlet
            process partition probabilities
            Update E^{r_j}, a_i^{r_j}, and w_i values for the new partition
        end for
    end for
    for r ← 1 to R_k partitions do
        Update E^r and A^r using R-SPICE
    end for
end for
```

When using Gibbs sampling, the final result is a distribution of values. However, in the current implementation, rather than examining the entire distribution of partitions created by this Gibbs sampler, the final result is set to the final sample after applying a post-processing step. This post-processing step assigns each data point to the partition whose associated likelihood value is largest. Also, an R-SPICE iteration is applied to the final partitioning. This post-processing step is effective at removing small partitions in the last sample whose points are adequately described by a larger partition's endmembers.

## 4. EXPERIMENTAL RESULTS

This method was applied to the June 1992 AVIRIS Indian Pines data set. This data was collected over the Indian Pines Test site in an agricultural area of northern Indiana. The image has $145 \times 145$ pixels with 220 spectral bands and contains approximately two-thirds agricultural land and one-third forest and other elements [10]. The Indian Pines data was reduced from 220 bands to 3 bands using Hierarchical Dimensionality reduction [11]. Figure 1 shows the result of PCE-RSPICE applied to woods and soybean-notill pixels [12].

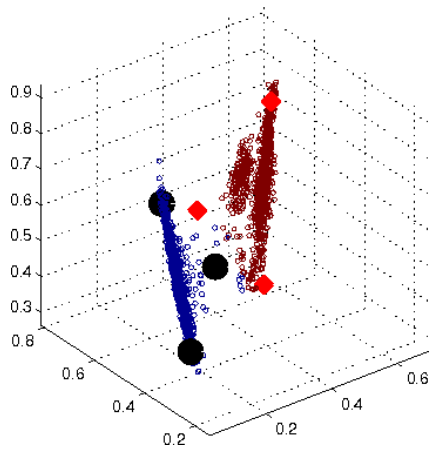
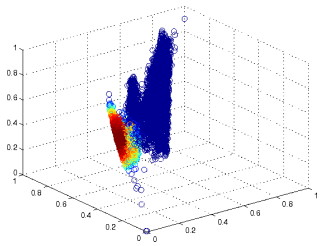



**Fig. 1**. Endmember and partitioning results on Woods and soybean-no till classes. Large circles indicate endmembers found for partition 1 (woods) and large squares indicate endmembers found for partition 2 (soybean-notill).

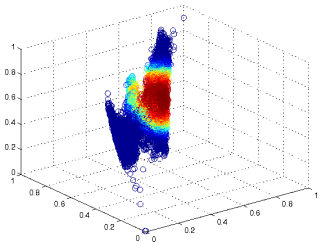
The parameters used in these results were $\alpha = 1$, $V_0 = 0.7$, $\nu = 0.05$, $\beta = 0.06$ and $\Gamma = 0.01$. The post-processing parameters used in this result were $\alpha = 100$, $V_0 = 0.7$, $\nu = 0.01$, $\beta = 0.06$ and $\Gamma = 0.01$. As seen in the figure, these points do not define a convex region which was also seen after applying PCA for dimensionality reduction. The PCE-RSPICE algorithm was effective at partitioning the points into two sets and finding appropriate endmembers for the two classes of the input data.

After using the PCE-RSPICE algorithm to find endmembers and contexts for the woods and soybean-no till classes, the probability values for each partition and a new partition were computed for the remaining labeled pixels in the data using Equations 6 and 8. These values can be used as classifier outputs for the soybean-notill, woods and the reject class. Figure 2 shows maps of the probability values computed for the all the labeled pixels. For areas in which the associated context is unknown, this Dirichlet process method produces high values in the probability map for the reject class. This is useful for classification tasks where uncertainty in a data point's context needs to be indicated. In contrast, endmember methods that produce only one set of endmembers, such as SPICE, cannot classify pixels into a reject class. In maps

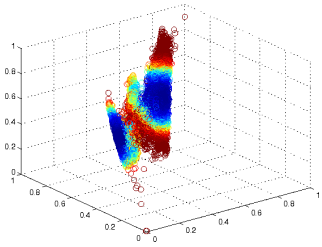
1 and 2, red points cover the regions where the input training pixels are located. In map 3, red indicates the area of the reject class.



(a) Probability map from woods partition



(b) Probability map from soybean partition



(c) Probability map from uncertainty class

**Fig. 2**. Probability maps computed for all labeled pixels in Indian Pines data. The area associated with the uncertainty class depends on the input samples. It has high probability outside of the convex regions found using PCE.

## 5. DISCUSSION AND CONCLUSIONS

PCE-RSPICE provides an autonomous method for partitioning the hyperspectral pixels into convex regions and computing endmembers, the number of endmembers, and abundances for each region. The Dirichlet process is employed to determine the number of needed convex regions. The results of this paper illustrate that the Dirichlet process provides a straight-forward way for defining regions of a reject class. Therefore, given some hyperspectral training set the PCE-RSPICE algorithm will find convex regions and associated endmembers. Then, probability values for test pixels can be computed for all partitions and a reject class. PCE-RSPICE's ability to compute endmembers, the number of endmembers, contexts and the number of contexts is significantly unique when compared to previous endmember detection algorithms. The method's natural ability to produce context classification outputs with a reject class is not provided by other endmember detection algorithms or even some classification methods.

## 6. REFERENCES


[1] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, pp. 44–57, 2002.

[2] M. E. Winter, "Fast autonomous spectral end-member determination in hyperspectral data," in *Proceedings of the Thirteenth International Conference on Applied Geologic Remote Sensing*, Vancouver, B.C., Canada, 1999, pp. 337–344.

[3] D. Lee and H. Seung, "Algorithms for non-negative matrix factorization," in *Advances in Neural Information Processing Systems 13*, 2000, pp. 556–562.

[4] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 3, pp. 765–777, Mar. 2007.

[5] T.-M. Tu, "Unsupervised signature extraction and separation in hyperspectral images: A noise-adjusted fast independent components analysis approach," *Optical Engineering*, vol. 39, no. 4, pp. 897–906, 2000.

[6] J. Wang and C.-I. Chang, "Applications of independent component analysis in endmember extraction and abundance quantification for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 9, pp. 2601–2616, Sept. 2006.

[7] A. Zare and P. Gader, "Sparsity promoting iterated constrained endmember detection for hyperspectral imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 4, no. 3, pp. 446–450, July 2007.

[8] C. Rasmussen, "The infinite Gaussian mixture model," in *Advances in Neural Information Processing Systems*, S. A. Solla, T. K. Leen, and K. R. Muller, Eds., vol. 12, pp. 554–560. MIT Press, 2000.

[9] A. Zare and P. Gader, "PCE: Piece-wise convex endmember detection," *IEEE Transactions on Geoscience and Remote Sensing*, Submitted.

[10] S. B. Serpico and L. Bruzzone, "A new search algorithm for feature selection in hyperspectral remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 7, pp. 1360–1367, July 2001.

[11] A. Martinez-Uso, F. Pla, J. M. Sotoca, and P. Garcia-Sevilla, "Clustering-based hyperspectral band selection using information measures," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 12, pp. 4158–4171, Dec. 2007.

[12] AVIRIS, "Free standard data products.," (2004, Sep) Jet Propulsion Laboratory, California Institute of Technology, Pasedena, CA. URL http://aviris.jpl.nasa.gov/html/aviris.freedata.html.