Piece-wise Convex Multiple Model Endmember Detection

Alina Zare, *Member, IEEE,* Ouiem Bchir, Hichem Frigui, *Member, IEEE,* and Paul Gader, *Senior Member, IEEE*

Abstract

A hyperspectral endmember detection and spectral unmixing algorithm that finds multiple sets of endmembers is presented. This algorithm, the Piece-wise Convex Multiple Model Endmember Detection (P-COMMEND) algorithm, models a hyperspectral image using a piece-wise convex representation. Each hyperspectral pixel is represented as a weighted sum of convex combinations of endmembers. By using a piece-wise convex representation, non-convex hyperspectral data are more accurately characterized. For example, the well-known Indian Pines hyperspectral image is used as an example of a piece-wise convex collection of pixels. The convex regions, weights, endmembers and abundances are found using an iterative fuzzy clustering method. Results indicate that the piece-wise convex representation provide endmembers that better represent hyperspectral data sets over methods that use a single convex region.

I. INTRODUCTION

The spectral signatures of the pure materials in a hyperspectral scene are referred to as *endmembers* [1]. *Spectral unmixing* is the task of decomposing pixels from a hyperspectral image into their respective endmembers and *abundances*. Abundances are the proportions of every endmember in each pixel in a hyperspectral image. The standard model used to perform spectral unmixing is the *convex geometry model* (also known as the *linear mixing model*). This model states that every pixel is a convex combination of endmembers in the scene. This has been shown in the literature to hold in cases where the spectra of the endmember are mixed by the spatial resolution of the imaging sensor [1–3]. If the convex geometry model holds, the endmembers are the spectra found at the corners of a convex region enclosing all the spectra in a hyperspectral scene. This model can be written as shown in Equation 1,

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

A. Zare and P. Gader are with the Department of Computer & Information Science & Engineering, University of Florida, Gainesville, FL, 32611 USA e-mail: azare@cise.ufl.edu, pgader@cise.ufl.edu. O. Bchir and H. Frigui are with the Department of Computer Engineering & Computer Science, University of Louisville, Louisville, KY, 40292 USA e-mail:h.frigui@louisville.edu

where N is the number of pixels in the image, M is the number of endmembers, ϵ_i is an error term, p_{ik} is the proportion of endmember k in pixel i, and $\mathbf{e}_k$ is the kth endmember. The proportions of this model satisfy the constraints in Equation 2,

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

In the algorithm presented here, the **P**iece-wise **Co**nvex **M**ultiple **M**odel **End**member Detection (P-COMMEND) algorithm, several sets of endmembers are found to describe hyperspectral image data. Each endmember set is found using the convex geometry model resulting in a piece-wise convex representation of the hyperspectral data. More precisely, the piece-wise convex model for an image pixel $\mathbf{x}_j$ is given by

$$\mathbf{x}_j = \sum_{i=1}^{C} \mu_{ij} \sum_{k=1}^{M_i} p_{ijk} \mathbf{e}_{ik} + \epsilon_i \tag{3}$$

where the $\sum_{i=1}^{C} \mu_{ij} = 1$, $\sum_{k=1}^{M_i} p_{ijk} = 1$, $p_{ijk} \geq 0$, M_i denotes the number of endmembers in convex model i, and p_{ijk} denotes the abundance for endmember k in convex model i.

Consider the AVIRIS Indian Pines [4] hyperspectral data set. When applying principal components analysis and reducing dimensionality down to three dimensions, the resulting data set appears piece-wise convex, as shown in Figure 1. This real hyperspectral data is not convex. Instead, a piece-wise convex representation would provide a better fit.
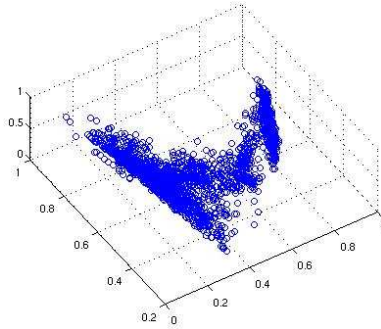

Fig. 1. The AVIRIS Indian Pines hyperspectral data set after applying PCA for dimensionality reduction to three dimensions. This illustrates that the Indian Pines hyperspectral data set is not convex but, instead, appears to be piece-wise convex.

The piece-wise convex representation for image pixels is appropriate since many, or perhaps all, of the pixels in an image will only be combinations of certain, and not all, materials in the scene. One could argue that a similar number of endmembers could be used in a single convex model as in a piece-wise convex model (with more than one piece). One would hope that for the single model, the pixels that are truly combinations of a subset of the endmembers will simply have zero abundance associated with the extra endmembers. Unfortunately, this is not the case, essentially because unions of convex sets are not necessarily convex. There can easily be cases for which the endmembers of convex pieces would never be chosen as the endmembers of the union of the convex pieces.

To be more concrete, consider the "spectral" samples in Figure 2(a) which look similar to the projection of the Indian Pines data but in two dimensions for illustrative purposes. Assume there are two distinct regions in the image so that every pixel in the image is truly generated by a convex combination of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_{true}$ or a convex combination of $\mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_5$. If endmembers are detected using a piece-wise model, the correct endmembers can be found. However, if they are detected using a single convex model, then the endmember $\mathbf{e}_{true}$ will not be found as an endmember, since it lies within the convex hull of the union of the two convex pieces. Hence, spectra in the left-hand triangle, T_1, will not be represented using the correct endmembers, and, therefore, their abundances would not be meaningful. One may take the position that $\mathbf{e}_{true}$ can be written as a convex combination of $\mathbf{e}_3$ and $\mathbf{e}_5$ so that if $\mathbf{x} \epsilon T_1$ with $\mathbf{x} = p_1\mathbf{e}_1 + p_2\mathbf{e}_2 + p_3\mathbf{e}_{true}$ then we would have $\mathbf{x} = p_1\mathbf{e}_1 + p_2\mathbf{e}_2 + p_3(a_1\mathbf{e}_3 + a_2\mathbf{e}_5)$ or $\mathbf{x} = p_1\mathbf{e}_1 + p_2\mathbf{e}_2 + p_3a_1\mathbf{e}_3 + p_3a_2\mathbf{e}_5$. However, this is not the correct answer either since the true endmember does not appear and cannot be recovered from the expression since only p_3a_1 and p_3a_2 are known and not p_3, a_1, or a_2. Hence, having more endmembers in a single model will not produce the correct results when the union of the convex pieces is not convex. The correct endmembers will not be found regardless of whether the abundances are zero or not.
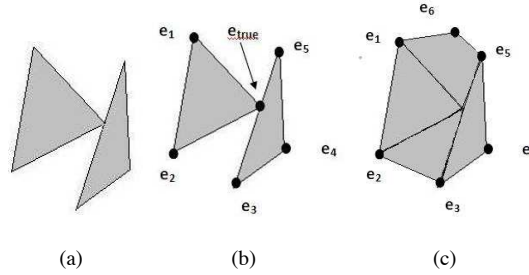

Fig. 2. Two-dimensional example illustrating the value of the piece-wise convex representation for non-convex data sets.

The piece-wise convex representation for hyperspectral imagery was first used by Zare in [5] as well as by Zare and Gader in [6] and [7]. These methods sample from a Dirichlet Process to partition the input hyperspectral data into convex regions. In contrast, the P-COMMEND algorithm applies a novel fuzzy clustering approach to the piece-wise convex representation for hyperspectral endmember detection and spectral unmixing. Although the sampling approach is more likely to find a global optimum, the P-COMMEND algorithm has the advantage over the Dirichlet process approach in that it is much faster.

II. RELATED WORK

A. Fuzzy Clustering

Clustering is the unsupervised classification of patterns into groups. It divides a set of unlabeled objects into clusters in such a way that objects in the same cluster are as similar as possible and objects in different clusters are as dissimilar as possible. In fuzzy clustering, data elements can belong to more than one cluster, and associated with each element is a set of membership levels. These indicate the strength of the association between that data

element and a particular cluster. Fuzzy clustering methods are good at dealing with the real world situation where clusters can overlap [8].

Fuzzy clustering is a process of assigning these membership levels, and then using them to assign data elements to one or more clusters. One of the most widely used fuzzy clustering algorithms is the Fuzzy C-Means (FCM) Algorithm [9]. The FCM algorithm [9] attempts to partition a set of feature vectors $X = \{\mathbf{X}_1, \ldots, \mathbf{X}_N\}$ into C fuzzy clusters. Given a set of feature vector, the algorithm returns a list of C cluster centers $\mathcal{C} = \{c_1, \ldots, c_C\}$ and a membership matrix $U = u_{ij} \in (0, 1)$, $i = 1, \ldots, C$, $j = 1, \ldots, N$ where each element tells the degree to which element $\mathbf{X}_j$ belongs to cluster c_i. The FCM [9] aims to minimize the following objective function:

$$J = \sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij}^{m} \|\mathbf{X}_j - c_i\|^2 \tag{4}$$

subject to

$$\sum_{i=1}^{C} u_{ij} = 1, \quad for\ 1 \leq j \leq N \tag{5}$$

In (4), m is the fuzzifier. It determines the level of cluster fuzziness. A large m results in smaller memberships u_{ij} and hence, fuzzier clusters. In the limit $m = 1$, the memberships u_{ij} converge to 0 or 1, which implies a crisp partitioning.

FCM has been generalized to find clusters of different shapes [10]. In [10], Bezdek et al. proposed the Fuzzy c-Varieties (FCV) algorithm to detect structures of data when all clusters are r-dimensional linear varieties, where r is less than the dimension of the data object. This is achieved by replacing the Euclidean distance in the FCM objective function (4) by the sum of the Euclidean distance and the scaled Euclidean distance mapped to the r principle scatter directions (r longest axes) of the data objects.

The gradual assessment of the membership of elements in a set was also applied in a regression framework [11]. Unlike classical regression methods which can only deal with a model at a time [12], the Multiple Model General Linear Regression MMGLR [11] attempts to estimate the regression parameters of multiple models simultaneously. The MMGLR [11] uses the residual of a point relative to a curve in a regression framework as a distance that represents the goodness of fit of a given model to the neighborhood of a point.

B. Endmember Detection Methods

Many endmember detection and spectral unmixing algorithms have been developed in the literature. A number of these methods rely on the pixel purity assumption and assume the endmembers can be found within the data set [13–16]. Methods have also been developed based on Non-Negative Matrix Factorization [17–20], Independent Components Analysis [21, 22] and others [23, 24]. However, all of these methods search for a single set of endmembers and, therefore, a single convex region to describe a hyperspectral scene. Since these algorithms assume a single convex region, they cannot find appropriate endmembers for non-convex data sets.

1) Vertex Component Analysis: Vertex Component Analysis, VCA, provides an extremely fast endmember detection method [15]. This method employs the pixel purity assumption and selects data points from the input image as endmembers. VCA proceeds by iteratively projecting the data onto a direction orthogonal to the space spanned by the previously selected endmembers. The data point lying at the extreme of the projection is chosen as the next endmember. VCA relies on the fact that endmembers following the convex geometry model are found at the corners of a simplex encapsulating the input data.

2) Iterated Constrained Endmembers Algorithm: The Iterated Constrained Endmembers, ICE, algorithm does not rely on the pixel purity assumption. Instead, endmembers and abundances for a data set are found by iteratively updating the objective function shown in Equation 6.

$$ICE^{obj} = (1 - \mu) \left[\sum_{i=1}^{N} \left(\mathbf{x}_i - \sum_{k=1}^{M} p_{ik} \mathbf{e}_k \right)^T \left(\mathbf{x}_i - \sum_{k=1}^{M} p_{ik} \mathbf{e}_k \right) \right] + \mu \left[\sum_{k=1}^{M-1} \sum_{l=k+1}^{M} \left(\mathbf{e}_k - \mathbf{e}_l \right)^T \left(\mathbf{e}_k - \mathbf{e}_l \right) \right] \quad (6)$$

where $\mathbf{x}_i$ is the ith data point, p_{ik} is the abundance value of the kth endmember for the ith data point, $\mathbf{e}_k$ is the kth endmember, N is the number of data points, and M is the number of endmembers.

The first term of this objective function computes the error incurred by estimating the input data using the current set of endmembers and abundances. The second term of the objective function computes the sum of squared differences (SSD) between all of the current endmembers. The SSD term is related to volume enclosed by the endmembers and is used to find endmembers that provide a tight fit around the input data set. This objective is iteratively minimized by cycling through solving for abundances while holding endmembers constant and solving for endmembers while holding abundances constant. During the abundance updates, quadratic programming is used to solve for the abundance values while adhering to the constraints in Equation 2.

III. Piece-wise Convex Multiple Model Endmember Detection

Because of the flexibility of fuzzy clustering with respect to the distance used in the objective function, we can use a more general interpretation of the distance from a point to a cluster, in particular, the error between the pixel spectra and the pixel estimate found using the endmembers and their proportions. Even though the regression residual of a point relative to a convex geometry model framework is not a physical distance, it does represent the goodness of fit of a given model in the neighborhood of a point whenever the model can be described in explicit from. In this case, the algorithm tries to find a robust estimate of the model's endmembers and abundances from the distribution of the data points. However, unlike classical endmember finding methods which only estimate parameters of a single convex model , we propose a new approach that estimates the endmembers and abundances parameters of multiple models simultaneously. Assuming a convex geometry mixing model, this method simultaneously fits C models to data and automatically identifies the endmembers corresponding to each model. In other words, a set of endmembers $\{\mathbf{e}_i\}_{i=1,...,C}$ and their corresponding set of abundances $\{\mathbf{p}_i\}_{i=1,...,C}$ are identified for the set of points assigned to the ith model. Moreover, the P-COMMEND algorithm is designed to be robust and allow for convex regions to overlap by minimizing the following expression

$$J^{MM} = \sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}^{m} \left(\mathbf{x}_j - \mathbf{p}_{ij}\mathbf{E}_i\right)\left(\mathbf{x}_j - \mathbf{p}_{ij}\mathbf{E}_i\right)^{T} \tag{7}$$

subject to

$$p_{ijk} \geq 0, \tag{8}$$

$$\sum_{k=1}^{M} p_{ijk} = 1 \; and \tag{9}$$

$$\sum_{i=1}^{C} u_{ij} = 1 \tag{10}$$

In (7), $\mathbf{x}_j$ is a $1 \times d$ vector representing the jth pixel of the image, N is the number of pixels in the image, M is the number of endmembers, P_i is a $N \times M$ matrix such that $\mathbf{p}_{ij}$, is the vector of proportions associated with pixel j with respect to model i, p_{ijk} is the proportion of endmember k in pixel j with respect to model i, and $\mathbf{E}_i$ is a $M \times d$ matrix such that each row of $\mathbf{E}_i$, $\mathbf{e}_{ik}$, is a $1 \times d$ vector representing the kth endmember with respect to model i. Further, in what follows, we use the notation $1_{S \times T}$ to denote an $S \times T$ matrix with all entries equal to 1. In a sense, the multiple models compete for a given data points $\mathbf{X}$ through a set of weights u_{ij} which represent the memberships or degrees of sharing of data points $\mathbf{X}$ in the ith model.

As we need to constrain the size of the simplex in such a way that the endmembers surround tightly the model, we add the SSD term used in Iterated Constrained Endmembers approach (ICE) [25]. This term is defined as follows:

$$SSD = \sum_{k=1}^{M-1}\sum_{j=k+1}^{M} \left(\mathbf{e}_{ik} - \mathbf{e}_{ij}\right)\left(\mathbf{e}_{ik} - \mathbf{e}_{ij}\right)^{T} \tag{11}$$

The expression in (11) can be rewritten in matrix form as

$$SSD = M \cdot trace(\mathbf{E}_i\mathbf{E}_i^{T}) - 1_{1 \times M}\mathbf{E}_i\mathbf{E}_i^{T}1_{M \times 1} \tag{12}$$

In order to consider the constraints defined by (8), (9) and (10), we use the Lagrange multiplier method and the Karush-Kuhn-Tucker (KKT) Conditions [26]. Thus, considering the regularization term described by (12) and the constraints defined by (8), (9) and (10), we can formulate the endmember detection problem as the problem of minimizing the following objective function

$$\begin{aligned} J &= \sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}^{m} \left(\mathbf{x}_j - \mathbf{p}_{ij}\mathbf{E}_i\right)\left(\mathbf{x}_j - \mathbf{p}_{ij}\mathbf{E}_i\right)^{T} \\ &+ \alpha\sum_{i=1}^{C}\left(M \cdot trace(\mathbf{E}_i\mathbf{E}_i^{T}) - 1_{1 \times M}\mathbf{E}_i\mathbf{E}_i^{T}1_{M \times 1}\right) \\ &- \sum_{i=1}^{C}\sum_{j=1}^{N}\lambda_i\left(1_{1 \times M}\mathbf{p}_{ij}^{T} - 1\right) \\ &- \sum_{i=1}^{C}\sum_{j=1}^{N}\gamma_{ij}\mathbf{p}_{ij}^{T} \\ &- \sum_{j=1}^{N}\beta_j\left(\sum_{i=1}^{C}u_{ij} - 1\right) \end{aligned} \tag{13}$$

We minimize the objective function described by equation (13) with respect to $\mathbf{E}_i$, $\mathbf{P}_i$ and u_{ij}. In order to compute the optimal values of $\mathbf{E}_i$, $\mathbf{P}_i$ and u_{ij}, we use the well known alternating optimization method. In other words, for the proposed objective function in (13), we will alternate the optimization of $\mathbf{E}_i$, $\mathbf{P}_i$ and u_{ij}. Unlike the ICE approach [25] which uses iterative quadratic programming for the optimization of $\mathbf{P}_i$, in the following, we propose to compute an explicit solution for $\mathbf{E}_i$, $\mathbf{P}_i$ and u_{ij}.

The derivative of (13) with respect to $\mathbf{E}_i$ is

$$
\begin{aligned}
\frac{\partial J}{\partial \mathbf{E}_i} &= \sum_{j=1}^{N} -2u_{ij}^m \mathbf{p}_{ij}^T \mathbf{x}_j + 2u_{ij}^m \mathbf{p}_{ij}^T \mathbf{p}_{ij} \mathbf{E}_i \\
&+ 2\alpha \left(M I_{M \times M} - 1_{M \times M} \right) \mathbf{E}_i
\end{aligned}
\tag{14}
$$

Setting (14) to zero gives

$$
\mathbf{E}_i = \left(\sum_j u_{ij}^m \mathbf{p}_{ij}^T \mathbf{p}_{ij} + 2\alpha D \right)^{-1} \left(\sum_j u_{ij}^m \mathbf{p}_{ij}^T \mathbf{x}_j \right)
\tag{15}
$$

where

$$
D = M I_{M \times M} - 1_{M \times M}
\tag{16}
$$

The derivative of (13) with respect to $\mathbf{p}_{ij}^T$ is

$$
\begin{aligned}
\frac{\partial J}{\partial \mathbf{p}_{ij}^T} &= -2u_{ij}^m \mathbf{E}_i \mathbf{x}_j^T + 2u_{ij}^m \mathbf{E}_i \mathbf{E}_i^T \mathbf{p}_{ij}^T \\
&- \lambda_i 1_{M \times 1} - \boldsymbol{\gamma}_{ij}
\end{aligned}
\tag{17}
$$

Setting (17) to zero gives

$$
\mathbf{p}_{ij}^T = \left(\mathbf{E}_i \mathbf{E}_i^T \right)^{-1} \left(\mathbf{E}_i \mathbf{x}_j^T - \frac{\lambda_i}{2} 1_{M \times 1} - \boldsymbol{\gamma}_{ij} 1_{M \times 1} \right)
\tag{18}
$$

KKT conditions [26] also gives us the complementary slackness, $\boldsymbol{\gamma}_{ij} \mathbf{p}_{ij}^T = 0$ and the sign condition for the inequality constraints, $\boldsymbol{\gamma}_{ij} \geq 0$. But, if $\boldsymbol{\gamma}_{ij} > 0$, then $\mathbf{p}_{ij}^T = 0$. Thus, $\boldsymbol{\gamma}_{ij}$ should be equal to zero and $\mathbf{p}_{ij}^T > 0$. So,

$$
\mathbf{p}_{ij}^T = \left(\mathbf{E}_i \mathbf{E}_i^T \right)^{-1} \left(\mathbf{E}_i \mathbf{x}_j^T - \frac{\lambda_i}{2} 1_{M \times 1} \right)
\tag{19}
$$

and

$$
\mathbf{p}_{ij}^{KKT} = max \left(\mathbf{p}_{ij}^T, 0 \right)
\tag{20}
$$

Using the fact that

$$
1_{1 \times M} \mathbf{p}_{ij}^T = 1
\tag{21}
$$

and substituting (18) in (21), gives

$$
1_{1 \times M} \left(\mathbf{E}_i \mathbf{E}_i^T \right)^{-1} \mathbf{E}_i \mathbf{x}_j^T = 1 - \frac{\lambda_i}{2} ones \left(1, M \right) \left(\mathbf{E}_i \mathbf{E}_i^T \right)^{-1}
\tag{22}
$$

Solving (22) with respect to λ_i gives

$$
\lambda_i = 2 \frac{1_{1 \times M} \left(\mathbf{E}_i \mathbf{E}_i^T \right)^{-1} \mathbf{E}_i \mathbf{x}_j^T - 1}{1_{1 \times M} \left(\mathbf{E}_i \mathbf{E}_i^T \right)^{-1} 1_{M \times 1}}
\tag{23}
$$

The derivative of (13) with respect u_{ij} is

$$(m-1)\, u_{ij}^{m-1} \left((\mathbf{x}_j - \mathbf{p}_{ij}\mathbf{E}_i)(\mathbf{x}_j - \mathbf{p}_{ij}\mathbf{E}_i)^T \right) - \beta_j \tag{24}$$

Setting (24) to zero, gives

$$u_{ij} = \left(\frac{\beta_j}{(m-1)(\mathbf{x}_j - \mathbf{p}_{ij}\mathbf{E}_i)(\mathbf{x}_j - \mathbf{P}_{ij}\mathbf{E}_i)^T} \right)^{\frac{1}{m-1}} \tag{25}$$

Using the fact that

$$\sum_{i=1}^{C} u_{ij} = 1 \tag{26}$$

and substituting (25) in (26), gives

$$\sum_{i=1}^{C} \left(\frac{\beta_j}{(m-1)(\mathbf{x}_j - \mathbf{p}_{ij}\mathbf{E}_i)(\mathbf{x}_j - \mathbf{p}_{ij}\mathbf{E}_i)^T} \right)^{\frac{1}{m-1}} = 1 \tag{27}$$

Solving (27) with respect to β_j gives

$$\beta_j^{\frac{1}{m-1}} = \frac{1}{\sum_{i=1}^{C} \left(\frac{1}{(m-1)(\mathbf{x}_j - \mathbf{p}_{ij}\mathbf{E}_i)(\mathbf{x}_j - \mathbf{p}_{ij}\mathbf{E}_i)^T} \right)^{\frac{1}{m-1}}} \tag{28}$$

Substituting (28) in (25) gives

$$u_{ij} = \frac{\left(\frac{1}{(\mathbf{x}_j - \mathbf{p}_{ij}\mathbf{E}_i)(\mathbf{x}_j - \mathbf{p}_{ij}\mathbf{E}_i)^T} \right)^{\frac{1}{m-1}}}{\sum_{q=1}^{C} \left(\frac{1}{(\mathbf{x}_j - \mathbf{p}_{qj}\mathbf{E}_q)(\mathbf{x}_j - \mathbf{p}_{qj}\mathbf{E}_q)^T} \right)^{\frac{1}{m-1}}} \tag{29}$$

The proposed Multiple Model endmember finding algorithm is summarized below.

Algorithm 1 The Multiple Model endmember finding algorithm

Fix number of clusters C and m $\in [1\infty)$;

Initialize the fuzzy partition matrix U;

Initialize alpha;

Initialize the set of abundance matrices $\{\mathbf{P}_i\}_{i=1,\ldots,C}$.

   REPEAT

        Update endmembers matrices $\{\mathbf{E}_i\}_{i=1,\ldots,C}$ using (15);

        Update the set of abundances $\{\mathbf{P}_i\}_{i=1,\ldots,C}$ using (19) and (23);

        Apply the KKT condition on $\{\mathbf{P}_i\}_{i=1,\ldots,C}$ using (20) and renormalize;

        Update the fuzzy membership using (29);

   UNTIL convergence

IV. Experimental Results

A. Two-Dimensional Data

The P-COMMEND and ICE algorithms were run on a simulated two-dimensional data set that was generated from three sets of three endmembers. The data set is shown in Figure 3(a). The results for the P-COMMEND algorithm are shown in Figure 3(b). These results were generated using the following parameter settings: $\alpha = 0.001$, $M = 3$, $C = 3$, and $m = 2$. As shown in the figure, the P-COMMEND algorithm was able to find well-suited endmembers

for this piece-wise convex data set. The results using the ICE algorithm are shown in Figure 3(c). These results were generated using the parameters $\mu = 0.001$ and $M = 9$. Although the same number of endmembers were used as in the P-COMMEND algorithm, the endmembers do not accurately represent the piece-wise convex nature of this data set.
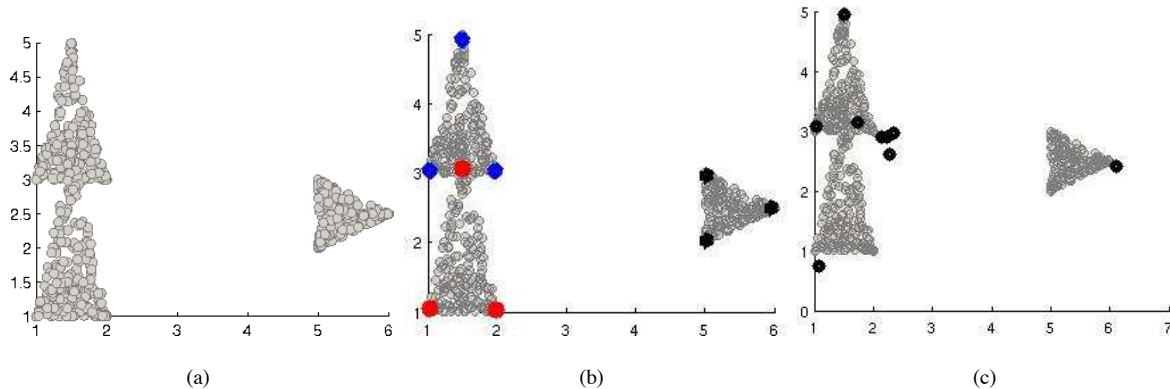

Fig. 3. (a) Simulated Two-Dimensional Piece-wise Convex data generated using three sets of three endmembers. (b) P-COMMEND endmember detection results on two-dimensional simulated data. The nine endmembers found are shown in bold. (c) ICE endmember detection results on two-dimensional simulated data. The nine endmembers found are shown in bold.

VCA endmember detection results could not be obtained for this data since the method requires the number of endmembers to be less than or equal to the dimensionality of the data set.

B. AVIRIS Indian Pines Data

The P-COMMEND and ICE endmember detection methods were applied to the June 1992 AVIRIS Indian Pines data set for comparison [4]. These data were collected over the Indian Pines test site in an agricultural area of northern Indiana. The image has 145×145 pixels with 220 spectral bands. The data contains approximately two-thirds agricultural land and one-third forest and other elements [27]. Figure 4 shows an image of band 10 (approximately 0.49 μm) and the ground truth for this data set.

Prior to applying the P-COMMEND and ICE algorithm, PCA was used to reduce the dimensionality of the data to three dimensions. Figure 1 shows a scatter plot of the data following this dimensionality reduction and illustrates that this data set is non-convex. The P-COMMEND algorithm was applied to this data using the following parameter set: $\alpha = 5 \times 10^{-4}$, $M = 3$, $C = 3$, and $m = 1.2$. A scatter plot displaying the endmembers that were estimated is shown in Figure 5(a). For comparison, the ICE algorithm was also run on this data with the following parameter set: $\mu = 0.001$ and $M = 9$. A scatter plot displaying the endmembers found by ICE is shown in Figure 5(b). The endmembers found by ICE enclose the data set as well as a large amount of empty space. In contrast, due to its piece-wise convex nature, the P-COMMEND algorithm encloses the data while reducing the amount of empty space enclosed by the endmembers. This is better illustrated in Figure 6. Figure 6 shows scatter plots of the data points partitioned according to their maximum membership, u, values and the corresponding endmembers for that partition. The endmembers for each partition provides a tight fit around the data points.
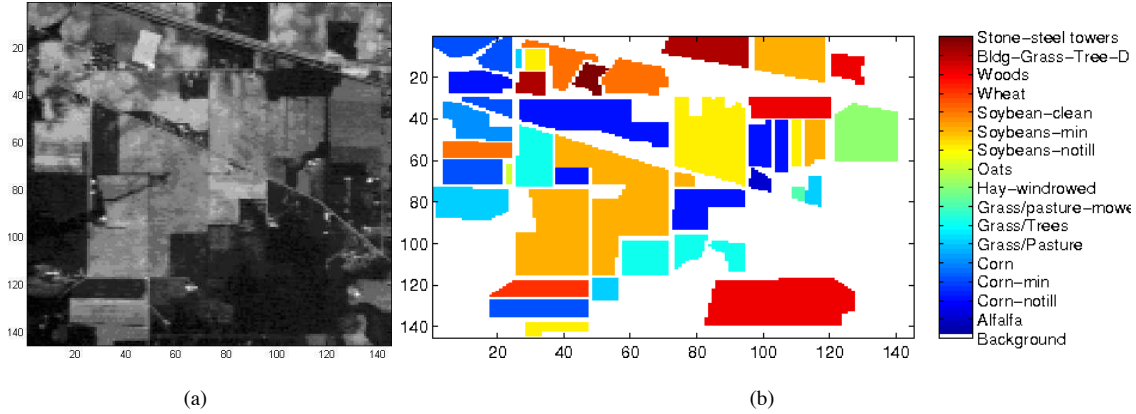
Fig. 4.   (a) Band 10 (0.49 μm) of the AVIRIS Indian Pines data set (b) The ground truth of the AVIRIS Indian Pines data set.
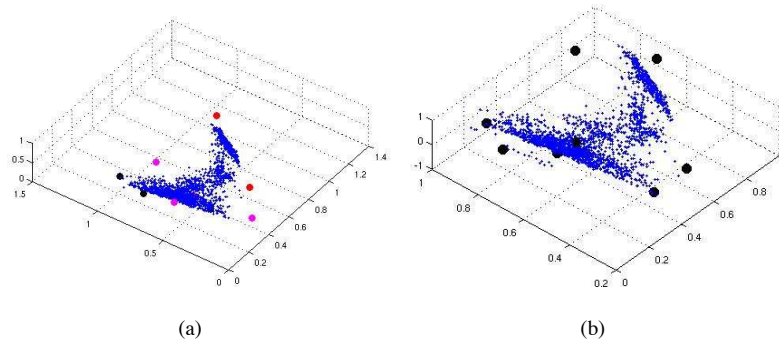

Fig. 5.   (a) Scatter plot of the AVIRIS Indian Pines data and the endmembers (in bold) found using the P-COMMEND algorithm. (b) Scatter plot of the AVIRIS Indian Pines data and the endmembers (in bold) found using the ICE algorithm.
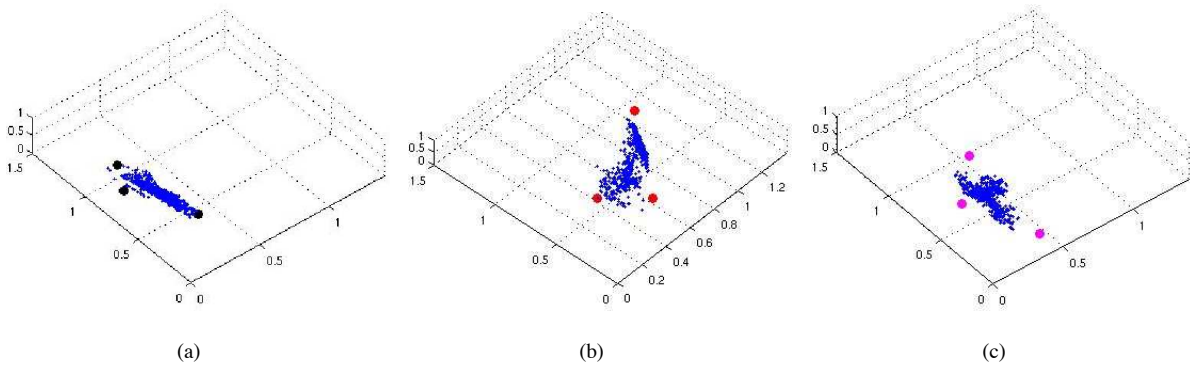

Fig. 6.   (a) Scatter plot of a subset of the AVIRIS Indian Pines data and the corresponding set of endmembers (in bold) found using the P-COMMEND algorithm. The data partitioned with the fuzzy clustering method used in the P-COMMEND algorithm. (b) Scatter plot of the second subset of the AVIRIS Indian Pines data and the corresponding set of endmembers (in bold) found using the P-COMMEND algorithm. (c) Scatter plot of a third subset of the AVIRIS Indian Pines data and the corresponding set of endmembers (in bold) found using the P-COMMEND algorithm. The data points were associated with the set of endmembers using the maximum membership value found by the P-COMMEND algorithm.
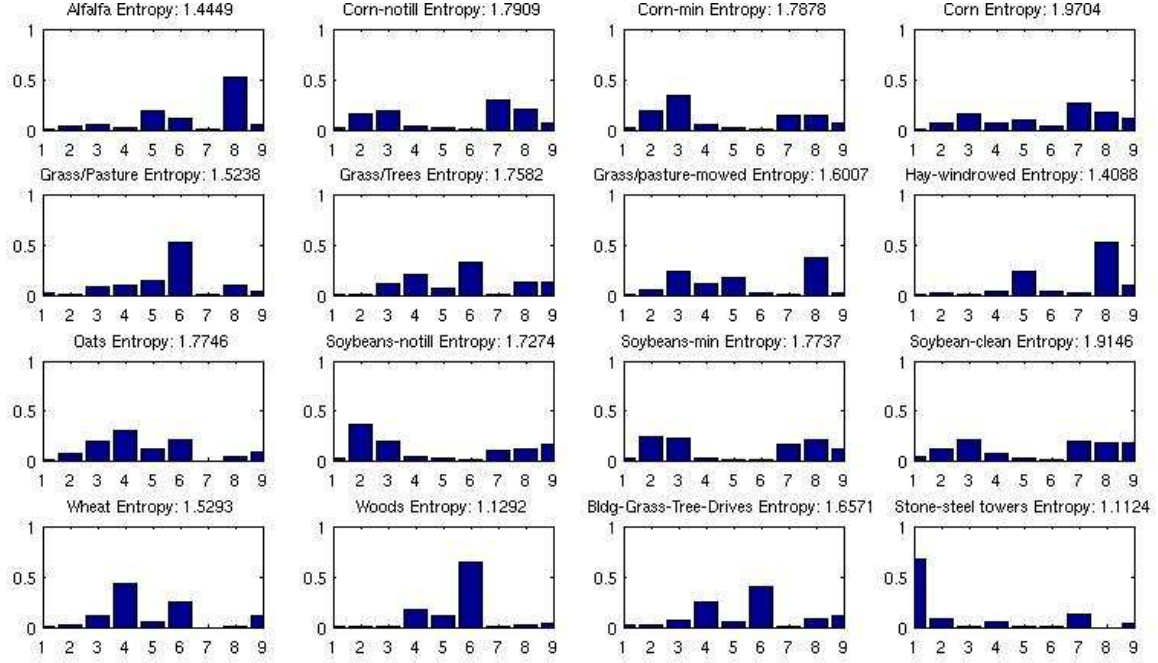
Fig. 7. The distribution of abundance values across the endmembers found using the ICE algorithm. The x-axis indicates endmember label and the y-axis are the normalized distribution of abundances associated with each ground truth label.

Figure 7 and Figure 8 further illustrate how P-COMMEND is able to generate endmembers that better suit the data. These figures plot the normalized distribution of abundance values across the endmembers for each class label. The abundance values found by P-COMMEND for each class are concentrated to a smaller number of endmembers than the abundance values found by ICE. For each class, Shannon's entropy of the normalized distribution of abundance values was computed. A smaller entropy value indicates that the abundances associated with a single class are concentrated onto a smaller number of endmembers. Table I lists the entropy values computed for each class using the abundances found by ICE and the abundance found by P-COMMEND. The P-COMMEND results are consistently smaller (except for the stone/steel towers) indicating that the abundances for each class are concentrated onto a smaller number of endmembers and, therefore, the endmembers provide a better representation of the input data. Consider the Grass/Pasture, Grass/Trees, Oats and Wheat classes. The abundances for all of these class are primarily concentrated to three endmembers using the P-COMMEND algorithm. In contrast, the ICE algorithm spreads the abundance values for these classes across 6, 6, 7, and 5 endmembers, respectively. These results can also be seen in the abundance maps for each endmember found by the P-COMMEND and ICE algorithms shown in Figures 9 and 10.

C. VIS/NIR Face Data

The P-COMMEND algorithm was also tested on VIS/NIR hyperspectral face imagery. This data was collected by Norsk Elektro Optikk using their HySpex hyperspectral sensor on participants at the 2009 IEEE Workshop
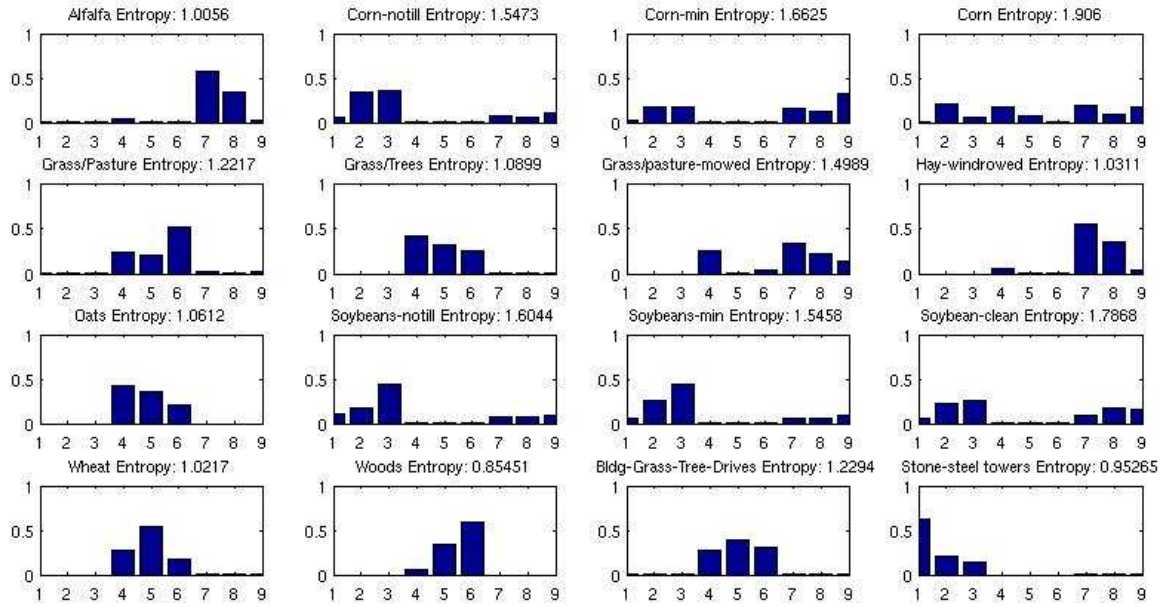
Fig. 8. The distribution of abundance values across the endmembers found using the P-COMMEND algorithm. The x-axis indicates endmember label and the y-axis are the normalized distribution of abundances associated with each ground truth label.

TABLE I

COMPARISON OF THE SHANNON ENTROPY VALUES FOR THE DISTRIBUTION OF ABUNDANCE VALUES ACROSS ENDMEMBERS ASSOCIATED WITH EACH GROUND TRUTH LABEL.

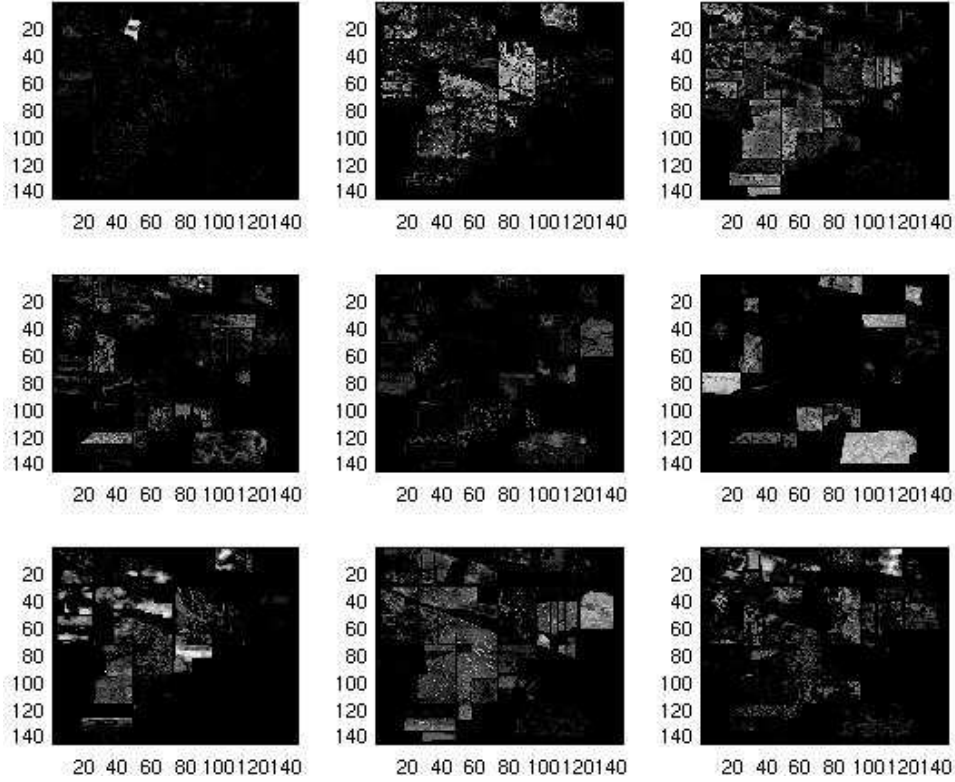|	P-COMMEND Entropies	ICE Entropies
Alfalfa	1.0	1.4
Corn-no till	1.5	1.8
Corn-min	1.7	1.8
Corn	1.9	2.0
Grass/Pasture	1.2	1.5
Grass/Trees	1.1	1.8
Grass/Pasture-Mowed	1.5	1.6
Hay-windrowed	1.0	1.4
Oats	1.1	1.8
Soybeans-notill	1.6	1.7
Soybeans-min	1.5	1.8
Soybeans-clean	1.8	1.9
Wheat	1.0	1.5
Woods	0.9	1.1
Bldg-Grass-Trees-Drives	1.2	1.7
Stone-steel Towers	1.1	1.0

Fig. 9. Abundance maps for the AVIRIS Indian Pines data associated with each of the nine endmembers found using the ICE algorithm. Colormapping is linear ranging from black (zero abundance) to white (abundance of one).

on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS09). The data contains 160 spectral bands ranging in wavelength from 410 to 990nm. Three images were selected from this data set and spatially sub-sampled by selecting every 20th row and column in the images. Band 100 (wavelength 772nm) for the three sub-sampled images selected are shown in Figure 11.

Using the PCA algorithm, the dimensionality of this data was reduced from 160 dimensions to 10 dimensions. The P-COMMEND algorithm was run on all three images using the following parameters: $\alpha = 0.001$, $M = 2$, $C = 3$, and $m = 2$. The weighted abundance maps associated with each endmember for the three images are shown in Figure 12. The abundances were weighted by multiplying each abundance value with the corresponding fuzzy membership value, u, for the cluster found using the P-COMMEND algorithm. The first three columns in Figure 12 correspond to the first set of endmembers and columns four through six correspond to the second set of endmembers found. The endmembers found using the P-COMMEND algorithm can be interpreted as follows: endmembers (a), (g), and (m) correspond to skin, endmembers (b), (h) and (n) correspond to shadows on the face and some hair, endmember (c) corresponds to teeth, endmembers (i) and (o) correspond to some shirt material, endmembers (d), (j) and (p) correspond to background, endmembers (e), (k) and (q) correspond to hair and some background, endmembers (f) and (l) correspond to shirt material and endmember (r) corresponds to some background scenery.
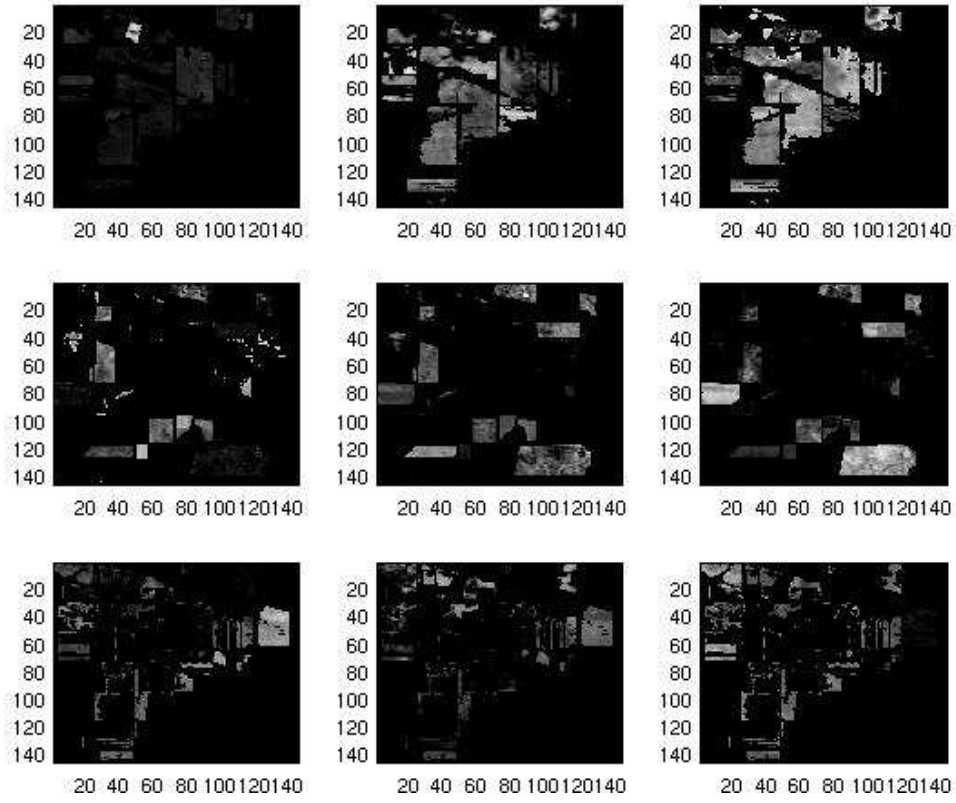
Fig. 10. Abundance maps for the AVIRIS Indian Pines data associated with each of the nine endmembers found using the P-COMMEND algorithm. Colormapping is linear ranging from black (zero abundance) to white (abundance of one).
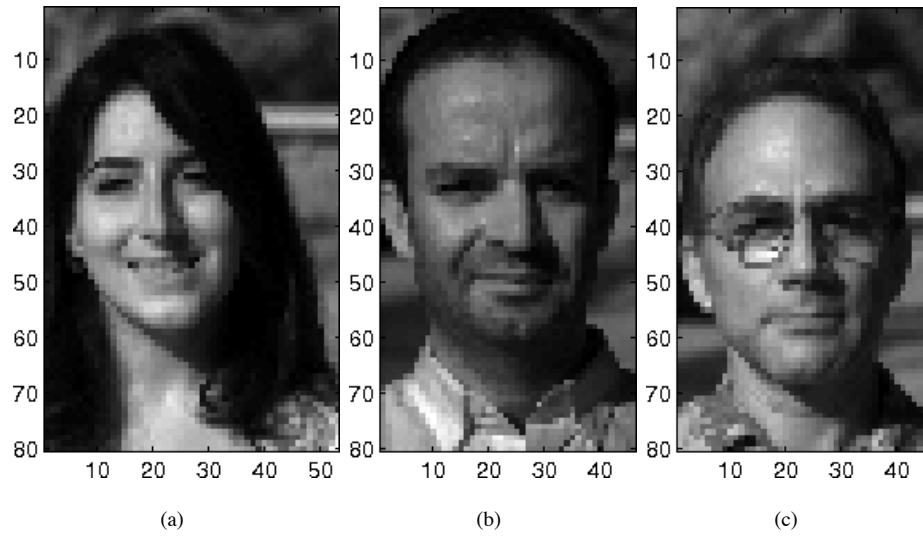

(a)   (b)   (c)

Fig. 11. Band 100 (wavelength 772nm) of three VIS/NIR hyperspectral face images after spatially sub-sampling the imagery by selecting every 20th row and column in the imagery.

Fig. 12. Abundance maps found using the P-COMMEND algorithm on the VIS/NIR hyperspectral face imagery. The first three columns correspond to endmembers from the first convex region. The last three columns correspond to endmembers from the second convex region. The same color mapping is used for all of the images ranging from zero abundance (black) to an abundance of one (white).

For comparison, the ICE and VCA algorithms were run on these images as well. The ICE algorithm was run with the following parameter settings: $\mu = 0.001$ and $M = 6$. The proportion maps from ICE are shown in Figure 13. The VCA algorithm was run to find 6 endmembers on these three images. Since the VCA method returns endmebmer spectra with out abundance values, the constrained unmixing method (using Equations 19, 23 and 20) employed by P-COMMEND was used to generate abundance maps for the images using the endmembers found from VCA. These abundance maps are shown in Figure 14.

Comparison of the P-COMMEND, ICE and VCA abundance maps on the face imagery indicate that the P-

Fig. 13. Abundance maps found using the ICE algorithm on the VIS/NIR hyperspectral face imagery. The same color mapping is used for all of the images ranging from zero abundance (black) to an abundance of one (white).

COMMEND algorithm provided endmembers that better distinguish between different portions of the imagery. Specifically, consider the second and fifth columns in the P-COMMEND results. The second column corresponds mainly to shadow whereas the fifth column corresponds mainly to hair and some background scenery. In contrast, both the ICE and VCA methods were not able to distinguish between hair and shadow. The P-COMMEND results also restricted any abundance associated with skin to the first convex region, or the first three endmembers. The overwhelming majority of the skin signature is associated with the first endmember of all of the images. In contrast, both ICE and VCA spread abundances associated with skin across many endmembers.
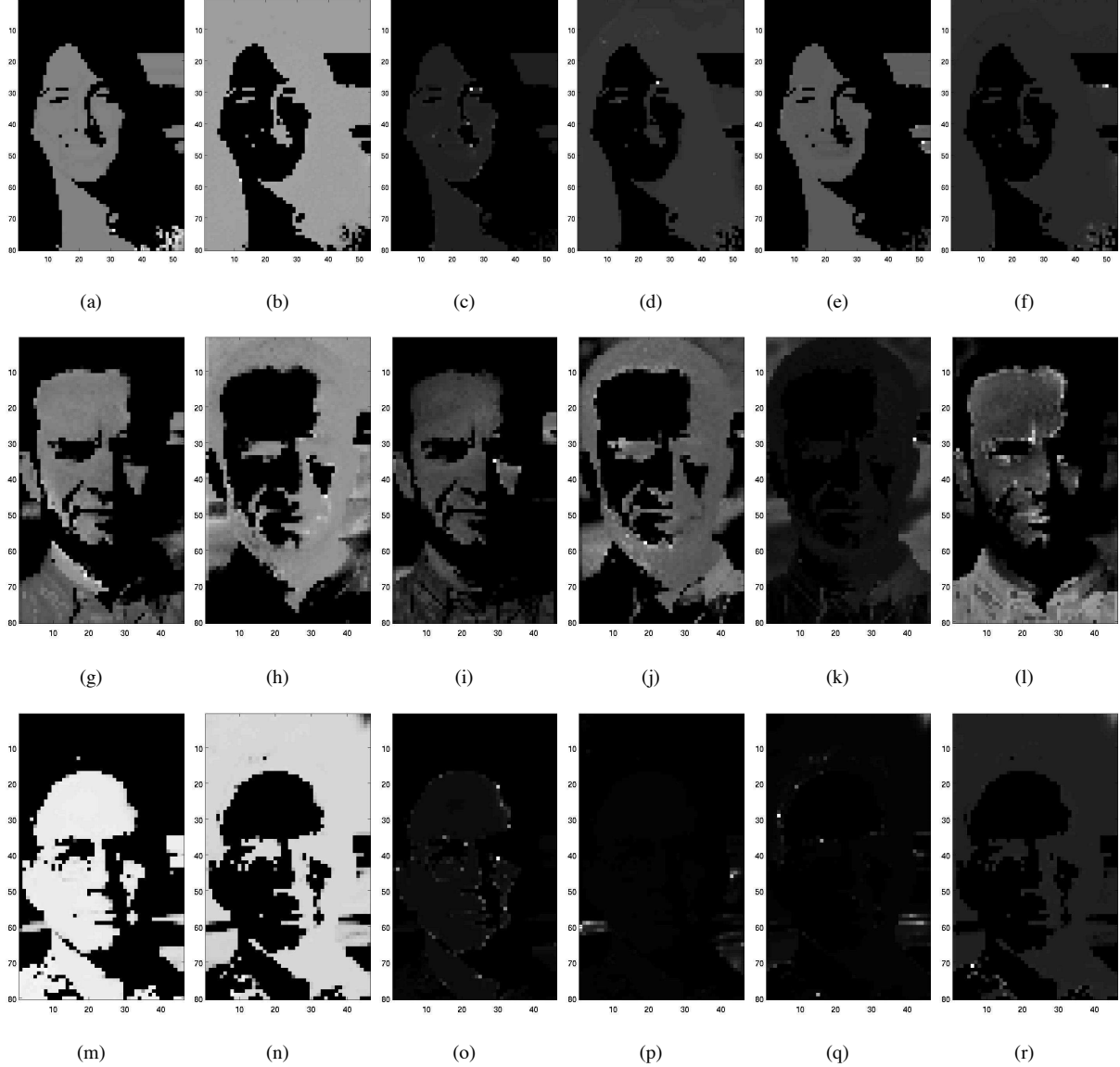
Fig. 14. Abundance maps found using the endmembers found by the VCA algorithm on the VIS/NIR hyperspectral face imagery. The same color mapping is used for all of the images ranging from zero abundance (black) to an abundance of one (white).

V. CONCLUSIONS AND FUTURE WORK

The P-COMMEND algorithm provides an endmember detection and spectral unmixing method that utilizes a piece-wise convex representation of hyperspectral imagery. The piece-wise convex representation allows several sets of endmembers to be determined. For each set of endmembers, the convex geometry model is applied and spectral unmixing is performed. As the results indicate, the piece-wise convex representation of hyperspectral imagery provides a better representation of the input data sets when compared to a single convex region.

Future work for this method include autonomously estimating the number of endmembers and the number of convex regions. Currently, the number of convex regions and endmembers need to be pre-determined. Furthermore,

methods to adaptively adjust the α regularization parameter during each iteration of the algorithm will be investigated as well.

REFERENCES

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

[2] J. M. P. Nascimento and J. M. Bioucas-Dias, "Does independent component analysis play a role in unmixing hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 1, pp. 175–187, Jan. 2005.

[3] D. Manolakis, D. Marden, and G. A. Shaw, "Hyperspectral image processing for automatic target detection applications," *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 79–116, 2003.

[4] (2004, Sep) AVIRIS free standard data products. Jet Propulsion Laboratory, California Institute of Technology, Pasedena, CA. URL http://aviris.jpl.nasa.gov/html/aviris.freedata.html.

[5] A. Zare, "Hyperspectral endmember detection and band selection using bayesian methods," Ph.D. dissertation, University of Florida, 2009.

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

[7] ——, "Context-based endmember detection for hyperspectral imagery," in *Proceedings of the First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, Grenoble, FR, Aug., pp. 1–4.

[8] R. G., B. Thomas, S. Tobgay, and T. S. Kumar, "Fuzzy clustering methods in data mining: A comparative case analysis," in *International Conference on Advanced Computer Theory and Engineering*, 2008, pp. 489–493.

[9] J. Bezdek, *Pattern Recognition with fuzzy objective function algorithm*. Plenum Press, 1981.

[10] J. Bezdek, C. Coray, R. Gunderson, and J. Watson, "Detection and characterization of cluster substructure: I. linear structure: Fuzzy clines," *SIAM J. Appl. Math*, vol. 40, pp. 339–357, 1981.

[11] H. Frigui and R. Krishnapuram, "A robust competitive clustering algorithm with applications in computer vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 450–465, May 1999.

[12] J. Nelder and R. Wedderburn, "Generalized linear models," *Journal of the Royal Statistical Society. Series A (General)*, vol. 135, no. 3, pp. 370–384, 1972.

[13] 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.

[14] J. Boardman, F. Kruse, and R. Green, "Mapping target signatures via partial unmixing of AVIRIS data," in *Summaries of the 5th Annu. JPL Airborne Geoscience Workshop*, R. Green, Ed., vol. 1. Pasadena, CA: JPL Publ., 1995, pp. 23–26.

[15] J. M. P. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix

hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898–910, Apr. 2005.

[16] A. Plaza, P. Martinez, R. Perez, and J. Plazas, "Spatial/spectral endmember extraction by multidimensional morphological operators," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 9, pp. 2025–2041, Sep. 2002.

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

[18] 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.

[19] V. P. Pauca, J. Piper, and R. J. Plemmons, "Nonnegative matrix factorization for spectral data analysis," *Linear Algebra Applications*, vol. 416, no. 1, pp. 321–331, Jul. 2005.

[20] S. Jia and Y. Qian, "Constrained nonnegative matrix factorization for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 1, pp. 161–173, Jan. 2009.

[21] 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.

[22] 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, Sep. 2006.

[23] M. D. Craig, "Minimum-volume transforms for remotely sensed data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 3, pp. 542–552, May 1994.

[24] G. X. Ritter, G. Urcid, and M. S. Schmalz, "Autonomous single-pass endmember approximation using lattice auto-associative memories," *Neurocomputing*, vol. 72, pp. 2101–2110, 2009.

[25] M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Donne, and J. F. Huntington, "ICE: A statistical approach to identifying endmembers in hyperspectral images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, pp. 2085–2095, Oct. 2004.

[26] H. W. Kuhn and A. W. Tucker, "Nonlinear programming," in *Proceedings of 2nd Berkeley Symposium*, 1951, pp. 481–492.

[27] 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.