A New Spectral Unmixing Approach based on ICA and Spectral Angle Measure for Hyperspectral Images

Zouhaier BEN RABAH*, Imed Riadh FARAH*, Basel SOLAIMAN** & Henda BEN GHZALA*

* RIADI GDL laboratory, School of Engineering and Computer Science, Manouba, Tunisia
zouhairbr@yahoo.fr
riadh.farah@ensi.rnu.tn
henda.benghzala@cck.rnu.tn

** ENST-Bretagne, ITI laboratory, Technopôle Brest Iroise CS 83818, 29238 Brest Cedex France
Basel.Solaiman@telecom-bretagne.eu

Abstract: The basic task underlying many hyperspectral image (HSI) applications is to identify different materials based on their reflectance spectrum. Indeed, a pixel in remotely sensed hyperspectral imagery is typically a mixture of multiple electromagnetic radiances coming from various ground cover materials commonly called ‘Endmember’, where the mixture coefficients correspond to the abundances of the constituting materials. Previous ICA-based abundance quantification methods hold in account one essential parameter, that’s the illumination of the observed pixels. As a consequence these methods are sensitive to the noise caused by this illumination. Through this study, we propose a new approach called Independent Component Analysis (ICA) and Spectral Angle Measure (SAM) based Spectral Unmixing (ICA-SAM-SU) which is insensitive to this noise. This approach use SAM technique not only to evaluate the spectral similarity between two spectral signatures but for abundance quantification. The major contribution of this paper is that our new approach allows first estimating abundance quantification only with SAM and not with pixel radiometric value of observed pixels. As a result, error factor was minimized. Second, physical constraints are respected. The experiment was conducted using simulated and ASTER image in order to evaluate and to validate our approach.

Key words: Abundance quantification, Endmember extraction, hyperspectral image, spectral angle measure.

INTRODUCTION

Hyperspectral imagery (HSI) provides valuable information helping us in classification tasks. This type of image use many contiguous bands of high spectral resolution covering the visible, near-infrared spectral bands (from 300 nm to 2500 nm) for identification of the composition of various materials called “Endmembers” (EM) and has been employed for various applications of detection, extraction and mapping of materials.

The signal captured by hyperspectral sensor at a given band and from a given pixel is a mixture of the energies scattered by the constituent substances located in the respective pixel spatial coverage [ADA 86]. Dealing with spectral signature compositions in mixed pixels EM extraction and abundance quantification are one of the most challenging problems in HSI analysis.

Using the spectral unmixing technique, mixing proportions of each ground cover material in a certain area could be obtained. Based on the abundances of a particular EM, correspondent map is produced. These abundance maps of pure material (EM) are 2D images whose pixel values, ranged between 0 and 1, indicate the surfacic proportion of this material spectrum in each vector pixel.

In order to obtain this abundance maps, we use ICA technique [HYV 97] which is a blind source separation (BSS) method based on the hypothesis that the EMs are statistically independent. This technique allows finding the linear transform decomposing HSI into 2D -images- the independent components (ICs). In [STR 80], G. Strang proposes the Least Square Error (LSE) method as an algorithm to solve the generalized linear spectral unmixing problem. In [JIN 06a], Wang and Chang propose an ICA-based Abundance Quantification Algorithm (ICA-AQA) that enables extraction EM and estimation of the EM abundances simultaneously. Nevertheless, this estimator is not adapted to some practical purposes.
which will be made explicit. In [ALE 07] Alexis H. and Mireille G. propose a method of hyperspectral image analysis. Post-processing step, applied after anomaly detection and extraction with ICA, enables estimation of anomaly abundances from the independent components with non negative abundances values. Certainly, all these methods shared the same objective but the same drawback too. The first is the need to solve the linear unmixing problem in the remote sensing field. The second is their sensitivity to the radiometric value of the observed pixels. To overcome this drawback, we propose a method called ICA and Spectral Angle Measure -based Spectral Unmixing (ICA-SAM-SU) that use ICA to extract EM and spectral angle measures to transform the equation system for the linear spectral unmixing problem, to a quadratic one (determined system) having unique solution independently of the radiometric values of the observed pixels. This method derives from a generalized mathematical relationship between the abundance estimation and the Spectral Angle Measure.

The paper is organized as follows. In Section 2, Linear Mixture Model (LMM) is presented as well as BSS method. In section 3, we present the traditional Spectral Angle Mapper (SAM). In Section 4, we define the ICA-SAM-SU and we discuss the methodology of our approach. Experimental is carried out in Section 5 emphasizes the robustness of our abundance estimator to ICA-AQA. Accuracies of the tow estimators are compared through synthetic image and ASTER image to quantify abundance. The last section concludes the paper.

1 Blind Sources Separation in HSI

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 the EM extraction seems natural and justifiable [JIN 06a].

1.1. Blind Sources Separation

The BSS technique has been applied in many fields like communication, speech signal processing, image processing, biology, telecommunication, econometrics, and recently in remote sensing [HYV 97], [STR 80].The ability of the use of BSS method was explored in hyperspectral remote sensing images to restore the independent components which can be used to estimate abundance quantification. The BSS problem consists of retrieving unknown sources when observing a mixture of them. The sources are assumed non-Gaussian signals and statistically independent one from another [MAN 00]. The investigation of these unknown sources is only based on their mixtures. If we admit that the mixing is linear, the model of BSS can be expressed by:

$$X = AS + N$$  \hspace{1cm} (1)

Where X is an n x p observed image matrix and each of its rows determines the reflectance of the observed image according to a given spectral band. S is an m x p source images matrix, each of its rows determines the reflectance of one source image. A is n x m mixing matrix; each of its columns is called the directional vector associated to the corresponding source. N is defined as an n x p matrix realized from a spatially additive white Gaussian noise considered as negligible [CAR 93] [LEU 07].

The goal of all BSS methods is to solve (Eq. 2), in which A and S are the unknown components. In fact, sources separation can be obtained by optimizing a contrast function that can be based on: entropy, mutual independency, higher order statistics, etc. [JUT 94] [CAR 93]. Many approximate methods [JUT 91] [HYV 97][CHO 00] are proposed in order to solve this equation (Eq.1). We selected Fast-ICA-2D algorithm to generate the source images or CIs using a matrix X representing the set images. This algorithm was adapted from the 1D field to the 2D one [FAR 03].

1.2. Linear Mixture Model

Linear mixture models (LMM) have been used extensively for characterizing spectral data. One major reason why the LMM has been broadly accepted for the spectral unmixing analysis is that the linear mixture assumption allows many mature mathematical skills and algorithms, to be easily applied to the spectral unmixing problem. The LMM assumes that each ground cover material only produces a single radiance, and the mixed spectrum is a linear combination of ground cover radiance spectra. So the spectrum for a given pixel is a linear combination of the EM spectra.

$$x_k = \sum_{i=1}^{P} a_{ik} m_{ik} + \varepsilon$$  \hspace{1cm} (2)

where $x_k$ is the i-th band of a given pixel, $m_{ik}$ is the k-th EM of the i-th band, $a_{ik}$ is the mixing proportions for the j-th pixel from the k-th EM, and $\varepsilon$ is additive Gaussian random error. It takes into account the sensor noise, the spectral variability, the atmosphere fluctuations and other model inadequacies. Indeed, the pixel compositions are assumed to be in percentages. Since each observed spectral signal is the result of an actual mixing process, the driving abundances must obey two rather common-sense constraints. First, all abundances must be non-negative (eq.3). Second, the sum of abundances for a given pixel must be unity (eq. 4) [CHA 03].

$$\forall h = 1,\ldots,F \quad u_k \geq 0$$  \hspace{1cm} (3)

$$\sum_{k=1}^{F} a_{ik} = 1$$  \hspace{1cm} (4)

Assuming a linear mixture model the image pixels are represented by a simplex space. A simplex space is the simplest geometric shape that can enclose a space
of a given dimension, such as a line (one dimension) and a triangle (two dimensions). The simplex can be defined by the vertices, which are called EM spectra. These correspond to pixels which contain the purest spectrum in the dataset. Ideally, these are unmixed spectra of pure materials, to facilitate automated methods for EM determination. Most automated approaches in retrieve EM are essentially methods for retrieving the purest pixels in a scene [CHA 03].

The unmixing technique solves a set of n linear equations for each pixel, where n is the number of bands in the image [ADA 86]. The unknown variables in these equations are the fractions of each EM in the pixel. To be able to solve the linear equations for the unknown pixel fractions it is necessary to have more equations than unknowns, which means that we need more bands than EM materials. With hyperspectral data this is almost always true.

2 Spectral Angle Mapper (SAM)

In this section, we present one spectral angle measure, commonly used to measure similarity between any two spectra pixel vectors. Assume that

\[ s_i = (s_{i1}, s_{i2}, ..., s_{iL})^T \]

and

\[ s_j = (s_{j1}, s_{j2}, ..., s_{jL})^T \]

are two pixel spectrums. The superscript T stands for a vector or matrix transpose.

The Spectral Angle Mapper (SAM) is successfully used to study soil degradation in South of Spain [MAR 01] and in North of Morocco [CHI 05]. This approach is implemented to study arid zones degradation given satisfactory results [SOH 99][YAN 99].

This method (SAM) [CLA 92] measures spectral similarity between the spectral signatures \( s_i \) and \( s_j \) in n-dimensions. The angle \( \alpha \) between the spectra treated as vectors in n-space is the "spectral angle", this is illustrated in 2 Dimensions.

This method assumes that the data have been reduced to apparent reflectance and uses only the "direction" of the spectra, and not their "length". Thus, this spectral angle insensitive to changes in pixel illumination because increasing or decreasing illumination doesn’t change the direction of the vector, only its magnitude (i.e., a darker pixel will plot along the same vector, but closer to the origin).

\[
\text{SAM}(s_i, s_j) = \cos^{-1} \left( \frac{\langle s_i, s_j \rangle}{\|s_i\|\|s_j\|} \right)
\]  

Where \( \langle s_i, s_j \rangle = \sum_{l=1}^{L} s_{il} s_{jl} \), \( \|s_i\| = \left( \sum_{l=1}^{L} s_{il}^2 \right)^{1/2} \) and \( \|s_j\| = \left( \sum_{l=1}^{L} s_{jl}^2 \right)^{1/2} \).

3 ICA and Spectral Angle Measure-based Spectral Unmixing

Spectral unmixing is a quantitative analysis procedure used to recognize constituent ground cover materials or EM and to obtain their mixing proportions (or abundances) from a mixed pixel. In order to reliably estimate these two variables, with respect to physical conditions (eq.3) and (eq.4), we present in this section the proposed unmixing method (ICA-SAM-SU) the proposed approach is presented in fig. 1.

![Image of proposed methodology](image)

**Figure 1. Proposed methodology for abundance quantification.**

3.1. Endmember extraction

To allow analysis of the data from the more useful perspective of reflectance, radiance measurements from the original real data were corrected.

We apply to these data linear spectral unmixing using one traditional ICA-based EM Extraction followed by SAM-based Abundance Quantification processing flow [JIN 06a].

The idea is to takes advantages of the FastICA-2D generated ICs that separate all extracted EM pixels in individual components. Unfortunately, in a 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 a 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 [JIN 06b] are used for this purpose. We propose to select the high-order statistics-based IC prioritization algorithm (HOS-ICPA) one which uses high-order statistics to affect a priority score to each IC. Following this Virtual Dimensionality (VD) [CHA 04] algorithm was performed to determine the number of ICs needed to be selected and which is the number of unknown signal sources in the HSI, let p this number.
For each of the selected $p$ FastICA-generated IC images, find a pixel with maximum absolute value, which is supposed to be an EM.

The same selected $p$ prioritized ICs was then used for abundance quantification for all image pixels.

In order to labeling each EM, SAM technique was performed by matching each subpixel component signature with spectral library one produced by Spectroscopic measure campaign and USGS website.

### 3.2. Abundance quantification

The main goal of the ICA-SAM-SU is to extract all EMs from the hyperspectral image and further generate one abundance map for each extracted EM spectra.

A new method can be developed in the following to use the Spectral Angle Measure for abundance quantification.

Based on the LMM and the SAM methods, a generalized mathematical relationship between the abundance estimation and the Spectral Angle Measure was derived.

The objective of the ICA-SAM-SU method is to obtain an optimum and unique estimate value for the abundances $\alpha_k$, for a given mixed pixel spectrum $x_i$ and EM spectra $m_k$. Starting with the linear mixture model of (eq.2) Where $\epsilon$ is defined as an additive white Gaussian noise considered as negligible [CAR 93] [LEU 07]. So equation (eq.2) can be rewritten to,  

$$x_i = \sum_{k=1}^{P} \alpha_k m_k$$  

(eq.7) can be rewritten as  

$$\sum_{k=1}^{P} \alpha_k M_k \bar{u}_k = \bar{v}_x$$  

$$\sum_{k=1}^{P} \alpha_k M_k \bar{v}_k = \bar{v}$$  

Let  

$$\alpha_k M_k \bar{u}_k = \bar{v}$$

In matrix form (eq.8) can be rewritten as a scalar product of $U$ and $A$, denoted by $<U,A>$ or

$$UU^T A = V$$  

With $U=[u_1, u_2, ..., u_p]$,  

$$A^T = [a_1, a_2, ..., a_p]$$  

$$V^T = [v_1, v_2, ..., v_l]$$  

$L$ is band number, and $u_i$ is the unit vector of $i$th EM  

$$u_i = [u_{i1}, u_{i2}, ..., u_{il}]$$

Multiplying (eq.9) on the left hand by $U$ we obtain

$$UU^T A = UV$$

Let  

$$\Gamma = UU^T$$

then (eq.10) can be rewritten as  

$$\Gamma A = UV$$  

$$\Gamma = UU^T = \begin{bmatrix} u_1^T & u_2^T & \cdots & u_p^T \end{bmatrix} \begin{bmatrix} u_1 & u_2 & \cdots & u_p \end{bmatrix}$$

$$\Gamma = \begin{pmatrix} u_1^Tu_1 & u_1^Tu_2 & \cdots & u_1^Tu_p \\ u_2^Tu_1 & u_2^Tu_2 & \cdots & u_2^Tu_p \\ \vdots & \vdots & \ddots & \vdots \\ u_p^Tu_1 & u_p^Tu_2 & \cdots & u_p^Tu_p \end{pmatrix}$$

Other definition of scalar product is  

$$u_i^T u_j = \cos(\theta_{ij}) \frac{||u_i||}{||u_j||}$$

Since $u_i$ is a unit vector then
\[ u_i^T u_j = \cos(\beta_{ij}); \]

\[ \beta_{ij} \] represent the spectral angle between \( j \) th EM vector and the \( i \) th one (fig. 2) and when \( i=j \) with \( \cos(\beta_{ij}) = 1; \)

\[ u_i^T u_j = \begin{pmatrix} 1 & \cos \beta_{12} & \cdots & \cos \beta_{1N} \\ \cos \beta_{12} & 1 & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ \cos \beta_{1N} & \cdots & \cdots & 1 \end{pmatrix} \]

**Figure 2.** Concept of the Spectral Angle Measure in a two dimensions space.

\( \Gamma \) is square matrix, symmetric and positive definite (since vectors constituting the matrix \( U \) are independent) then it is an invertible matrix.

\( A \) can be written as

\[ A = \Gamma^{-2} UV \]  \hspace{1cm} (11)

\( \Gamma^{-2} U \) is calculated using only the sources and can be computed once for all image analysis.

Then, the image is analyzed pixel by pixel, for each pixel we calculate \( A \)

\[ a_h = \alpha_h \frac{M_h}{X}; \]

\[ a_h = \alpha_h \frac{M_h}{X}; \quad \alpha = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_P \end{pmatrix} \cdot X \]  \hspace{1cm} (12)

Since the \( \alpha_h \) are Weighted coefficient of \( M_h \) its was known in a coefficient meadows and (eq.12) can be rewritten as

\[ \alpha_h = \frac{a_h}{M_h} \]  \hspace{1cm} (13)

In order to respect (eq.4) the \( \alpha_h \) must be normalized to one by

Once that the EMs and their proportions are calculated for each pixel, these data can be used to correct the observed pixels values.

### 4 Experiments, results and discussion

#### 4.1 Simulation image

In our first experiment we use mineral data from the United States Geological Survey (USGS) website to build artificial mixtures for evaluating our ICA-SAM-SU method. Five target EM where chosen (Alunite (A), Buddingtonite (B), Calcite (C), Kaolinite (K), and muscovite (M)). And their spectral signatures were plotted in Fig. 3.

**Figure 3.** five mineral spectra

Using these signatures, a synthetic image scene with size of 64 x 64 pixels was simulated according to Fig. 4 where the objects have been set into lines and columns.

**Figure 4.** Synthetic image with specified abundance fractions

The image background was made up of 50% Alunite and 50% Kaolinite with an additive 30 dB Gaussian noise, and three other minerals Buddingtonite, Calcite, and Muscovite were used to simulate 24 panels of different abundance fractions with those panels in first, second, and third columns.
specified by Buddingtonite, Calcite and Muscovite, respectively. These 24 panels were then implanted into the image background in a way that the corresponding background pixels were replaced by the pixels in the 24 panels. Except three panels in the first row labeled by 100%, which contained four pure mineral pixels, all other panels were single-pixel panels with various abundance fractions as specified in Fig. 4. It is worth noting that the panel pixels with abundance fractions less than 100 were mixtures of the panel signature with the background. For example, the panel pixel labeled by 80% in the first column is a mixture of 80% Buddingtonite with 20% background signature (BKG).

Table 1. Abundance estimation of the Buddingtonite EM

<table>
<thead>
<tr>
<th>Ground truth</th>
<th>ICA-AQA</th>
<th>ICA-SAM-SU</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimation</td>
<td>Error EM</td>
</tr>
<tr>
<td>100</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>80</td>
<td>80.24</td>
<td>0.24</td>
</tr>
<tr>
<td>60</td>
<td>60.46</td>
<td>0.46</td>
</tr>
<tr>
<td>40</td>
<td>41.10</td>
<td>1.10</td>
</tr>
<tr>
<td>30</td>
<td>30.78</td>
<td>0.78</td>
</tr>
<tr>
<td>20</td>
<td>20.91</td>
<td>0.91</td>
</tr>
<tr>
<td>10</td>
<td>10.96</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Figure 5. Abundance quantification for BKG and each Endmember using ICA-AQA

Figure 6. Abundance quantification for BKG and each Endmember using ICA-SAM-SU

As explained in previous Section, obtained abundance maps are images, whose grey level values accuracy depend of the method robustness and the respect to the physical constraints (eq.3) and (eq.4). In Fig.5, abundance maps of the three objects, which are EM, are estimated with ICA-AQA. On the other hand those who are presented in Fig.6 are estimated with ICA-SAM-SU method. The panels in the left column represents the IC corresponding to the Buddingtonite mineral (b), the middle column represents the IC corresponding to the Calcite mineral (c) and the right column represents the IC corresponding to the Muscovite mineral (m).

We notice that in any case, ICA-AQA or ICA-SAM-SU return abundances near to 1 if the EM is expected in the pixel and approximately null in the background (Fig.5,6 , line 1,2). If we focus on the first EM (Buddingtonite) abundances whose results seem accurate for both estimators with a little improvement for the proposed method, we can draw up a table (Table.1) comparing estimated abundances with the ground-truth ones.

For both methods, we notice that the lower the EM abundance, the higher the relative error of estimation. This is caused by the signal to noise ratio that decreases relatively with the EM abundance.

4.2. Validation with ASTER image

The data used in the following experiments were directly extracted from the ASTER image scene of size 225X207 pixels shown in Fig. 7. The ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) sensor is a 9 channel imaging spectrometer covering the 0.4-2.5 µm spectral range.
ASTER images is not a typical data source for validate performance of the proposed method to hyperspectral image because its band numbers are small. But for lack of availability of popular hyperspectral image we adopt them.

In order to show the effectiveness of our proposed approach, we apply first a classic approach to the ASTER image subset. So we generate different ICs using the FAST-ICA2D then the VD in order to estimate the number of EM. As a result this number is estimated to 5. A priority score is then calculates for each IC and the 5 EMs was extracted from HSI labeled using signature spectral data base. The next steps consist to launch first the classical ICA-AQA abundance quantification method then the proposed ICA-SAM-SU method. The fig. 8 shows their results (a) ICA-AQA; (b) ICA-SAM-SU.

According to these results presented in fig. 8 and summarized in table 2, the abundance quantification defined by ICA-SAM-SU was shown to perform significantly more efficient and more objective than did the commonly used pixel level-based ICA-AQA in abundance quantification of subpixels and mixed pixels for real hyperspectral images.

Table 2. Result of the abundance quantification sum given by the two methods for the same pixel positions.

<table>
<thead>
<tr>
<th>Pixel position</th>
<th>X=18</th>
<th>Y=31</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methods</td>
<td>ICA-AQA</td>
<td>ICA-SAM-SU</td>
</tr>
<tr>
<td>% in the 1st EM</td>
<td>73</td>
<td>81</td>
</tr>
<tr>
<td>% in the 2nd EM</td>
<td>33</td>
<td>0</td>
</tr>
<tr>
<td>% in the 3rd EM</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>% in the 4th EM</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>% in the 5th EM</td>
<td>30</td>
<td>9</td>
</tr>
<tr>
<td>Rate sum</td>
<td>152</td>
<td>90</td>
</tr>
</tbody>
</table>

Indeed, comparing results of abundance quantification relative to the same Oasis EM pixels given by fig. 8 (a) and fig. 8 (b), we note that the abundance of the proposed approach 8 (b) is more efficient then the ICA-AQA one 8 (a) and more objective since the sum-to-one constraint (eq.4) and non negativity (eq.3) are respected. This is due to the insensitivity of the Spectral Angle Measure to the number of spectral bands and to signal pixel amplitude. As a consequence this reduces signal noise effect, eliminating waste and maximizing value-added work.
5 Conclusion

This paper presents an ICA-SAM-SU method for abundance quantification that contributing to hyperspectral image analysis. Linear unmixing method and spectral angle measure technique are combined by means of mathematical theoretics to generate abundance maps.

It is a processing which allows EM abundance estimation using generalized mathematical relationship between the abundance estimation and the Spectral Angle Measure. First, this new method was performed using synthetic image and was compared in term of abundance quantification rate, to one classical method. Second, it was performed using real image.

The comparison of the results obtained using the two different approaches outlined shows that the ICA-SAM-SU approach provides more accurate results compared to the results obtained with ICA-AQA.

The results obtained through synthetic and real-world HSI experiments are very promising so we obtain an optimum and unique estimate value for the abundances quantification. In addition the unity and the no negativity constraints were respected and observed pixel value can be corrected using only the sources and abundance values.

REFERENCES


