

© 2006 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Wang, J., 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 44, no. 9 (21 August 2006): 2601–16. <https://doi.org/10.1109/TGRS.2006.874135>.

<https://doi.org/10.1109/TGRS.2006.874135>

Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

### **Please provide feedback**

Please support the ScholarWorks@UMBC repository by emailing [scholarworks-group@umbc.edu](mailto:scholarworks-group@umbc.edu) and telling us what having access to this work means to you and why it's important to you. Thank you.

# Applications of Independent Component Analysis (ICA) in Endmember Extraction and Abundance Quantification for Hyperspectral Imagery

Jing Wang, *Student Member, IEEE* and Chein-I Chang, *Senior Member, IEEE*

**Abstract**—Independent component analysis (ICA) has shown success in many applications. This paper investigates a new application of the ICA in endmember extraction and abundance quantification for hyperspectral imagery. An endmember is generally referred to as an idealized pure signature for a class whose presence is considered to be rare. When it occurs, it may not appear in large population. In this case, the commonly used principal components analysis (PCA) may not be effective since endmembers usually contribute very little in statistics to data variance. In order to substantiate our findings, an ICA-based approach, called ICA-based abundance quantification algorithm (ICA-AQA) is developed. Three novelties result from our proposed ICA-AQA. First, unlike the commonly used least squares abundance-constrained linear spectral mixture analysis (ACLSMA) which is a 2<sup>nd</sup> order statistics-based method, the ICA-AQA is a high order statistics-based technique. Second, due to the use of statistical independence it is generally thought that the ICA cannot be implemented as a constrained method. The ICA-AQA shows otherwise. Third, in order for the ACLSMA to perform abundance quantification, it requires an algorithm to find image endmembers first then followed by an abundance-constrained algorithm for quantification. As opposed to such a two-stage process, the ICA-AQA can accomplish endmember extraction and abundance quantification simultaneously in one-shot operation. Experimental results demonstrate that the ICA-AQA performs at least comparably to abundance-constrained methods.

**Index Terms**— Abundance-constrained linear spectral mixture analysis (ACLSMA). Abundance quantification. Endmember extraction. FastICA. High order statistics-based IC prioritization algorithm (HOS-ICPA). Independent components (ICs). IC prioritization. Independent component analysis (ICA). ICA-based endmember extraction algorithm (ICA-EEA). ICA-based abundance quantification algorithm (ICA-AQA). Initialization driven-based IC prioritization algorithm (ID-ICPA). Virtual dimensionality (VD).

## I. INTRODUCTION

Hyperspectral imaging has recently emerged as a very active area in remote sensing [1]. Many applications of

statistical signal processing techniques to hyperspectral imaging have been documented in [2]. Two major issues are of interest in remotely sensed imagery and are not encountered in classical two-dimensional (2D) image processing, which are subpixel targets and mixed pixels. A subpixel target is generally referred to a target with size smaller than a pixel resolution and therefore, is embedded in a single pixel. Accordingly, a subpixel target cannot be identified spatially by visual inspection. A mixed pixel occurs due to high spectral resolution or low spatial resolution when more than one material substance present in a pixel, in which case these substances are considered to be mixed linearly or nonlinearly in the pixel. Because of that, traditional pure pixel-based image processing techniques are generally not applicable to mixed pixel analysis. Unfortunately, due to the use of high spectral resolution by hyperspectral imaging sensors, a hyperspectral image pixel is most likely to be mixed by several substances or contain a subpixel target. As a result, the presence of pure pixels is rather rare, but provides critical information in data analysis, such as rare minerals in geology, rare species in agriculture and ecology, toxic waste in environmental monitoring, drug trafficking in law enforcement, vehicles in combat field, etc., all of which appear either at subpixel scale or in a mixed pixel rather than a pure pixel. Consequently, extracting pure pixels from a hyperspectral image presents a great challenge and has become increasingly important in hyperspectral data exploitation. However, there is a distinction between pure pixel and pure signature. Based on the definition given in [3], an endmember is an idealized, pure signature for a class. So, an endmember is a pure signature, but is not necessary to be a pure pixel. This is because a pure signature may come from a data base or spectral library which has been calibrated. So, it does not have to appear as a pixel in real data. When it does, it is called pure pixel. This implies that it may have many pure pixels which are specified by the same pure signature. Since the term “endmember” has been widely used in the remote sensing community, it will be used throughout this paper to indicate pure signature supposedly extracted by an endmember extraction algorithm (EEA). When an endmember extracted by an EEA as a pixel, it is referred to as endmember pixel.

As noted, an endmember represents a class, that is, finding endmembers is equivalent to finding different classes. Many endmember extraction algorithms (EEAs) have been developed for this purpose, for example, pixel purity index

Manuscript received October 19, 2005, revised December 13, 2005 and February 7, 2006.

J. Wang and C.-I Chang are with the Remote Sensing Signal and Image Processing Laboratory, Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, Baltimore, MD 21250 USA (e-mail: jingw1@umbc.edu, cchang@umbc.edu).

(PPI) [4], N-finder (N-FINDR) algorithm [5], iterative error analysis (IEA) [6], automated morphological endmember extraction (AMEE) algorithm [7], minimum volume transform [8], convex geometry [9], convex cone analysis [10], vertex component analysis (VCA) [11] etc. All these approaches share one thing in common which is the use of convexity to determine endmembers. However, they all also suffer from one common drawback which is no guideline suggested to determine how many endmembers required to be generated.

This paper presents a rather different approach, which uses statistical independence as a criterion as opposed to the convexity on which most EEAs are based. Our rationale is based on several observations. First of all, since distinct endmembers represent different classes, their statistical dependency must be least correlated. Second, due to the nature of endmembers as pure signatures the probability of the occurrence of endmembers is generally low. Third, when endmembers do occur, the spatial extent of their presence is usually very small. All these profiles seem to fit the strengths of the independent component analysis (ICA) [12] which uses mutual information to measure statistical independency. It is also well known that one of major applications for the ICA is blind source separation, which demixes a linear mixture of signal sources. If the signal sources to be demixed are interpreted as sources of pure signatures, its applicability to endmember extraction seems natural and justifiable.

Applications of the ICA to solving linear mixing problems for remotely sensed imagery are not new, just name a few [13-16]. However, most ICA-based approaches have mainly focused on separation of signal sources, but not on abundance quantification of signal sources. A second but rather important issue is determination of number of signal sources to be separated. To the authors' best knowledge, this issue has been left open in the ICA and has not been really addressed in the past. Recently, the issue was investigated in [17] where a new concept introduced in [2,18], called virtual dimensionality (VD) to estimate the number of independent components (ICs) required to be generated for dimensionality reduction. Interestingly, a similar idea of using the VD to estimate number of endmembers was first explored in [19-21] for PPI [4] and N-FINDR [5] and VCA in [21]. The work in [17] was the first to use the VD to perform the ICA-based dimensionality reduction. Unlike the principal components analysis (PCA) which ranks its principal components in accordance with the magnitude of variances or eigenvalues, the ICA does not prioritize its generated ICs as does the PCA. More specifically, the ICs generated earlier are not necessarily more important or significant than those produced later on. This dilemma is caused by the fact that there is no criterion implemented to prioritize components in the way that the PCA uses the eigenvalues to prioritize its components. Because the ICA generally uses random vectors as initial projection vectors to initialize the algorithm, the order that ICs are generated is completely random.

This issue is critical to endmember extraction because some ICs in which endmembers are present may be generated very last. In this case, all the ICs must be selected for endmember extraction in order to avoid missing of any possible endmember. Unfortunately, in hyperspectral image analysis,

there are generally hundreds of spectral bands, each of which corresponds to a component image, but only tens of components may contain useful and important information such as endmembers. Under this circumstance, selecting appropriate ICs for data analysis becomes crucial. To cope with this problem, two algorithms developed in [17] are used for this purpose. One is called high order statistics-based IC prioritization algorithm (HOS-ICPA) which uses high order statistics to prioritize ICs. The other is called initialization driven-based IC prioritization algorithm (ID-ICPA) which implements a custom-designed initialization algorithm to produce an appropriate set of initial projection vectors in sequence so that the ICs can be generated by the ICA in the same order that the initial projection vectors are generated. With the use of such an initialization algorithm the random generation of ICs can be avoided. Unfortunately, since the main theme in [17] was focused on dimensionality reduction and endmember extraction was only used to demonstrate one of its potential applications, a detailed study and analysis on endmember extraction was not conducted in [17]. In order for the ICA to perform abundance quantification, the ICA must satisfy two constraints, abundance sum-to-one constraint (ASC) and abundance non-negativity constraint (ANC). However, it is known that due to the use of statistical independence as its source separation criterion the ICA-generated results are scale invariant and may be also not necessarily positive, both of which violate the ASC and ANC. This is due to the fact that if the ICA is implemented as an abundance-constrained technique; otherwise, the endmembers will not be statistically independent. Because of that, using the ICA as an abundance quantifier is never thought to be possible. This paper takes up this challenging issue which was never addressed and investigated in the literature including [17] and shows otherwise.

In general, abundance quantification cannot be accomplished by any EEA including those in [4-11]. This is because an EEA only identifies endmembers which are assumed to be pure. As a result, there is no need to estimate abundance ratios of the extracted endmembers. In order for an EEA also to achieve abundance quantification of all pixels, an abundance quantification algorithm such as fully constrained least squares linear mixture analysis (LSMA) [2,20] must be implemented after endmembers are extracted. Compared to EEAs, our proposed ICA-based abundance quantification approaches can perform endmember extraction and abundance quantification simultaneously with no need of any additional abundance quantification algorithm. Its ability in abundance quantification can be derived from the same ICs selected for dimensionality reduction. Since each of endmembers represents a distinct class, all endmembers must be extracted in a separate and individual IC for discrimination and classification. Such a significant benefit cannot be gained by any EEA or abundance constrained LSMA. As a matter of fact, the use of the selected ICA-generated ICs to perform abundance quantification is similar to that of the LSMA to generate abundance fraction maps for each of image endmembers used in linear mixture model [2,22-23] except that the latter imposes constraints on abundance fractions.

In order to substantiate the utility of our proposed ICA-based approach in endmember extraction and abundance quantification, a set of synthetic images-based computer simulations is conducted to simulate various scenarios for quantitative analysis. Additionally, real hyperspectral image experiments are also performed to validate our proposed ICA-based approach in real applications.

The remainder of this paper is organized as follows. Section II briefly reviews the ICA and its application to linear spectral mixing problems in hyperspectral imagery. Section III develops two algorithms to prioritize ICs generated by the ICA. Section IV presents ICA-based algorithms for endmember extraction and abundance quantification. Section V and Section VI conduct computer simulations and real hyperspectral image experiments respectively to demonstrate the utility of our proposed ICA-based approach. Section VII concludes some remarks. Finally, an appendix on VD is included for completeness for those who are not familiar with and would like to learn more about the VD concept.

## II. INDEPENDENT COMPONENT ANALYSIS

Independent component analysis (ICA) has been widely used in blind source separation. Its application to LSMA in remote sensing image processing has been shown to be promising [13-16]. In what follows, we briefly describe its concepts with application to the LSMA.

Suppose that  $L$  is the number of spectral bands and  $\mathbf{r}$  is an  $L$ -dimensional image pixel vector. Assume that there are  $p$  targets,  $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_p$  present in an image scene. Let  $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_p$  denote their corresponding target signatures, which are generally referred to as digital numbers (DN). A linear mixture model of  $\mathbf{r}$  models the spectral signature of  $\mathbf{r}$  as a linear combination of  $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_p$  with appropriate abundance fractions specified by  $\alpha_1, \alpha_2, \dots, \alpha_p$ . More precisely,  $\mathbf{r}$  is an  $L \times 1$  column vector and  $\mathbf{M}$  be an  $L \times p$  target spectral signature matrix, denoted by  $[\mathbf{m}_1 \mathbf{m}_2 \dots \mathbf{m}_p]$ , where  $\mathbf{m}_j$  is an  $L \times 1$  column vector represented by the spectral signature of the  $j$ -th target  $\mathbf{t}_j$  resident in the pixel vector  $\mathbf{r}$ . Let  $\boldsymbol{\alpha} = [\alpha_1 \alpha_2 \dots \alpha_p]^T$  be a  $p \times 1$  abundance column vector associated with  $\mathbf{r}$  where  $\alpha_j$  denotes the fraction of the  $j$ -th target signature  $\mathbf{m}_j$  present in the pixel vector  $\mathbf{r}$ . A classical approach to solving a mixed pixel classification problem is linear unmixing which assumes that the spectral signature of the pixel vector  $\mathbf{r}$  is linearly mixed by  $\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_p$ , the spectral signatures of the  $p$  targets,  $\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_p$  as follows.

$$\mathbf{r} = \mathbf{M}\boldsymbol{\alpha} + \mathbf{n} \quad (1)$$

where  $\mathbf{n}$  is noise or can be interpreted as a measurement or model error.

In order for the ICA to perform the mixed pixel classification specified by (1), we reinterpret the target signature matrix and abundance vector  $\boldsymbol{\alpha} = [\alpha_1 \alpha_2 \dots \alpha_p]^T$  as mixing matrix and  $p$  abundance signal sources, denoted by

$\alpha_1, \alpha_2, \dots, \alpha_p$  respectively. There are two major differences of such interpretation from the linear mixture model used in the LSMA. One is that the  $p$  abundance fractions  $\alpha_1, \alpha_2, \dots, \alpha_p$ , which are commonly assumed to be unknown and non-random constants in the LSMA, are now assumed to be random parameters and statistically independent signal sources. Another is that at most one of these  $p$  signal sources is allowed to be Gaussian. A linear spectral mixture analysis implemented with these two assumptions was proposed in [14] and referred to as linear spectral random mixture analysis (LSRMA).

When the ICA is implemented, it generally assumes that the mixing matrix is a square matrix. When the number of rows of a mixing matrix,  $L$  is less than the number of sources,  $p$  to be separated, (1) deals with an under-determined system where the ICA is referred to as over-complete ICA. On the other hand, if  $L > p$ , (1) represents an over-determined system. In this case, the ICA is referred to as under-complete ICA and generally requires a dimensionality reduction to produce a square unmixing matrix. A common practice is to apply a technique such as PCA to reduce data dimensionality from  $L$  to  $p$ . However, such an approach is generally not effective for hyperspectral imagery because the subtle information such as subpixel, anomalies may be lost in the course of dimensionality reduction. So, as an alternative, this paper takes another approach. Instead of reducing data dimensionality by any means, it prioritizes ICs in accordance with a criterion determined by a specific application. The idea of IC prioritization is new and was reported in [17]. In this paper, it will be briefly reviewed in Section III.

Since we can interpret the  $L$  and  $p$  in a hyperspectral image as the number of bands and the number of abundance sources  $\alpha_1, \alpha_2, \dots, \alpha_p$  respectively with the assumption that  $L > p$  is always true, the ICA has  $p$  under-complete bases that are used to uncover the  $p$  abundance signal sources,  $\alpha_1, \alpha_2, \dots, \alpha_p$ . In this paper, the FastICA algorithm developed by Hyvarinen and Oja [24] will be used to find  $p$  ICs that correspond to the  $p$  abundance signal sources,  $\alpha_1, \alpha_2, \dots, \alpha_p$ . In doing so, each spectral band image, is converted to a vector. More specifically, assume that a hyperspectral image cube has size of  $M \times N \times L$  pixels where  $L$  is the number of spectral bands and  $MN$  is the size of each spectral band image. The hyperspectral image cube can then be represented by a data matrix  $\mathbf{X}$  of size  $L \times MN$  with  $L$  rows and  $MN$  columns. In other words, each row in the data matrix  $\mathbf{X}$  is specified by a particular spectral band image. As a result, a total of  $L$  ICs can be generated by the FastICA. However, according to (1) only  $p$  ICs are of interest and the ICA must be implemented with under-complete bases. In this case, the deflation approach used in FastICA is applied to generate ICs one by one sequentially where each of ICs is produced by maximizing the negentropy measured by kurtosis. When each IC is generated by the FastICA, an initial projection vector is required for initialization where a general approach is to randomly generate a vector and use it as its initial projection vector. As



a consequence, the ICs generated by the FastICA in different runs will not be necessary the same. In particular, an IC generated earlier does not necessarily imply that it is more important than the one generated later on. This dilemma will be resolved by introducing IC prioritization algorithms developed in Section III that can prioritize ICs by certain criteria.

### III. ALGORITHMS FOR PRIORITIZING ICs

Since the number of spectral bands  $L$  in hyperspectral imagery is generally much greater than the number of endmembers,  $p$ , the ICA in this case is actually under-complete ICA where the number of mixtures (i.e., observations) is much larger than the number of signal sources. As a result, how to find those ICs which contain endmembers is a key issue in endmember extraction. Furthermore, it should be noted that the FastICA-generated ICs are not necessarily arranged in order of information significance as the way that the principal components (PCs) are generated by the PCA or the MNF in accordance with decreasing magnitude of variances, eigenvalues or SNRs. Also the ICs generated by the FastICA in different runs generally appear in different orders. These issues are primarily due to the use of the initial projection unit vectors that are randomly generated by the Fast ICA to produce ICs. This implies that if the FastICA is repeated again, its produced ICs will appear in different orders. Under such circumstance, an endmember which appears in an early IC generated by one run may actually appear in a later IC generated by another run.

In order to resolve the issue resulting from the nature of the random initial projection unit vectors, a concept of IC prioritization is proposed. It determines a priority score for each of the ICA-generated ICs according to its significance measured by a certain criterion. These priority scores are then used to rank all the ICs in order of their priorities. Two IC prioritization algorithms, called High Order Statistics-Based IC prioritization algorithm (HOS-ICPA) and Initialization Driven-based IC prioritization algorithm (ID-ICPA) are developed for this purpose. In the HOS-ICPA, we consider each generated IC as a random variable. In this case, we assume that the  $i$ -th  $IC_i$  can be described by a random variable  $\zeta_i$  with values specified by the gray level value of the  $n$ -th pixel in the  $IC_i$ , denoted by  $z_n^i$ . In light of this interpretation, the FastICA-generated ICs can be ranked and prioritized by priority scores measured by high-order statistics-based criteria derived from the random variable  $\zeta_i$ . By contrast, the ID-ICPA takes a complete opposite approach. In order to remove the random nature caused by the initial projection unit vectors used by the FastICA, the ID-ICPA implements a particularly-designed initialization algorithm to produce an appropriate set of initial projection vectors in sequence for the FastICA. Consequently, the priority scores for the ICs generated by the ID-ICPA are always ranked by the sequential order that the initial projection vectors are generated.

#### A. High Order Statistics-Based IC Prioritization (HOS-ICPA)

The idea of the HOS-ICPA is to first determine the number of ICs needed to be retained,  $p$  which can be estimated by the VD. It then prioritizes the FastICA-generated ICs according to a high-order statistics criterion to select the first  $p$  prioritized ICs.

*High Order Statistics-Based IC-Prioritized Algorithm (HOS-ICPA)*

- 1 Find VD to determine the number of ICs required for FastICA to generate,  $p$ .
- 2 Implement the FastICA to produce all ICs where each  $IC_i$  uses a randomly generated unit vector as an initial projection vector to produce the final desired projection vector for the  $IC_i$ .
- 3 For each  $IC_i$  calculate its priority score based on the following high-order statistics that combines 3<sup>rd</sup> and 4<sup>th</sup> orders of statistics for  $\zeta_i$ .

$$priority\_score(IC_i) = (1/12)[\kappa_i^3]^2 + (1/48)[\kappa_i^4 - 3]^2 \quad (2)$$

where  $\kappa_i^3 = E[\zeta_i^3] = \frac{1}{MN} \sum_{n=1}^{MN} (z_n^i)^3$  and

$\kappa_i^4 = E[\zeta_i^4] = \frac{1}{MN} \sum_{n=1}^{MN} (z_n^i)^4$  are sample means of 3<sup>rd</sup> and 4<sup>th</sup> orders of statistics in the  $IC_i$ .

- 4 Prioritize the  $\{IC_i\}$  in accordance with the  $priority\_score(IC_i)$ .
- 5 Select the  $p$  ICs with  $p$  highest priority scores.

It should be noted that in (2) the priority score for each FastICA-generated IC is determined by only the third and fourth order statistics. Any statistics higher than four can be also used to calculate priority scores. However, based on our experience and experiments, such higher statistics do not offer significant advantages over the one used in (2).

#### B. Initialization Driven-Based IC Prioritization (ID-ICPA)

Since the initial projection unit vector used for each FastICA-generated IC is randomly generated, the ICs produced by the FastICA in different runs generally appear in different orders and also are not necessarily the same. The HOS-ICPA resolves this issue by implementing the criterion specified by (2) to calculate priority score for each of the FastICA-generated ICs that can be used to prioritize all the ICs. In this subsection, we present an alternative algorithm, called initialization driven-based IC prioritization algorithm (ID-ICPA) which does not calculate priority scores the way that the HOS-ICPA does for IC prioritization. Instead, the ID-ICPA resolves the issue of randomly generated ICs by using a custom-designed initialization algorithm which produces an appropriate set of initial projection vectors for the FastICA so that the FastICA-generated ICs are prioritized by the same order that the initial projection vectors are generated.

*Initialization Driven-Based IC-Prioritized Algorithm (ID-ICPA)*

- 1 Find VD to determine the number of ICs required for FastICA to generate,  $p$ .

- 2 Custom-design an initialization algorithm to produce an appropriate set of initial projection vectors  $\{\mathbf{t}_i\}_{i=1}^p$  to be used by the FastICA.
- 3 For each  $\text{IC}_i$ , the FastICA uses  $\mathbf{t}_i$  as its initial projection vector to produce the final desired projection vector for the  $\text{IC}_i$  and its priority score is then assigned to be “ $i$ ”, i.e.,  $\text{priority\_score}(\text{IC}_i) = i$ .

Three comments on ID-ICPA are noteworthy.

- (a) It should be noted that the IC prioritization is ranked by priority scores that are determined by the sequential order that the initial projection vectors  $\{\mathbf{t}_i\}_{i=1}^p$  are generated by a specific initialization algorithm. So, a different initialization algorithm generally produces a different set of initial projection vectors  $\{\mathbf{t}_i\}_{i=1}^p$ . As a consequence, the priority scores produced for the FastICA-generated ICs will be also different. In this paper, an automatic target generation process (ATGP) used in an automatic target detection and classification algorithm (ATDCA) developed by Ren and Chang [25] is selected to implement the ID-ICPA.
- (b) It is also worth noting that all the initial projection vectors  $\{\mathbf{t}_i\}_{i=1}^p$  are present in separate ICs and no two initial projection vectors are the same IC.
- (c) Unlike the HOS-ICPA which requires the FastICA to generate all ICs for prioritization, the ID-ICPA only needs the FastICA to generate  $p$  ICs.
- (d) An intuitive selection of initial projection vectors for  $\{\mathbf{t}_i\}_{i=1}^p$  seems to be eigenvectors. Unfortunately, according to our experiments the eigenvector-selected initial projection vectors do not always work and often result in inconsistent results [26]. This may be due to the fact that eigenvectors are results of second-order statistics and cannot serve well for the FastICA which is developed based on statistical independence.

#### IV. ICA-BASED ENDMEMBER EXTRACTION AND ABUNDANCE QUANTIFICATION

Many algorithms which have been developed for endmember extraction cannot be also used to perform abundance quantification. For such an endmember extraction algorithm to do so, it must be implemented in conjunction with an abundance quantification algorithm, for example a fully abundance-constrained method, fully constrained least squares (FCLS) in [2,23]. This section presents an ICA-based approach, called ICA-based abundance quantification algorithm (ICA-AQA) which can achieve both endmember extraction and abundance quantification simultaneously. The idea is to select the first  $p$  prioritized IC for endmember extraction and further use the same selected  $p$  prioritized ICs for abundance quantification for all image pixels. In other words, since each IC represents a specific class of targets extracted from the image data, it can be used to serve as abundance fraction map for this particular class.

##### A. Endmember Extraction

Most EEAs extract all desired endmembers in a single map to show their spatial presence in the image in a similar way that a classification algorithm produces a class map. The following proposed ICA-based endmember extraction algorithm performs otherwise. It takes advantages of the FastICA-generated ICs that separate all extracted endmember pixels in individual components so that no two endmember pixels that represent two different classes will be present in the same component.

##### ICA-Based Endmember Extraction Algorithm (ICA-EEA)

1. Find VD to determine the number of ICs needed to be generated,  $p$ .
2. Implement either HOS-ICPA or ID-ICPA to find first  $p$  prioritized ICs in accordance with their priority scores.
3. For each of the selected  $p$  FastICA-generated IC images, find a pixel with maximum absolute value, which is referred to as endmember pixel. The spectral signature of such found pixel is then selected as an endmember.
4. The  $p$  endmember pixels produced in step 3 denoted by  $\{\mathbf{e}_i\}_{i=1}^p$  are the final set of endmember pixels whose spectral signatures are our desirable endmembers.

It should be noted that there may have more than one endmember pixel extracted in one IC, in which all the extracted endmember pixels are considered to be in the same class as will be shown in our experiments.

##### B. Abundance Quantification

The main goal of the ICA-EEA is to extract all the desired endmembers from a hyperspectral image from a set of  $p$  appropriately prioritized ICs and further show each of these extracted endmember pixels in a separate and individual component. Since the ICA does not impose constrain on abundance fractions (otherwise the signal sources will not be statistically independent), using the ICA to perform abundance quantification seems out of its reach. Interestingly, an approach can be developed in the following to use the same  $p$  ICs obtained by the ICA-EEA for abundance quantification in a similar manner that the LSMA generates one abundance fraction map for each of its image endmembers.

##### ICA-Based Abundance Quantification Algorithm (ICA-AQA)

1. Implement ICA-EEA using either the HOS-ICPA or the ID-ICPA to find  $p$  endmember pixels,  $\{\mathbf{e}_i\}_{i=1}^p$ .
2. For each endmember pixel  $\mathbf{e}_i$ , let  $\text{IC}_i$  be the IC from which  $\mathbf{e}_i$  was extracted and  $\text{IC}_i(\mathbf{r})$  denote the value of each pixel  $\mathbf{r}$  in  $\text{IC}_i$ . We normalize the absolute value of  $\text{IC}_i(\mathbf{r})$ ,  $|\text{IC}_i(\mathbf{r})|$  with respect to  $|\mathbf{e}_i|$ , the absolute value of  $\mathbf{e}_i$  and define its corresponding abundance fraction  $\alpha_{\text{IC}_i}(\mathbf{r})$  by

$$\alpha_{\text{IC}_i}(\mathbf{r}) = \frac{|\text{IC}_i(\mathbf{r})| - \min_{\mathbf{r}} |\text{IC}_i(\mathbf{r})|}{|\mathbf{e}_i| - \min_{\mathbf{r}} |\text{IC}_i(\mathbf{r})|} \quad (3)$$

It is worth noting that the endmember  $\mathbf{e}_i$  in (3) is actually  $|\mathbf{e}_i| = \max_{\mathbf{r}} |\text{IC}_i(\mathbf{r})|$  which is the maximum of  $|\text{IC}_i(\mathbf{r})|$  over all the image pixels in the  $\text{IC}_i$ .

3. The set of  $\{\alpha_{IC_i}(\mathbf{r})\}_{\mathbf{r} \in IC_i}$  is the desired abundance fraction

map of the  $i$ -th independent component,  $IC_i$ .

A comment on (3) in step 2 is noteworthy. In order for the ICA\_AQA to perform quantification, (3) is implemented to satisfy both ASC and ANC by normalization. However, this is a heuristic approach to make an unconstrained method a fully constrained method. It does not really perform a fully constrained method as it should be such as FCLS. This is because the normalization performed in (3) varies with different ICs.

It should be also noted that the ICA-EEA and ICA-AQA proposed in this section make an assumption that the image data to be processed contain endmembers. This is a reasonable assumption since image data generally have a number of distinct classes and each class is assumed to be specified by an endmember. If there is no endmember in image data, the ICA-EEA extracts pixels which are most likely to be purest pixels that can represent endmembers. Then the ICA-AQA treats these extracted pixels as if they were endmembers to perform abundance quantification. As experimental results demonstrated in the following section, the ICA-AQA indeed performs at least comparably to fully constrained least squares abundance quantification method in both cases. Since the use of the VD and design of IC prioritization algorithms, HOS-ICPA and ID-ICPA are not our contribution which can be found in [17], in what follows we describe three major novelties derived from our proposed ICA-AQA.

1. One of major novelties is the idea of using independent component analysis (ICA) for abundance quantification that has not been explored before. Since the ICA uses statistical independence as blind source separation, the ICA cannot be implemented as a constrained method, thus it cannot be used as abundance quantification. The ICA-AQA proposed in this paper shows otherwise.
2. A second novelty is that our proposed ICA-based abundance quantification can perform both image endmember extraction and abundance quantification simultaneously which cannot be accomplished by any existing abundance quantifier or endmember extraction. For example, for ACLSMA to perform abundance quantification, it first requires the ACLSMA to find image endmembers first, which needs an algorithm to do so. After the image endmembers are determined, an AC-LSMA method must be followed up to estimate the abundance fractions. In this case, the algorithms for finding the image endmembers and the ACLSMA are generally different. The ICA-AQA uses the same ICA-generated independent components to perform image endmember extraction as well as abundance quantification.
3. A third major novelty is that the ICA-based abundance quantification is a high order statistics-based technique as opposed to commonly used ACLSMA methods which are least squares-based and second order statistics-based techniques. To authors' best knowledge, there exist no such high order statistics-based abundance quantification techniques reported in the literature.

## V. COMPUTER SIMULATIONS

In order to quantitatively evaluate the effectiveness of the ICA-based endmember extraction and abundance quantification, synthetic image-based computer simulations are conducted in this section for performance analysis. The synthetic images used for experiments were simulated by a library of five mineral spectra alunite (A), buddingtonite (B), calcite (C), kaolinite (K) and muscovite (M) available at the USGS website [26] shown in Fig. 1.

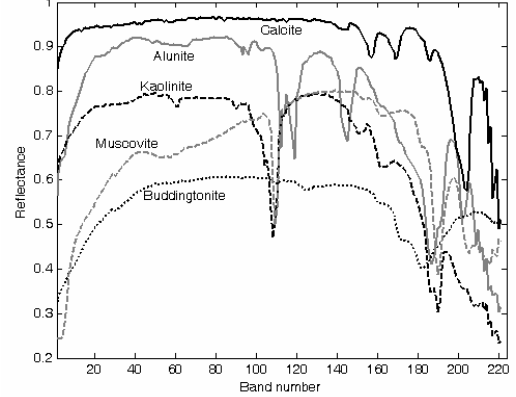


Figure 1. Spectra of five pure pixels corresponding to minerals: alunite (A), buddingtonite (B), calcite (C), kaolinite (K) and muscovite (M) provided by the USGS

Using the five signatures provided by Fig. 1 a synthetic image scene with size of  $64 \times 64$  pixels was simulated according to Fig. 2 where the image background was made up of 50% Alunite and 50% Kaolinite, and three other minerals Buddingtonite, Calcite, and Muscovite were used to simulate 27 panels of different abundance fractions with those panels in 1<sup>st</sup>, 2<sup>nd</sup> and 3<sup>rd</sup> columns specified by Buddingtonite, Calcite and Muscovite respectively.

These 27 panels were then implanted into the image background in a way that the corresponding background pixels were replaced by the pixels in the 27 panel. Except three panels in the first row in Fig. 2 labeled by 100% which contained 4 pure mineral pixels, all other panels in Fig. 2 were single-pixel panels with various abundance fractions as specified in Fig. 2. It is worth noting that the panel pixels with abundance fractions less than 1 were mixtures of the panel signature with the background. For example, the panel pixel labeled by 80% in the 1<sup>st</sup> column is a mixture of 80% Buddingtonite with 20% background signature.

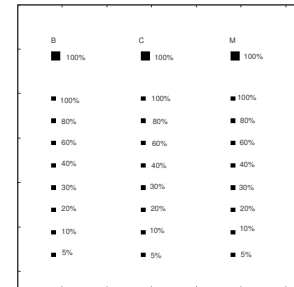


Figure 2. A synthetic image with specified abundance fractions

This synthetic image was used for experiments to simulate two scenarios, clean panels implanted into a noisy background and clean panels implanted into clean background with additive Gaussian noise. The former was design to evaluate the ability of an EEA in extracting endmembers in a noisy environment, while the latter is to evaluate the effectiveness of an EEA in extraction of endmembers corrupted by Gaussian noise, in which case there were no pure signatures.

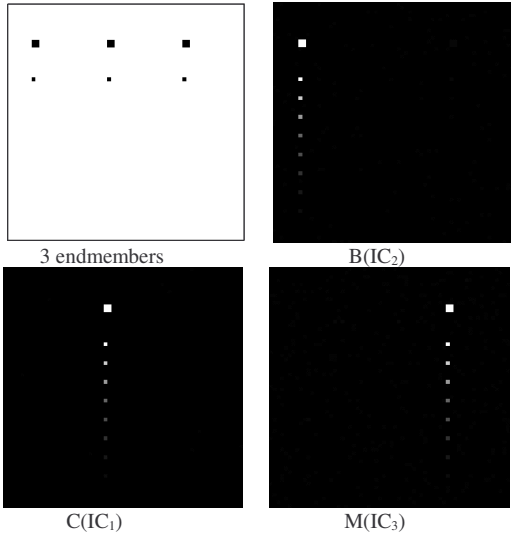
Experiment 1: Scenario with Endmembers (Clean panels implanted into a noisy background)

In this experiment we simulated a scenario that the 27 clean panels in Fig. 2 were implanted into a noisy image background which was also simulated by mixing 50% Alunite and 50% Kaolinite, but was also corrupted by an additive Gaussian noise to achieve signal-to-noise ratio (SNR) 30:1. So, the resulting synthetic image had 27 clean panels implanted in a noisy background in such a way that the pixels in the 27 panels replaced their corresponding noisy background pixels so that the 27 panels were not corrupted by the Gaussian noise. The VD for this synthetic image with various false alarm probabilities is tabulated in Table 1. So, for our experiments VD = 3 was chosen.

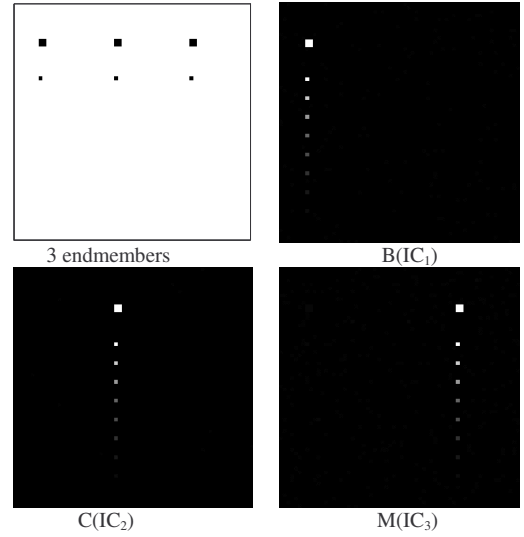
**Table 1.** VD estimates for the synthetic image in Ex.1 with various false alarm probabilities

	$P_F = 10^{-1}$	$P_F = 10^{-2}$	$P_F = 10^{-3}$	$P_F = 10^{-4}$	$P_F = 10^{-5}$
VD	3	3	3	3	3

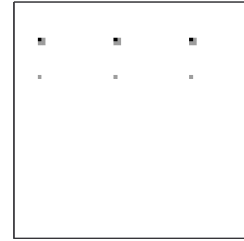
Fig. 3(a-c) shows the results of 3 endmembers extracted in three different ICs by the ICA-EEA using HOS-ICPA, ICA-EEA using ID-ICPA and PPI [4]. Since no noise was added to the 27 panels, the ICAEEA using the HOS-ICPA, ID-ICPA were able to extract all the five pure pixels in the 1<sup>st</sup> and 2<sup>nd</sup> rows in Fig. 3(a-b) for each of three minerals, B, C and M as endmember pixels. That is, a total of 15 panel pixels were extracted, 12 panels in three 2×2 panels in the 1<sup>st</sup> row and three 1×1 panel pixels in the 2<sup>nd</sup> row.



(a) 3 endmembers extracted by ICA-EEA using HOS-ICPA



(b) 3 endmembers extracted by ICA-EEA using ID-ICPA



(c) 3 endmembers extracted by PPI

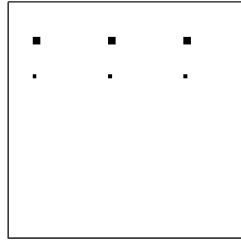
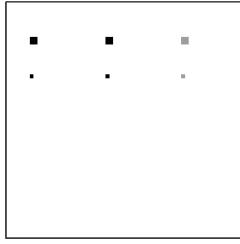
Figure 3. 3 endmembers extracted by ICA-EEA using HOS-ICPA ICA-EEA using ID-ICPA PPI

It is worth noting that all the ICs generated in Fig. 3(a-b) were real-valued and are shown in gray scales. The panel pixels which are not 100% pure were also detected in all ICs. But their gray level intensities were proportional to the abundance fractions of endmembers contained in the panel pixels, that is, the more the abundance the brighter the panel pixel. Interestingly, the PPI only extracted three panel pixels which were at the upper left corners of the three 2×2 panels in the 1<sup>st</sup> row in Fig. 3(c).

However, unlike the ICs in Fig. 3(a-b), the image in Fig. 3(c) was actually a binary map where the gray pixels are included to show those pixels which were pure panel pixels, but were not extracted by the PPI. It should be also noted that when the PPI was implemented, it required dimensionality reduction which was performed by the maximum noise fraction (MNF) [27] and the number of dimensions to be retained was set to  $p$ , in this case,  $p = 3$ .

In order to evaluate the performance of the ICA-AQA, an unsupervised fully constrained least squares (UFCLS) method [2,23] with  $p = 3$  was also used for comparison. The 3 panel pixels extracted by the UFCLS is shown in Fig. 4(a) where the UFCLS missed the mineral signature M in Fig. 4(a).





(a) UFCLS with  $p = 3$  (b) UFCLS with  $p = 4$   
Figure 4. 3 pixels extracted by UFCLS with  $p = 3$  and UFCLS with  $p = 4$

Table 2 also tabulates the results of the ICA-AQA using the HOS-ICPA and ID-ICPA along with the UFCLS in abundance quantification of all panel pixels in the 27 panels in Fig. 2.

**Table 2.** Abundance quantification results by ICA-AQA and UFCLS with  $p = 3$

Ground truth (%)	B (%)			C (%)			M (%)		
	ICA-AQA HOS-ICPA	ICA-AQA ID-ICPA	UFCLS ( $p = 3$ )	ICA-AQA HOS-ICPA	ICA-AQA ID-ICPA	UFCLS ( $p = 3$ )	ICA-AQA HOS-ICPA	ICA-AQA ID-ICPA	UFCLS ( $p = 3$ )
100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	66.75
100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	66.75
100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	66.75
100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	66.75
100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	66.75
80	79.96	79.96	80.33	79.96	79.96	80.41	79.96	79.96	53.73
60	59.92	59.92	60.65	59.92	59.92	60.82	59.93	59.93	40.70
40	39.89	39.88	40.98	39.88	39.88	41.24	39.89	39.89	27.68
30	29.87	29.86	31.14	29.86	29.86	31.44	29.87	29.87	21.16
20	19.85	19.85	21.30	19.84	19.84	21.65	19.85	19.85	14.65
10	9.83	9.83	11.46	9.82	9.82	11.85	9.83	9.83	8.14
5	4.82	4.82	6.54	4.81	4.81	6.96	4.82	4.83	4.88

As also shown in Table 2, all the 15 shaded panel pixels extracted in Fig. 3 and the 10 shaded panel pixels extracted in Fig. 4(a) were quantified to 100% pure pixels and the UFCLS failed to quantify the five panel pixels that corresponded to the mineral M to 100% pure pixels. The ICA-AQA performed significantly better than the UFCLS in quantification of other panel pixels as well. Since the background signature was spectrally distinct from the three endmembers, it should be accounted as the fourth signature in the UFCLS. Therefore, if the UFCLS was implemented with  $p = 4$ , Fig. 4(b) shows the same results as those obtained in Fig. 3 where a total of the 15 panel pixels in the 1<sup>st</sup> and 2<sup>nd</sup> rows were also extracted by the UFCLS.

Table 3 tabulates the UFCLS-quantified abundance fractions of the panel pixels in all the 27 panels in Fig. 2. Also included in Table 3 for comparison are the results produced by the FCLS using complete prior knowledge of three mineral signatures B,C,M with/without the background signature (BKG) where the FCLS using the complete knowledge of four signatures, B,C,M and BKG produced perfect results. However, if the FCLS used only three signatures, B,C,M for abundance quantification, it produced the worst performance and even worse than those obtained by the ICA-AQA in Table 2.

**Table 3.** Abundance quantification results of B, C, M by UFCLS and FCLS

Ground truth (%)	B (%)			C (%)			M (%)		
	UFCLS ( $p = 4$ )	FCLS		UFCLS ( $p = 4$ )	FCLS		UFCLS ( $p = 4$ )	FCLS	
		B,C,M	B,C,M,+ BKG		B,C,M	B,C,M + BKG		B,C,M	B,C,M, + BKG
100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
80	80.15	86.53	80.00	80.36	89.28	80.00	80.27	84.18	80.00
60	60.29	73.07	60.00	60.72	78.57	60.00	60.54	68.36	60.00
40	40.44	59.60	40.00	41.07	67.85	40.00	40.81	52.54	40.00
30	30.51	52.87	30.00	31.25	62.50	30.00	30.94	44.63	30.00
20	20.58	46.14	20.00	21.43	57.14	20.00	21.08	36.72	20.00
10	10.65	39.40	10.00	11.61	51.78	10.00	11.21	28.82	10.00
5	5.69	36.04	5.00	6.70	49.10	5.00	6.28	24.86	5.00

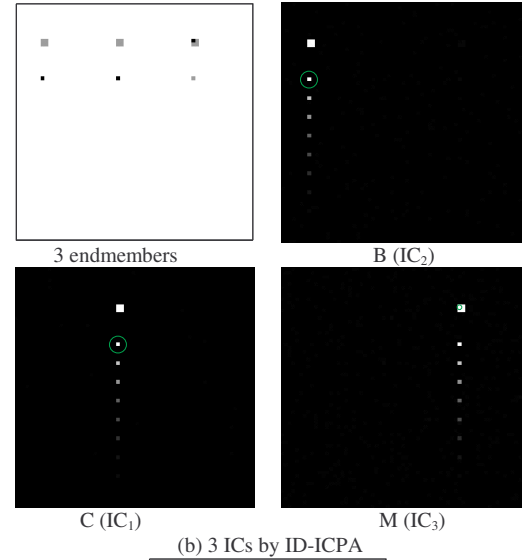
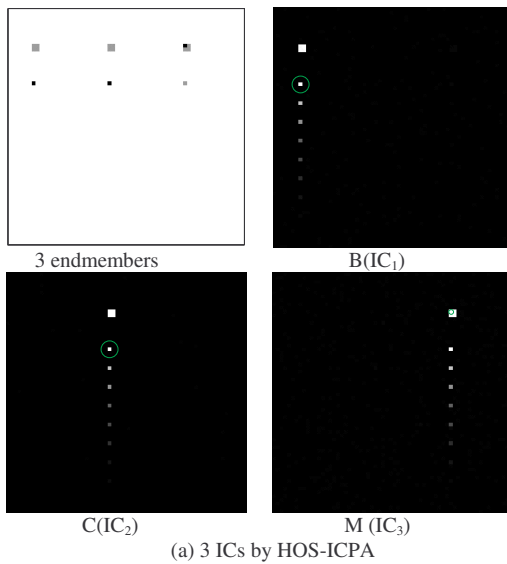
Experiment 2: Scenario with No Endmembers (Clean panels implanted into clean background with additive Gaussian noise)

In Experiment 1 we simulated a scenario where there were three endmembers present in a noisy background. In this experiment, we simulated another scenario where there was no true endmember present in a synthetic image. In doing so, an additive Gaussian noise was added to the image in Fig. 2 to achieve signal-to-noise ratio (SNR) 30:1 where the 27 clean panels and clean image background were all corrupted by an additive Gaussian noise. The VD for this synthetic image with various false alarm probabilities was estimated and tabulated in Table 4.

**Table 4.** VD estimates for the synthetic image in Experiment 2 with various false alarm probabilities

	$P_F = 10^{-1}$	$P_F = 10^{-2}$	$P_F = 10^{-3}$	$P_F = 10^{-4}$	$P_F = 10^{-5}$
VD	4	3	3	3	3

For our experiments  $VD = 3$  was chosen with  $P_F \leq 10^{-2}$ . Fig. 5(a-b) shows the results of 3 endmembers that are marked by circles and were extracted in three different ICs by the ICA-EEA using the HOS-ICPA and ID-ICPA respectively where both extracted two pixels in the 2<sup>nd</sup> row, each of which was a single panel pixel corresponding to the minerals B and C respectively and a third pixel which was the upper left corner panel pixel in the  $2 \times 2$  panel in the 3<sup>rd</sup> column that corresponded to the mineral M. So, the ICA-EEA still successfully extracted three endmembers even they were noise-corrupted. On the other hand, the PPI behaved quite differently. It extracted a total of 5 panel pixels in Fig. 5(c) with 1 single panel pixels in the 2<sup>nd</sup> row and 4 panel pixels in the 1<sup>st</sup> row, one located at bottom left corner of the  $2 \times 2$  panel in the 1<sup>st</sup> column, one located at upper right corner of the  $2 \times 2$  panel in the 2<sup>nd</sup> column and 2 at upper left and right corners of the  $2 \times 2$  panel in the 3<sup>rd</sup> column.



(c) endmembers extracted by PPI

Figure 5. Endmembers extracted by ICA-EEA using HOS-ICPA, ICA-EEA using ID-ICPA and PPI

Compare Fig. 5(a-b) to Fig. 5(c), the ICA-EEA performed as well as the PPI from an endmember extraction point of view since all the three endmembers corresponding to the three distinct minerals, B,C,M were extracted correctly.

Similarly, in analogy with Experiment 1 the UFCLS with  $p = 3$  and 4 was also implemented for comparison. Fig. 6(a) shows 3 pixels extracted by the UFCLS with  $p = 3$ , of which only one was a pure panel pixel located at the bottom right corner of the  $2 \times 2$  panel in the 1<sup>st</sup> row and the 2<sup>nd</sup> column. Fig. 6(b) shows 4 pixels extracted by the UFCLS with  $p = 4$ , of which three were pure panel pixels with one located at the bottom left corner of the  $2 \times 2$  panel at the 1<sup>st</sup> row and the 1<sup>st</sup> column, a second panel pixel being a single panel pixel at the 2<sup>nd</sup> row and the 2<sup>nd</sup> column and a third pixel located at the upper left corner of the  $2 \times 2$  panel at the 1<sup>st</sup> row and the 3<sup>rd</sup> column.

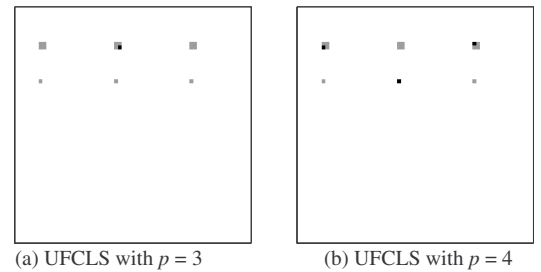


Figure 6. 3 pixels extracted by UFCLS with  $p = 3$  and UFCLS with  $p = 4$

As shown in Fig. 6(a-b), the UFCLS with  $p = 4$  outperformed the UFCLS with  $p = 3$  and successfully extracted three noise corrupted endmembers which

corresponded to three distinct minerals, B,C,M as the ICA-EEA did in Fig. 5(a-b).

Despite the fact that there were no endmembers present in the scene, the panel pixels extracted in Fig. 5(a-b) were assumed to be endmember pixels and their abundance fractions were then assigned to 100%. All other panel pixels were then normalized to the range of [0,1] by (3) accordingly. Table 5 tabulates quantification results of abundance fractions for the panel pixels in the 27 panels estimated by the ICA-AQA where the shaded pixels were the endmember pixels extracted by the ICA-EEA. Since only one of the five panel pixels for each of B, C and M was extracted as an endmember pixel and assigned to 100% abundance, the abundance fractions for the rest of the four pure panel pixels were normalized by (3). Nevertheless, their quantified abundance fractions were no less than 97.38%. For subpixels with abundance fractions less than 100%, the ICA-AQA actually performed well in abundance quantification except those subpixels with 5% abundance.

In comparison with the ICA-EEA the quantification results produced by the UFCLS with  $p = 3$  for the panel pixels in the 27 panels are also included in Table 5 where the UFCLS performed substantially worse than the ICA-AQA due to the fact that the background signature was not included in the UFCLS as the fourth signature. In this case, only one endmember pixel (shaded) corresponding to C was extracted by the UFCLS and the other two signatures B and M were

missed by the UFCLS. However, if the UFCLS was implemented with  $p = 4$ , Fig. 6(b) shows the 4 extracted pixels, three of which turned out to be panel pixels representing three distinct minerals. The quantitative results produced by the UFCLS for the panel pixels in the 27 panels are tabulated in Table 6 where three endmember pixels (shaded) corresponding to three mineral signatures B,C,M extracted by the UFCLS in Fig. 6(b). In this case, the UFCLS performed better than the ICA-AQA.

Table 6 suggested that in order for the UFCLS to perform well, using only endmembers as its image endmembers was not sufficient. It must include all potential spectrally distinct signatures other than endmembers such as background signatures and interfering signatures. For the ICA-EEA such signatures are extracted in ICs with low priorities. To further verify the results, a supervised fully constrained least squares (FCLS) method [2,23] was implemented using complete prior knowledge of three mineral signatures (B,C,M) with/without the background signature (BKG). Their results are also tabulated in Table 6 where the FCLS using the B, C, M with the background signature BKG performed better than both the ICA-EEA and UFCLS. By contrast, the FCLS using only three minerals, B, C, M without the background signature BKG performed very poorly particularly for subpixels because the background signature was not included as an undesired signature.

**Table 5.** Endmembers extracted by ICA-EEA and abundance quantification results by ICA-AQA and UFCLS

Ground truth (%)	B (%)			C (%)			M (%)		
	HOS-ICPA	ID-ICPA	UFCLS ( $p = 3$ )	HOS-ICPA	ID-ICPA	UFCLS ( $p = 3$ )	HOS-ICPA	ID-ICPA	UFCLS ( $p = 3$ )
100	98.04	98.03	98.67	99.15	99.17	98.99	100.00	100.00	19.54
100	97.50	97.47	99.08	99.39	99.41	99.24	97.38	97.38	18.91
100	98.91	98.90	98.73	99.20	99.21	99.09	98.43	98.43	20.19
100	98.43	98.42	99.39	99.97	99.99	100.00	98.78	98.78	19.97
100	100.00	100.00	99.99	100.00	100.00	99.65	98.72	98.72	19.39
80	76.54	76.54	79.49	77.17	77.16	79.65	77.18	77.18	16.68
60	57.49	57.49	61.02	58.05	58.04	60.21	57.55	57.55	12.11
40	37.88	37.87	41.33	38.60	38.60	40.89	38.45	38.45	8.42
30	28.74	28.74	30.75	28.32	28.32	31.12	28.37	28.37	7.21
20	18.68	18.68	21.67	19.27	19.27	20.84	17.61	17.61	5.52
10	8.56	8.55	11.33	8.70	8.69	11.03	9.17	9.17	3.36
5	4.26	4.26	7.39	5.58	5.58	6.41	3.73	3.73	2.64

**Table 6.** Abundance quantification results of B, C, M by UFCLS and FCLS

Ground truth (%)	B (%)			C (%)			M (%)		
	UFCLS ( $p = 4$ )	FCLS		UFCLS ( $p = 4$ )	FCLS		UFCLS ( $p = 4$ )	FCLS	
		B,C,M	B,C,M,+ BKG		B,C,M	B,C,M + BKG		B,C,M	B,C,M, + BKG
100	97.99	99.21	99.21	98.68	99.48	99.33	100.00	100.00	100.00
100	98.07	99.54	99.54	99.09	99.82	99.63	96.67	99.23	98.70
100	99.09	100.00	100.00	98.73	99.79	99.35	96.84	99.13	99.13
100	100.0	100.00	100.00	99.39	100.00	99.96	97.34	99.46	99.46
100	99.65	100.00	100.00	100.00	100.00	100.00	96.97	99.42	99.31
80	78.46	85.87	79.71	79.44	88.30	79.69	78.44	84.08	80.07
60	59.94	73.23	60.09	60.77	78.29	60.60	59.35	68.93	60.10
40	40.33	59.58	40.15	41.22	67.95	40.45	40.34	53.85	40.38

30	30.87	53.12	30.32	30.71	62.60	29.98	29.95	44.89	29.81
20	20.79	46.23	19.73	21.67	57.74	20.56	18.87	35.84	19.04
10	10.55	39.07	9.74	11.07	51.64	9.92	10.11	28.71	9.51
5	5.51	35.44	4.59	7.32	49.37	5.57	4.81	24.63	4.56

However, it should be noted that the UFCLS was performed by two stage processes where an unsupervised least square error-based algorithm was implemented to find a set of image endmembers including endmembers and undesired signatures in the first stage, then followed by a supervised FCLS in the second stage to quantify abundance fractions. If the set of found image endmembers was not representative for the image data to be processed, the UFCLS would generally fail as shown in Table 2 and Table 3. Compared to the UFCLS, the ICA-AQA performs endmember extraction and abundance quantification simultaneously in a one-shot single operation. This advantage may give the ICA-AQA better utility in real applications. Similarly, Experiment 2 was also conducted for SNR = 20:1 and 10:1 where the results were very similar to what we presented here. Therefore, their experimental results are not included.

## VI. REAL HYPERSPECTRAL IMAGE EXPERIMENTS

In this section, two real hyperspectral image data were used for experiments. One is a well-known Airborne Visible/InfraRed imaging spectrometer (AVIRIS) Cuprite image scene and the second one is HYperspectral Digital Imagery Collection Experiment (HYDICE) image scene.

### A. AVIRIS Image Experiments

The AVIRIS Cuprite image scene shown in Fig. 7 is available at website [26] and was collected over the Cuprite mining site, Nevada, in 1997. It is a 224-band image with size of  $350 \times 350$  pixels and well understood mineralogically where bands 1-3, 105-115 and 150-170 have been removed prior to the analysis due to water absorption and low SNR in those bands. As a result, a total of 189 bands were used for experiments. Although there are more than five minerals on the data set, the ground truth available for this region only provides the locations of the pure pixels: Alunite (A), Buddingtonite (B), Calcite (C), Kaolinite (K) and Muscovite (M). The locations of these five pure minerals are labeled by A, B, C, K and M respectively and shown in Fig. 7.

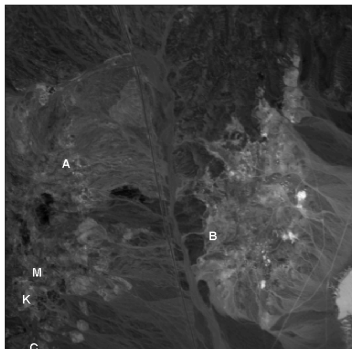


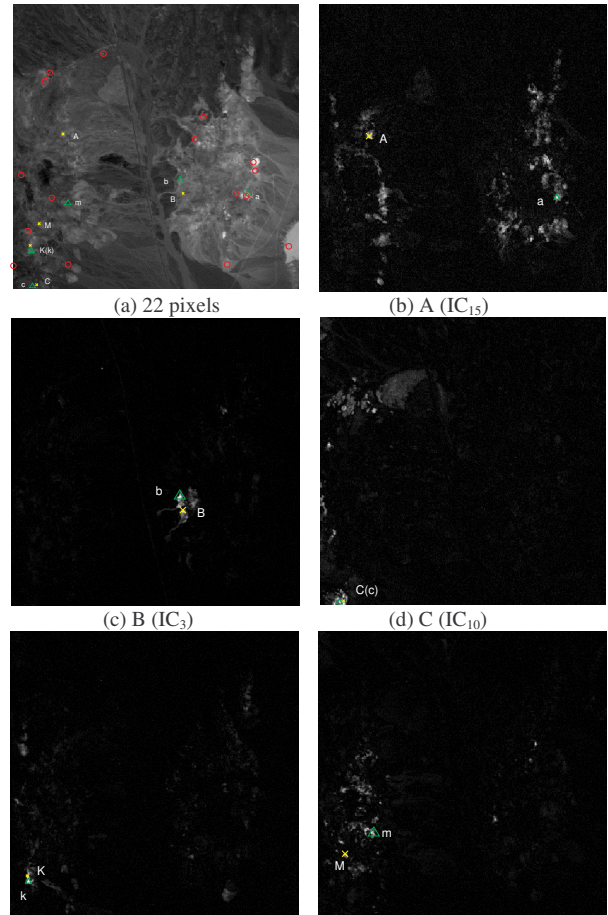
Figure 7. Cuprite AVIRIS image scene g with spatial positions of five pure pixels corresponding to minerals: alunite (A), buddingtonite (B), calcite (C), kaolinite (K) and muscovite (M)

This scene is well understood mineralogically and has reliable ground truth. It becomes a standard test site for comparison of endmember extraction algorithms. The VD for this image scene was estimated with various false alarm probabilities and tabulated in Table 7. For our experiments VD = 22 was chosen with  $P_F = 10^{-4}$ .

**Table 7.** VD estimates for the AVIRIS scene in Fig. 7 with various false alarm probabilities

	$P_F = 10^{-1}$	$P_F = 10^{-2}$	$P_F = 10^{-3}$	$P_F = 10^{-4}$	$P_F = 10^{-5}$
VD	34	30	24	22	20

Figs. 8-9 show the spatial locations of the 22 pixels extracted in Fig. 8(a) and Fig. 9(a) by the ICA-EEA using the HOS-ICPA and the ID-ICPA along with their corresponding five different ICs in Fig. 8(b-f) and Fig. 9(b-f) that extracted the five pixels marked by a, b, c, k, m with triangles were shown to be closest to the spatial locations of the five ground truth mineral endmember pixels marked by A, B, C, K, M with crosses.





(e) K (IC<sub>5</sub>)(f) M (IC<sub>7</sub>)

Figure 8. 22 pixels extracted by ICA-EEA using HOS-ICPA for endmember extraction

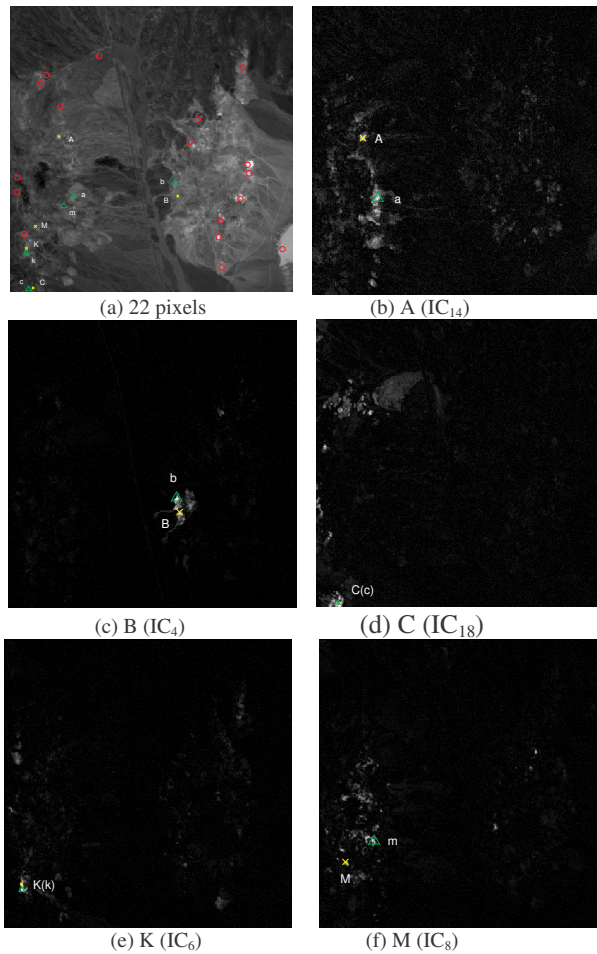


Figure 9. 22 pixels extracted by ICA-EEA using ID-ICPA for endmember extraction

The indexes of ICs in Fig. 8-9 were included to indicate the order that they were generated. The spectral similarity measure used for mineral identification was the correlation-matched filter based distance (RMFD) [28]. The reason that the RMFD was used in this experiment is because it was shown in [28] that the RMFD was preferred to the commonly used spectral similarity measure, spectral angle mapper (SAM) for real hyperspectral imagery. The remaining 17 pixels marked by circles in Fig. 8(a) and Fig. 9(a) were not endmember pixels that corresponded to the five minerals, but could be some other unidentified substances which were not available from the ground truth. Fig. 10 shows the result obtained by the PPI with dimensions reduced to  $p = 22$  for comparison where the pixels marked by triangle were the extracted endmembers correspond to 5 minerals, and circles were unidentified endmembers. As noted, the PPI missed one mineral, which is Buddingtonite. Comparing the PPI result in Fig. 10 to the results produced in Figs. 8-9 by the ICA-EEA, the ICA-EEA performed better than the PPI did.

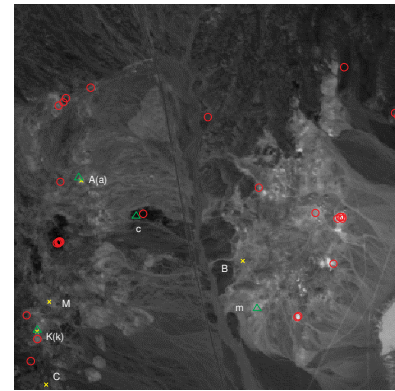
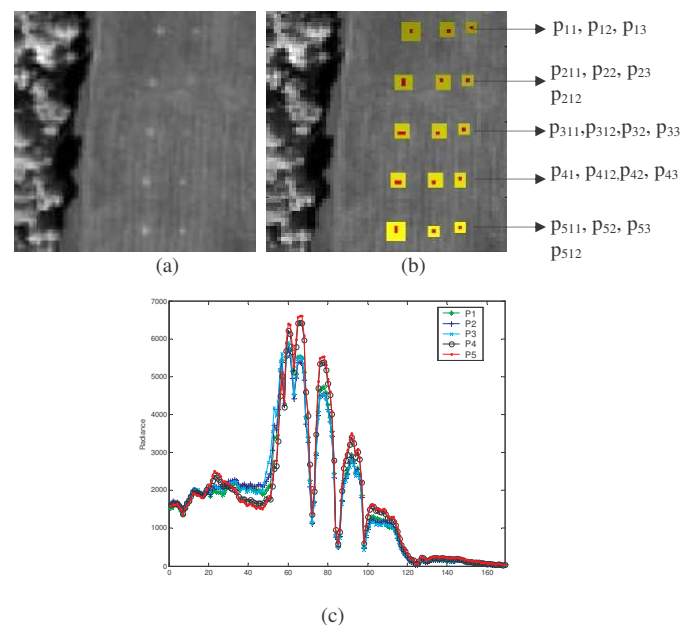


Figure 10. 22 pixels extracted by PPI for endmember extraction

It should be noted that due to the lack of information about abundance fractions of pixels in the image, the ICA-AQA was not implemented for this scene.

### B. HYDICE Image Experiments

Another image scene to be studied is HYperspectral Digital Imagery Collection Experiment (HYDICE) image scene shown in Fig. 11(a) was used for experiments, which has a size of  $64 \times 64$  pixel vectors with 15 panels in the scene and the ground truth map in Fig. 11(b) [1]. It was acquired by 210 spectral bands with a spectral coverage from  $0.4\mu\text{m}$  to  $2.5\mu\text{m}$ . Low signal/high noise bands: bands 1-3 and bands 202-210; and water vapor absorption bands: bands 101-112 and bands 137-153 were removed. So, a total of 169 bands were used. The spatial resolution is  $1.56\text{m}$  and spectral resolution is  $10\text{nm}$ . Within the scene in Fig. 11(a) there is a large grass field background, and a forest on the left edge. Each element in this matrix is a square panel.

Figure 11. (a) A HYDICE panel scene which contains 15 panels; (b) Ground truth map of spatial locations of the 15 panels; (c) Spectral signatures of  $p_1$ ,  $p_2$ ,  $p_3$ ,  $p_4$  and  $p_5$

For each row  $i = 1, 2, \dots, 5$ , the three panels were painted by the same material but have three different sizes. For each column  $j = 1, 2, 3$ , the five panels have the same size but were painted by five different materials. It should be noted that the panels in rows 2 and 3 are made by the same material with different paints, so did the panels in rows 4 and 5. Nevertheless, they were still considered as different materials. The sizes of the panels in the first, second and third columns are  $3\text{m} \times 3\text{m}$ ,  $2\text{m} \times 2\text{m}$  and  $1\text{m} \times 1\text{m}$  respectively. So, the 15 panels have five different materials and three different sizes. Fig. 11(b) shows the precise spatial locations of these 15 panels where red pixels (R pixels) are the panel center pixels and the pixels in yellow (Y pixels) are panel pixels mixed with background. The 1.56m-spatial resolution of the image scene suggests that the panels in the 2<sup>nd</sup> and 3<sup>rd</sup> columns, denoted by  $p_{12}$ ,  $p_{13}$ ,  $p_{22}$ ,  $p_{23}$ ,  $p_{32}$ ,  $p_{33}$ ,  $p_{42}$ ,  $p_{43}$ ,  $p_{52}$ ,  $p_{53}$  in Fig. 11(b) are one pixel in size. Additionally, except the panel in the 1<sup>st</sup> row and 1<sup>st</sup> column, denoted by  $p_{11}$  which also has size of one pixel, all other panels located in the 1<sup>st</sup> column are two-pixel panels which are the panel in the 2<sup>nd</sup> row with two pixels lined up vertically, denoted by  $p_{211}$  and  $p_{221}$ , the panel in the 3<sup>rd</sup> row with two pixels lined up horizontally, denoted by  $p_{311}$  and  $p_{312}$ , the panel in the 4<sup>th</sup> row with two pixels also lined up horizontally, denoted by  $p_{411}$  and  $p_{412}$ , and the panel in the 5<sup>th</sup> row with two pixels lined up vertically, denoted by  $p_{511}$  and  $p_{521}$ . Since the size of the panels in the third column is  $1\text{m} \times 1\text{m}$ , they cannot be seen visually from Fig. 11(a) due to the fact that its size is less than the 1.56m pixel resolution.

Fig. 11(c) plots the 5 panel spectral signatures  $\mathbf{p}_i$  for  $i = 1, 2, \dots, 5$  obtained by averaging R pixels in the  $3\text{m} \times 3\text{m}$  and  $2\text{m} \times 2\text{m}$  panels in row  $i$  in Fig. 11(b). It should be noted the R pixels in the  $1\text{m} \times 1\text{m}$  panels are not included because they are not pure pixels due to that fact that the spatial resolution of the R pixels in the  $1\text{m} \times 1\text{m}$  panels is 1 m smaller than the pixel resolution, 1.56 m. These panel signatures along with the R pixels in the  $3\text{m} \times 3\text{m}$  and  $2\text{m} \times 2\text{m}$  panels were used as required prior target knowledge for the following comparative studies.

First of all, the VD was used to estimate number of bands,  $p$ , required for band selection. Table 8 calculated the values of VD for the HYDICE image in Fig. 11(a) with various false alarm probabilities. For our experiments, the VD was chosen to be 9. The selection of  $p = 9$  was empirical based on the false alarm fixed at probabilities  $P_F = 10^{-4}$ .

**Table 8.** VD estimates for the HYDICE scene in Fig. 11 with various false alarm probabilities

	$P_F = 10^{-1}$	$P_F = 10^{-2}$	$P_F = 10^{-3}$	$P_F = 10^{-4}$	$P_F = 10^{-5}$
VD	14	11	9	9	7

Fig. 12-13 shows the spatial locations of 9 pixels extracted in Fig. 12(a) by the ICA-EEA using the HOS-ICPA and the ID-ICPA along with their corresponding five different ICs indexed by the order that they appeared, Fig. 12(b-f) and Fig. 13(b-f) that extracted the five distinct panel

signatures where endmember pixels marked by triangles were those corresponded to ground truth R pixels marked by crosses for comparison. The remaining 4 endmember pixels marked by circles in Fig. 12(a) and Fig. 13(a) were not R pixels but could correspond to tree, forest, grass or road in the image background. Comparing Fig. 12(b-f) and Fig. 13(b-f) against the ground truth map in Fig. 9(b), the ICA-EEA using HOS-ICPA and ID-ICPA extracted the same five panel pixels,  $p_{11}$ ,  $p_{221}$ ,  $p_{312}$ ,  $p_{411}$  and  $p_{521}$  that represented the five distinct panel signatures  $\mathbf{p}_1$ ,  $\mathbf{p}_2$ ,  $\mathbf{p}_3$ ,  $\mathbf{p}_4$  and  $\mathbf{p}_5$  despite that both found four different remaining pixels marked by circles in Fig. 12(a) and Fig. 13(a). Fig. 14 shows the results obtained by the PPI with dimensions reduced to  $p = 9$  extracted 11 pixels, of which only four endmember pixels,  $p_{311}$ ,  $p_{312}$ ,  $p_{412}$  and  $p_{521}$  that corresponded to three distinct panel signatures (panel signatures  $\mathbf{p}_3$ ,  $\mathbf{p}_4$  and  $\mathbf{p}_5$  in rows 1, 3 and 5) and missed the other two panel signatures  $\mathbf{p}_1$  and  $\mathbf{p}_2$ . The remaining endmember pixels marked by circles were not R pixels, but rather background pixels.

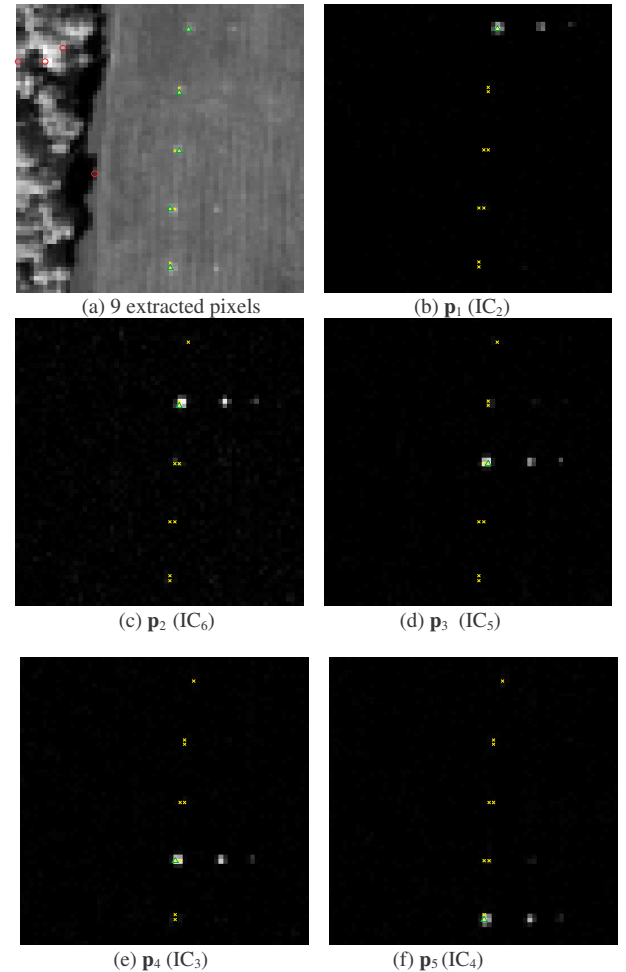


Figure 12. 9 pixel extracted by the ICA-EEA using HOS-ICPA

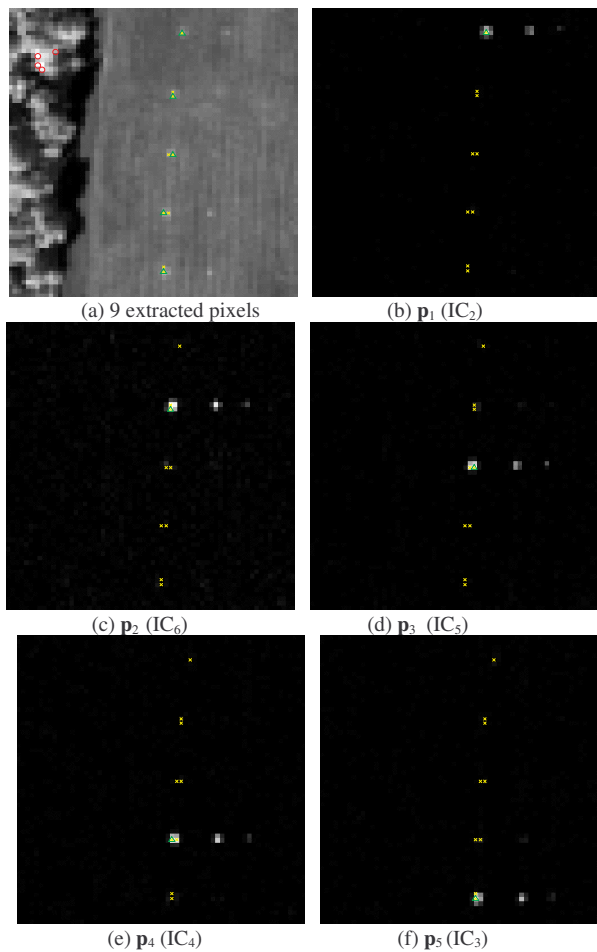


Figure 13. 9 pixel extracted by the ICA-EEA using ID-ICPA

This experiment showed that the ICA-EEA performed significantly better than the PPI in endmember extraction where the former found all five endmember pixels including the three extracted by the latter which missed the second and fourth signatures.

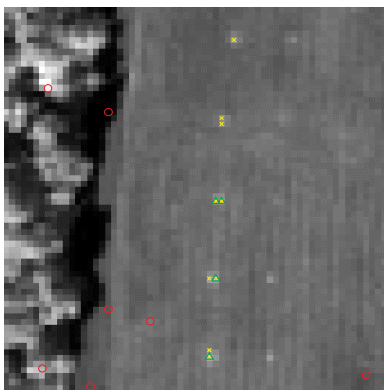
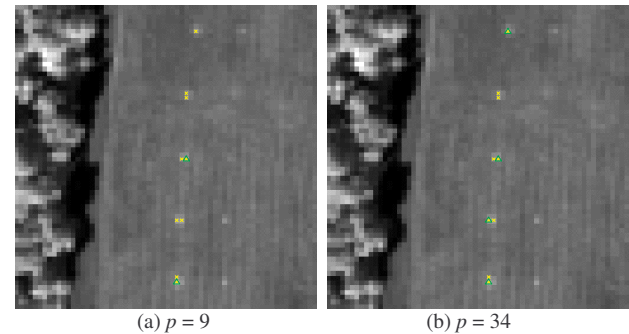


Figure 14. 9 pixels extracted by PPI

Since the abundance fractions of the 19 R pixels in Fig. 11(b) can be calculated based on the ratio of their panel size to the spatial resolution [29], the ICA-AQA was implemented and evaluated in comparison with the UFCLS for performance analysis. Since it has been shown in [2,23] that in order for the UFCLS to perform well it required at least 34 signatures to be used for linear unmixing, Fig.

15(a-b) shows the results of pixels extracted by the UFCLS with  $p = 9$  and 34 respectively where two pixels,  $p_{312}$  and  $p_{521}$  were extracted by the UFCLS with  $p = 9$  and 4 pixels,  $p_{11}$ ,  $p_{312}$ ,  $p_{411}$  and  $p_{521}$  were extracted by the UFCLS with  $p = 34$ .

Figure 15. Pixels extracted by the UFCLS with  $p = 9$  and 34

Using the pixels extracted in Fig. 12(b-f) and Fig. 13(b-f) as endmember pixels, Table 9 tabulates the quantification results of the 19 R panel pixels produced by the ICA-AQA where the shaded pixels were the pixels extracted by algorithms in Fig. 12(b-f) and Fig. 13(b-f). Also included in Table 9 are the abundance quantification results of 19 R panel pixels produced by the UFCLS with  $p = 9$  and  $p = 34$  with the shaded pixels extracted in Fig. 15(a-b).

**Table 9.** Abundance fractions of all the 19 R pixels by ICA-AQA and UFCLS in Fig. 6

	ICA-AQA		UFCLS	
	HOS-ICPA	ID-ICPA	$p = 9$	$p = 34$
$p_{11}$	1	1	0.3498	1
$p_{12}$	0.4458	0.4458	0.1982	0.4098
$p_{13}$	0.1745	0.1744	0.0712	0.0499
$p_{211}$	0.9986	0.9988	0.4291	0.5255
$p_{212}$	1	1	0.4358	0.3141
$p_{22}$	0.9215	0.9209	0.3560	0.6917
$p_{23}$	0.2733	0.2719	0.1978	0.4221
$p_{311}$	0.8943	0.8981	0.9273	0.8647
$p_{312}$	1	1	1	1
$p_{32}$	0.5358	0.5396	0.5127	0.5343
$p_{33}$	0.3356	0.3381	0.3735	0.3285
$p_{411}$	1	1	0.7380	1
$p_{412}$	0.9	0.9012	0.8066	0.3821
$p_{42}$	0.7907	0.7889	0.6397	0.7034
$p_{43}$	0.1918	0.1897	0.1772	0.2242
$p_{511}$	0.7004	0.7001	0.7242	0.7203
$p_{521}$	1	1	1	1
$p_{52}$	0.7306	0.7322	0.7742	0.7789
$p_{53}$	0.1301	0.1301	0.1584	0.1466

From Table 9, the ICA-AQA performed significantly better than the UFCLS in both cases. The poor performance of the UFCLS with  $p = 9$  was due to insufficient number of signatures used for UFCLS. Even though the performance of the UFCLS with  $p = 34$  significantly improved the UFCLS with  $p = 9$ , but its performance was still not as good as that of the ICA-AQA. Comparing results in Table 9 both the ICA-AQA and the UFCLS performed rather



differently on panels in the 1<sup>st</sup> and 2<sup>nd</sup> rows. Since there is no complete prior knowledge about the HYDICE image scene other than those 15 panels, no experiments were conducted for the supervised FCLS for comparison. Table 9 also revealed an intriguing difference between the ICA-EEA and a fully abundance-constrained method such as UFCLS. If the image data contain pure pixels, i.e., endmember pixels, the ICA-AQA can do as well as a fully abundance constrained method does. On some occasions, it may perform even better. On the other hand, if there are no pure pixels present in the image data, the ICA-AQA is forced to find one which may be a mixed pixel, but very likely a purest pixel. Under this circumstance, the abundance fraction of the found mixed pixel is normalized to one and the abundance fractions of all other pixels are also normalized by (3) accordingly. Consequently, the obtained abundance fractions may not be accurate. Such phenomena were demonstrated by Fig. 5 and Table 5 conducted for Experiment 2 and also by the HYDICE experiments where both pixels,  $p_{211}$  and  $p_{212}$  in Fig. 6(c) are R pixels but they are not pure pixels according to the experiments conducted in [29]. In this case, the ICA-AQA found the pixel  $p_{212}$  as a desired endmember pixel and normalized it to one. The abundance fraction of the second pixel  $p_{211}$  in the same 2-pixel panel was then quantified as 0.9986 (HOS-ICPA) or 0.9988 (ID-ICPA). However, according to the UFCLS the abundance fractions of these two pixels were estimated as 0.5255 and 0.3141 respectively. Therefore, there was discrepancy between the ICA-AQA and the UFCLS in abundance quantification when no pure pixels are in the image data. Under this circumstance, UFCLS produces more accurate abundance estimation than the ICA-AQA. In the latter case, the ICA-AQA found the purest pixel as an endmember pixel and then treated as if it was a pure pixel by normalizing its abundance fraction to one. Then all other pixels are quantified accordingly via (3) as the way that it did for quantification of the panel that is made up of two panel pixels,  $p_{211}$  and  $p_{212}$ .

One of major advantages derived from the proposed ICA-AQA is a one-shot process in endmember extraction and quantification compared to any other unsupervised abundance quantification algorithm which is generally carried out in two stages, endmember extraction followed by abundance quantification. As expected, there is a significant saving in computational complexity for the ICA-AQA. In order to see this, Table 10 tabulates the CPU time required for the ICA-AQA and UFCLS to generate the results shown in Table 9 where the computer environment is tabulated in Table 11. Table 10 shows that the ID-ICPA yielded the best time as opposed to the worst time produced by the UFCLS generating 34 target pixels.

**Table 10.** CPU time required by ICA-AQA and UFCLS

	ICA-AQA		UFCLS	
	HOS-ICPA	ID-ICPA	$p = 9$	$p = 34$
CPU Time (seconds)	7.56	3.79	19.16	276.17

**Table 11.** Computer environment used to run the algorithms

CPU	Memory	OS	Matlab Version
Intel(R) Pentium(R) 4 CPU 2.66GHz	1G	Windows XP	6.5

## VII. CONCLUSIONS

This paper presents an ICA-based approach to endmember extraction and abundance quantification, called ICA-abundance quantification algorithm (ICA-AQA). Of particular interest is a detailed study using synthetic image-based computer simulations conducted for performance analysis. The proposed ICA-AQA first uses a new concept of virtual dimensionality (VD) recently developed in [2,18] to resolve a long standing and challenging issue, determination of number of ICs required to be generated,  $p$  for endmember extraction. It then implements an ICA-based endmember extraction algorithm (ICA-EEA) to extract desired image endmembers where two IC prioritization algorithms, high order statistics-based IC prioritization algorithm (HOS-ICPA) and initialization driven-based IC prioritization algorithm (ID-ICPA) are used to prioritize the ICs so that an appropriate set of  $p$  ICs can be selected for endmember extraction. Finally, the same ICA-EEA selected  $p$  prioritized ICs are further used for abundance quantification, a task that cannot be accomplished by any endmember extraction algorithm. Both synthetic image-based computer simulations and real hyperspectral image experiments are conducted to substantiate the ICA-AQA for quantitative analysis as well as to validate their utility in real applications. The experiments demonstrate that the ICA-EEA performs as well as or even better than a commonly used endmember extraction algorithm, PPI. Additionally, the ICA-AQA also performs comparably to fully abundance-constrained linear spectral mixture analysis methods such as UFCLS [23]. It should be noted that similar conclusions drawn for the PPI are also true for another popular endmember extraction algorithm, N-FINDR [5]. Since both results are very similar, the experiments conducted for the N-FINDR are not included in this paper.

## ACKNOWLEDGMENT

The authors would like to acknowledge the use of the FastICA available on the website. They would also like to thank Professor A. Plaza who provides the Cuprite AVIRIS data with ground truth and valuable discussions including PPI's implementation and results for Cuprite data. Finally, the authors would like to thank anonymous reviewers for their thoughtful comments that greatly improve this paper's quality and presentation.

## APPENDIX

The purpose of this appendix is to provide a brief introduction of the concept of the VD and a method, called Harsanyi-Farrand-Chang (HFC) method developed in [30] to estimate the VD. The details about the VD can be found in [18]. The name of virtual dimensionality (VD) was



originally coined in [2] and later in [18]. It was designed to determine the number of spectrally distinct signatures. If a component such as PC or IC is used to accommodate a spectrally distinct signature for classification and identification, the number of required components happens to be the number of spectrally distinct signatures, which is the VD. Despite several methods were developed in [18], the method developed by Harsanyi et al. [30], referred to as Harsanyi-Farrand-Chang's (HFC) method is selected for two reasons. One is simple to implement. Another is that it was shown to be effective in determining the number of spectrally signatures for AVIRIS data [30]. Its idea is very simple. It first calculates the sample correlation matrix,  $\mathbf{R}$ , and sample covariance matrix,  $\mathbf{K}$ , then finds the difference between their corresponding eigenvalues.

More specifically, let  $\{\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_L\}$  and  $\{\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L\}$  be two sets of eigenvalues generated by  $\mathbf{R}$  and  $\mathbf{K}$ , called correlation eigenvalues and covariance eigenvalues, respectively where the  $L$  is the number of spectral channels. By assuming that signal sources are nonrandom unknown positive constants and noise is white with zero mean, we can expect that

$$\hat{\lambda}_l > \lambda_l \text{ for } l = 1, \dots, VD, \quad (\text{A-1})$$

and

$$\hat{\lambda}_l = \lambda_l \text{ for } l = VD + 1, \dots, L. \quad (\text{A-2})$$

Using (A-1)-(A-2), the eigenvalues in the  $l$ -th spectral channel can be related by

$$\hat{\lambda}_l > \lambda_l > \sigma_{n_l}^2 \text{ for } l = 1, \dots, VD$$

and

$$\hat{\lambda}_l = \lambda_l = \sigma_{n_l}^2 \text{ for } l = VD + 1, \dots, L \quad (\text{A-3})$$

where  $\sigma_{n_l}^2$  is the noise variance in the  $l$ -th spectral channel.

In order to determine the VD, Harsanyi et al. [25] formulated the VD determination problem as a binary hypothesis problem as follows.

$$H_0 : z_l = \hat{\lambda}_l - \lambda_l = 0$$

$$\text{versus} \quad \text{for } l = 1, 2, \dots, L \quad (\text{A-4})$$

$$H_1 : z_l = \hat{\lambda}_l - \lambda_l > 0$$

where the null hypothesis  $H_0$  and the alternative hypothesis  $H_1$  represent the case that the correlation-eigenvalue is equal to its corresponding covariance eigenvalue and the case that the correlation-eigenvalue is greater than its corresponding covariance eigenvalue, respectively. In other words, when  $H_1$  is true (i.e.,  $H_0$  fails), it implies that there is an endmember contributing to the correlation-eigenvalue in addition to noise, since the noise energy represented by the eigenvalue of  $\mathbf{R}$  in that particular component is the same as the one represented by the eigenvalue of  $\mathbf{K}$  in its corresponding component.

Despite the fact that the  $\hat{\lambda}_l$  and  $\lambda_l$  in (A-1)-(A-3) are unknown constants, according to [26], we can model each pair of eigenvalues,  $\hat{\lambda}_l$  and  $\lambda_l$ , under hypotheses  $H_0$  and

$H_1$  as random variables by the asymptotic conditional probability densities given by

$$p_0(z_l) = p(z_l | H_0) \equiv N(0, \sigma_{z_l}^2) \text{ for } l = 1, 2, \dots, L \quad (\text{A-5})$$

and

$$p_1(z_l) = p(z_l | H_1) \equiv N(\mu_l, \sigma_{z_l}^2) \text{ for } l = 1, 2, \dots, L \quad (\text{A-6})$$

respectively, where  $\mu_l$  is an unknown constant and the variance  $\sigma_{z_l}^2$  is given by for  $l = 1, 2, \dots, L$

$$\sigma_{z_l}^2 = \text{var}[\hat{\lambda}_l - \lambda_l] = \text{var}[\hat{\lambda}_l] + \text{var}[\lambda_l] - 2\text{cov}(\hat{\lambda}_l, \lambda_l). \quad (\text{A-7})$$

It has been shown that when the total number of samples,  $N$  is sufficiently large,  $\text{var}[\hat{\lambda}_l] \approx 2\hat{\lambda}_l^2 / N$  and  $\text{var}[\lambda_l] \approx 2\lambda_l^2 / N$ . Therefore, the noise variance  $\sigma_{z_l}^2$  in (A-6) can be estimated and approximated using (A-7).

From (A-5), (A-6) and (A-9), we define the false alarm probability and detection power (i.e., detection probability) as follows:

$$P_F = \int_{\tau_l}^{\infty} p_0(z) dz \quad (\text{A-8})$$

$$P_D = \int_{\tau_l}^{\infty} p_1(z) dz. \quad (\text{A-9})$$

A Neyman-Pearson detector for  $\hat{\lambda}_l - \lambda_l$ , denoted by  $\delta_{NP}(\hat{\lambda}_l - \lambda_l)$  for the binary composite hypothesis testing problem specified by (A-4) can be obtained by maximizing the detection power  $P_D$  in (A-9), while the false alarm probability  $P_F$  in (A-8) is fixed at a specific given value,  $\alpha$ , which determines the threshold value  $\tau_l$  in (A-8)-(A-9). So a case of  $\hat{\lambda}_l - \lambda_l > \tau_l$  indicating that  $\delta_{NP}(\hat{\lambda}_l - \lambda_l)$  fails the test, in which case there is signal energy assumed to contribute to the eigenvalue,  $\hat{\lambda}_l$ , in the  $l$ -th data dimension. It should be noted that the test for (A-4) must be performed for each of  $L$  spectral dimensions. Therefore, for each pair of  $\hat{\lambda}_l - \lambda_l$ , the threshold  $\tau$  is different and should be  $l$ -dependent, that is  $\tau_l$ .

## REFERENCES

- [1] SPIE conference on Defense and Security (previously known as AeroSense), *Algorithms and Technologies for Multispectral, Hyperspectral and Ultraspectral Imagery*.
- [2] C.-I Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, Kluwer Academic/Plenum Publishers, 2003.
- [3] R.A. Schwengerdt, *Remote Sensing: Models and Methods for Image Processing*, 2<sup>nd</sup>. Ed., Academic Press, 1997, p. 447.
- [4] J.W. Boardman, F.A. Kruse and R.O. Green, "Mapping target signatures via partial unmixing of AVIRIS data," *Summaries of JPL Airborne Earth Science Workshop*, Pasadena, CA, 1995.
- [5] M.E. Winter, "N-finder: an algorithm for fast autonomous spectral endmember determination in hyperspectral data," *Image Spectrometry V*, Proc. SPIE 3753, pp. 266-277, 1999.
- [6] R.A. Neville, K. Staenz, T. Szeredi, J. Lefebvre and P. Hauff, "Automatic endmember extraction from hyperspectral data for mineral exploration," *4<sup>th</sup> International Airborne Remote Sensing Conf. and Exhibition/21<sup>st</sup> Canadian Symposium on Remote Sensing*, Ottawa, Ontario, Canada, pp. 21-24, June 1999.
- [7] A. Plaza, P. Martinez, R. Perez and J. Plaza, "Spatial/spectral endmember extraction by multidimensional morphological

- operations," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 40, no 9, pp. 2025-2041, September 2002.
- [8] M.D. Craig, "Minimum-volume transforms for remotely sensed data," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 32, no 3, pp. 542-552, May 1994.
- [9] J.W. Boardman, "Geometric mixture analysis of imaging spectrometry data," *Proc. Int. Geoscience and Remote Sensing Symp.*, vol. 4, pp. 2369-2371, 1994.
- [10] A. Ifarraguerri and C.-I Chang, "Hyperspectral image segmentation with convex cones," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 37, no 2, pp. 756-770, March 1999.
- [11] J.M.P. Nascimento and J.M. Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Trans. Geoscience and Remote Sensing*, vol. 43, no 4, pp. 898-910, April 2005.
- [12] A. Hyvarinen, J. Karhunen and E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.
- [13] J. Bayliss, J.A. Gualtieri and R.F. Crompt, "Analyzing hyperspectral data with independent component analysis," *Proceedings of SPIE*, vol. 3240, pp. 133-143, 1997.
- [14] C.-I Chang, S.S. Chiang, J.A. Smith and I.W. Ginsberg, "Linear spectral random mixture analysis for hyperspectral imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 40, no. 2, pp. 375-392, February 2002.
- [15] X. Zhang and C.H. Chen, "New independent component analysis method using higher order statistics with applications to remote sensing images," *Optical Engineering*, vol. 41, pp. 1717-1728, July 2002.
- [16] J.M.P. Nascimento and J.M.B. Dias, "Does an independent component analysis play a role in unmixing hyperspectral data," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 43, no 1, pp.175-187, January 2005.
- [17] J. Wang and C.-I Chang, "Dimensionality reduction by independent component analysis for hyperspectral image analysis," *IEEE International Geoscience and Remote Sensing Symposium*, Seoul, Korea, July 25-29, 2005 and *IEEE Trans. on Geoscience and Remote Sensing*, (to appear)
- [18] C.-I Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 42, no. 3, pp. 608-619, March 2004.
- [19] A. Plaza and C.-I Chang, "Fast implementation of pixel purity index algorithm," *Conference on Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI*, SPIE Symposium on Defense and Security, SPIE Vol. 5806, Orlando, Florida, 28 March 28-1 April, 2005.
- [20] A. Plaza and C.-I Chang, "An improved N-FINDR algorithm in implementation," *Conference on Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI*, SPIE Symposium on Defense and Security, SPIE Vol. 5806, Orlando, Florida, 28 March 28-1 April, 2005
- [21] C.-I Chang and A. Plaza, "Fast iterative algorithm for implementation of pixel purity index," *IEEE Trans. on Geoscience and Remote Sensing Letters*.
- [22] Adams, J.B., M.O. Smith, and A.R. Gillespie, "Image spectroscopy: interpretation based on spectral mixture analysis," *Remote Geochemical Analysis: Elemental and Mineralogical Composition*, edited by C.M. Pieters and P.A. Englert, Cambridge University Press, pp. 145-166, 1993.
- [23] D. Heinz and C.-I Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 39, no. 3, pp. 529-545, March 2001.
- [24] A. Hyvarinen and E. Oja, "A fast fixed-point for independent component analysis," *Neural Computation*, vol. 9, no. 7, pp. 1483-1492, 1997.
- [25] H. Ren and C.-I Chang, "Automatic spectral target recognition in hyperspectral imagery," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1232-1249, October 2003.
- [26] <http://speclab.cr.usgs.gov/cuprite.html>
- [27] A.A. Green, M. Berman, P. Switzer and M.D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. on Geosciences and Remote Sensing*, vol. 26, pp. 65-74, 1988.
- [28] C.-I Chang, W. Liu and C.-C. Chang, "Discrimination and identification for subpixel targets in hyperspectral imagery," *IEEE International Conference on Image Processing*, Singapore, Oct. 24-27, 2004
- [29] C.-I Chang, H. Ren, C.-C. Chang, J.O. Jensen and F. D'Amico, "Estimation of subpixel target size for remotely sensed imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 42, no. 6, pp. 1309-1320, June 2004.
- [30] J.C. Harsanyi, W. Farrand and C.-I Chang, "Detection of subpixel spectral signatures in hyperspectral image sequences," *Annual Meeting, Proceedings of American Society of Photogrammetry & Remote Sensing*, Reno, pp. 236-247, 1994.

**Jing Wang** (S'03) received the B.S. degree in electrical engineering and M.S degree in computer engineering from the Beijing University of Post and Telecommunications in 1998 and 2001. She also received the M.S degree in electrical engineering from the University of Maryland Baltimore County, Baltimore (UMBC), in 2005., where she is currently a Ph.D. candidate. She received a UMBC dissertation fellowship in 2005 and is currently a research assistant in the Remote Sensing, Signal and Image Processing Laboratory, UMBC. Her research interests include signal and image processing, pattern recognition and data compression.

**Chein-I Chang** (S'81-M'87-SM'92) received his B.S. degree from Soochow University, Taipei, Taiwan, M.S. degree from the Institute of Mathematics at National Tsing Hua University, Hsinchu, Taiwan and M.A. degree from the State University of New York at Stony Brook, all in mathematics. He also received his M.S., M.S.E.E. degrees from the University of Illinois at Urbana-Champaign and Ph.D. degree in electrical engineering from the University of Maryland, College Park, in 1987. Dr. Chang has been with the University of Maryland Baltimore County (UMBC) since 1987, as a Visiting Assistant Professor from January 1987 to August 1987, Assistant Professor from 1987 to 1993, Associate Professor from 1993-2001 and Professor in the Department of Computer Science and Electrical Engineering since 2001. He was a Visiting Research Specialist in the Institute of Information Engineering at the National Cheng Kung University, Tainan, Taiwan, from 1994-1995. He received an NRC (National Research Council) senior research associateship award from 2002-2003 sponsored by the US Army Soldier and Biological Chemical Command, Edgewood Chemical and Biological Center, Aberdeen Proving Ground, Maryland. He is currently a distinguished lecturer at the National Chung Hsing University sponsored by the Department of Education and National Science Council in Taiwan from 2005-2006. He has three patents and several patents pending. He is on the editorial board on the *Journal of High Speed Networks* and was the guest editor of a special issue of the same journal on telemedicine and applications. His research interests include multispectral/hyperspectral image processing, automatic target recognition, medical imaging, information theory and coding, signal detection and estimation and neural networks. He has authored a book, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification* published by Kluwer Academic Publishers. Dr. Chang is an Associate Editor in the area of hyperspectral signal processing for IEEE Transaction on Geoscience and Remote Sensing and a Fellow of SPIE and a member of Phi Kappa Phi and Eta Kappa Nu.