INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X

Impact factor: 4.295

(Volume 4, Issue 3)

Available online at: www.ijariit.com

Classification of hyperspectral images using PPI and LMM

Pavithra Sukumar
pavithrasukumar6@gmail.com
Marian Engineering College,
Thiruvananthapuram, Kerala

Sreena V G
sreenavg@gmail.com
Marian Engineering College,
Thiruvananthapuram, Kerala

ABSTRACT

Hyperspectral images are the treasure of information since it contains hundreds of spectral bands. Classification of Hyperspectral images is the process of identifying the components in each pixel. For this purpose, the pure and mixed pixels of the image should be identified and the endmember signatures and components are identified with the help of spectral libraries. In this paper, it is attempted to identify the minerals in the hyperspectral data 'Cuprite' that covers the Cuprite mines in Las Vegas, Naveda, the U.S. The pure pixels are identified by using PPI algorithm and their endmember signatures are obtained. The abundance maps of mixed pixels are obtained by using LMM. The total variation based regularization and joint sparsity of abundance maps are exploited in this paper.

Keywords: *HSI (Hyperspectral Image), Spectral reflectance curve, Pure pixel, Mixed pixel, Endmember, PPI (Pixel Purity Index), LMM (Linear Mixing Model).*

1. INTRODUCTION

There are many things on earth which are still invisible to human beings even though the technology has grown up to a great extend. One of the reasons for this inability is that human eyes are sensitive only to the visible region of the electro- magnetic spectrum.

Human eyes can see the color of visible light in mostly three frequencies – red, green and blue. Spectral imaging divides the spectrum into many bands. Hyperspectral imaging combines the merits of digital imaging and spectroscopy. The goal of hyperspectral imaging is to obtain the spectra of each pixel in the image of a scene with the purpose of finding and identifying minerals or detecting processes.

A pixel in a satellite image corresponds to an extensive area on earth. This spatial region constituting that pixel may be covered by a single object or multiple objects. A pure pixel contains only one component in it while the mixed pixel consists of more than one component in it. The term fractional abundance indicates the percentage of a particular endmember in a pixel. Thus the abundance map shows the distribution of a particular endmember over a region. The pure pixels have fractional abundance 1 whereas mixed pixels have fractional abundance between 0 and 1.

The components of these pixels are identified by plotting the spectral reflectance curve. Reflectance is the ratio of the reflected energy to the incident energy as a function of wavelength. Representative spectral reflectance curves for several common earth surface materials over the visible light to reflected IR spectral range. Reflectance is a unitless quantity that ranges in value from 0 to 1, or it can be expressed as a percentage. These values are either measured directly or derived from measurements of light reflected from a standard reference material with known spectral reflectance.

The overall spectral reflectance curve and the position and strength of absorption bands in many cases can be used to identify and discriminate different materials. For example, vegetation has a higher reflectance in the near infrared range and lower reflectance of red light than soils. Several libraries of reflectance spectra of natural and man-made materials are available for public use. These libraries provide a source of reference spectra that can aid the interpretation of hyperspectral and multispectral images. ASTER Spectral Library, USGS Spectral Library are some of such spectral libraries. The total number of endmembers available from these spectral libraries are huge, but only a few of these endmembers are present in a given hyperspectral image. When the size of the ground resolution cell is large, it is more likely that more than one material contributes to an individual spectrum measured by the

sensor. The result is a composite or mixed spectrum, and the pure spectra that contribute to the mixture are called endmember spectra.

This paper attempts to analyze the performance of PPI algorithm on the hyperspectral data, Cuprite to extract endmembers and there by identifying the pure pixels of minerals in the image scene and use it for estimating the fractional abundance of sub-pixels existent within each pixel. The further classification of the data is done by LMM.

2. LITERATURE SURVEY

The journal [1] attempts to study the performance of PPI and NFINDER algorithms. Also, LMM is also used in this paper since this model has gained considerable significance in deriving mixed pixel components and generates fraction images of each ground cover class. Thus the endmembers are identified by two target detection algorithms that have been used for unmixing the hyperspectral data and calculate the fractional abundance of pure species in the study area. This study enables a time-saving species identification mechanism.

The paper [8] propose a method that neither depends on the availability of pure pixels in the original HSI nor on the ability of an endmember extraction algorithm to identify such endmembers. It aims at estimating the fractional abundance of pure spectral signatures in each mixed pixel using LMM. Also, this paper analyzes the influence of the internal characteristics of the spectral libraries on the accuracy of sparse unmixing algorithms. Two relevant indicators such as mutual coherence of the library and the spectral dissimilarity of the signatures present in the library are also taken into account.

The work [7] presents a generalized version of LMM to handle the natural variation that presents in the real world hyperspectral imagery. According to this model, a bimodal distribution in the histogram with peaks at 0 and 1, with a relatively small number of pixels containing abundance between 0 and 1(mixed pixels).

The journal [6] propose a method to make sparse unmixing more flexible for all types of HIS unmixing. This method, called Robust Collaborative Sparse Regression (RCSR) simultaneously takes the collaborative sparse property of the abundance and sparsely distributed additive property.

The paper [2] attempts to use PPI in an elegant way to avoid some issues of PPI algorithm. RPPI considers the results produced by PPI using a set of initial random vectors as skewers as a realization of a random algorithm. The number of skewers is selected on a trial-and-error basis. This experiment shows that RPPI can be competitive as the PPI.

In journal [3], the multidimensional extension of PPI algorithm is focused. Here the relation between the endmember simplex and the convex hull of the data is presented. Pixels having non-zero abundances re considered to be in the interior of the hull. Pixels with at least one abundance equal to zero can lie on the convex hull.

The work [5], attempts to do the unmixing process by using total variation based regularization and joint sparsity of abundance maps. Also, the Gaussian noise and Sparse noise are also removed.

3. METHODOLOGY OF THE STUDY

3.1 Data Acquisition

The hyperspectral data used in this paper is 'Cuprite' which is the most benchmark data set for hyperspectral research that was collected over the Cuprite mining site, Naveda in Las Vegas, the US in 1997. There are 224 bands, ranging from 370 nm to 2480 nm with size 512 X 614. The image was acquired from the website https://speclab.cr.usgs.gov /cuprite.html. The dataset consists of 12 endmembers such as Alunite, Buddingtonite, Andridite, Pyrope, Nontronite, Kaolinite1, Kaolinite2, Muscovite, Dumortierite, Montmorillonite, Sphene, and Chalcedony.

3.2 End Member Detection

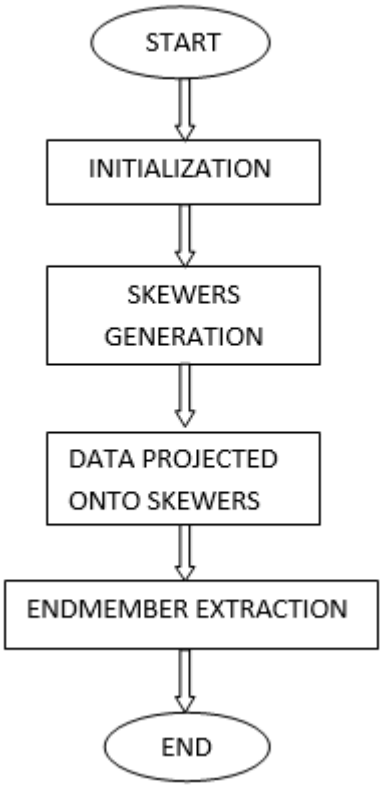
End member detection is done to find the spectral signatures exist in the hyperspectral image. Algorithms such as PPI and N-FINDER have been applied

Pixel Purity Index:

First of all, PPI requires prior knowledge about the number of skewers, NSK. This value is selected empirically based on a trial and error basis. Usually, the larger number of skewers NSk is, the better the obtained results become since we have generated more samples. Typically it is chosen on the order 10^4. According to the geometry of convexity[3], an end member is considered as a pure signature, should locate at the end points of some skewers with maximum or minimum projection. The endmember extraction can be considered as locating the vertices that span the simplex surrounding the data in spectral space.

The main idea behind PPI algorithm is that the end members should be located at the spatial extremes of data cloud in spectral space. The algorithm counts the number of times that a number has been identified as an extreme and hence associates a score with each pixel. The pixels with highest scores are designated as endmembers.

Let M1 = $(m_1,m_2,\ldots\ldots m_N)$ is a data cube of size (d, N) in a data set of N points in a d dimensional vector space R^d.Here it is considered as the endmembers are present in the dataset, and also the possible preprocessing steps like dimensionality reduction or noise removal are not considered. The algorithm works in several steps.



Flowchart of PPI

(a) **Initialization:**

First a large number NSk and a threshold T_2. Now generate a set of NSk random vectors $\{skewes\}_{i=1}^{NSk}$ called skewers. Then PPI scores are generated which count how many times a pixel is found as an extreme.

(b) **Iteration: (**Loop over the NSk skewers.)

- The data is projected orthogonally on to this skewers.
  tmp = M^T*(skewers)
- Then the indices of the maximum and minimum of these projections are calculated.
- Increase the corresponding PPI scores by one
- Thresholding: The endmember candidate is those points with PPI score larger than T_2.

(c) **Endmember extraction:** A data sample 'o' is selected as an endmember if its PPI count is greater than threshold T_2.

Linear Unmixing

Spectral unmixing has intension in finding the fractional abundance of pure spectral signatures (endmembers) in each mixed pixel. The linear mixing model considers the observed spectrum of a pixel can be expressed as a linear combination of the spectra of the endmembers present in the respective pixel. It is mathematically expressed as

$$z_i = \sum_{j=1}^{q} y_{i,j},\alpha_j + n_i \qquad (1)$$

Where yi is the measured value of the reflectance at spectral band i, mij is the reflectance of the j^{th} endmember and n_i represents the error term of the spectral band i.

The matrix formulation of LMM is:

$$Z = \alpha Y + N \qquad (2)$$

Where, Z is the collected mixtures matrix, with Y is the abundance matrix, α is the endmember signature and N is the additive noise. Thus hyperspectral vectors are approximated by a linear combination of a small number of spectral signatures in the spectral library and the number of columns that are equal to the number of pixels.

All the endmembers present in the real hyperspectral image may not be present in the spectral library. So the abundance sum may not be equal to 1.

$$\min_{A,S} \|Z - \alpha Y - S_p\|^2_F + \gamma_1\|\alpha\|_{2,1} + \gamma_2\|S_p\|_1. \qquad (4)$$

Here, the first term is data fidelity term that is equivalent to minimizing the variance of Gaussian noise $G_n = Z - \alpha Y - S_p$. First regularization term is an $l_{2,1}$-norm minimization term on abundance matrix α which is also called joint-sparse regularization term. This term is based on the observation that in most hyperspectral images, fewer endmembers are present compared to the available endmembers. This observation is mathematically modeled as joint-sparse regularization on matrix A with few nonzero rows, but each nonzero row is allowed to be dense. The second regularization term corresponds to minimizing l_1-norm of sparse noise matrix S. Here, 1_1-norm is minimized due to modeling assumption that sparse noise affects few pixels in the image.

Also one of the most important factors that we can exploit is that most of the natural images are piece wise smooth. Therefore, the abundance maps can be considered as piece wise smooth. The piece wise smoothness can be modeled as total-variation regularization.

$$\min_{A,S} \| Z - \alpha Y - S_p \|_F^2 + \gamma_1 \| \nabla \alpha^T \|_1 + \gamma_2 \| S_p \|_1 \quad (5)$$

Here, ∇ is a 2D total variation operator that applies variation along both vertical and horizontal direction. This operator is applied on α^T because each abundance map is along the rows of α.

For the utilization of the joint sparsity and spatial smoothness of the abundance maps, equation (5) can be expressed as

$$\text{Min}_{A,S} \| Z - \alpha Y - S_p \|_F^2 + \gamma_1 \| \nabla \alpha^T \| +$$

$$\gamma_2 \| \alpha \| + \gamma_3 \| S_p \| \quad (6)$$

$\gamma_1, \gamma_2, \gamma_3$ are the regularisation parameters corresponding to total variation term, joint sparsity term and sparse noise term respectively [5]. The JSTV algorithm is used for solving (5) is described in [5] which help to get the abundance maps.

In this study, LMM has been developed entirely in MATLAB, and the fractional abundance of each mineral is obtained

4. RESULT AND ANALYSIS

MATLAB has been used as a platform for the hyperspectral image processing. The main objective of this study is to find the pure pixels using PPI algorithm and obtaining the components in a mixed pixel using LMM. This will help in identifying the minerals in the Cuprite hyperspectral data. The minerals corresponding to the obtained pure pixels are identified by comparing their reflectance values with the minerals in the spectral library of minerals by calculating the sum of squared distances.

The PPI algorithm has been executed with10000 iterations and could extract 5 distinct pure pixels (fig 1) and their corresponding endmembers (fig 2). A threshold of 12000 has been considered to count the number of times a pixel have been considered pure. The pixels below 12000 count display spectral profiles that are ether similar to the profiles of higher count or do not resemble mineral profiles. Hence they have been discarded.

Here the marked pixels represents pure pixels of the minerals Alunite(A), Buddingtonite(B), Calcite(C), Kaolinite(K), Mascovite(M) as in Fig 1. The pixels other than these pure pixels are mixed pixels. Fig 2 shows the spectral signatures of the minerals.

LMM is used to extract the components of these mixed pixels. For the proper working of the algorithm, the data size is reduced to 128X128X224. Thus the abundance maps of five minerals are obtained as shown in fig 3.

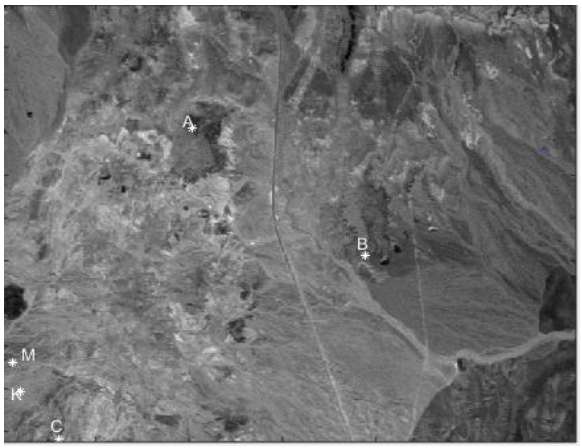


Fig 1

Spatial positions of five pure pixels corresponding to the minerals: A, B, C, K and M

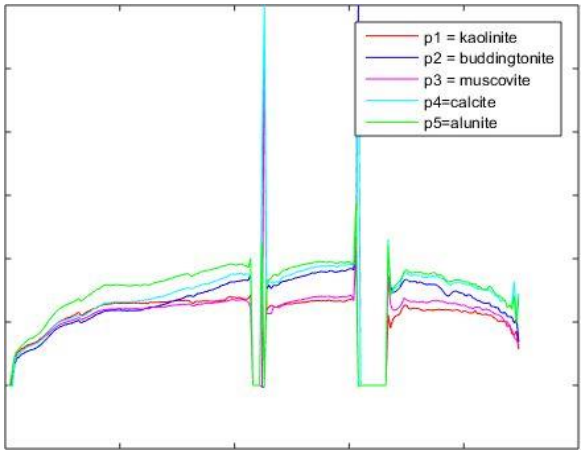
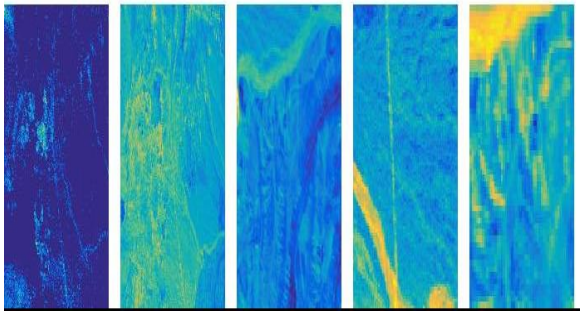


Fig 2

Endmembers identified with PPI algorithm



C K M A B

Fig 3

Abundance maps

C, K, M, A, and B represent the abundance maps of Calcite, Kaolinite, Muscovite, Alunite and Buddingtonite respectively.

5. CONCLUSION

The hyperspectral images have a large amount of information in them. The classification is done to obtain the information about the components in it. The PPI is the standard algorithm to identify the pure pixels and hence extracting the endmembers of the corresponding pixels.

In this paper, the hyperspectral data Cuprite is analyzed and the pure and mixed pixels are identified. The minerals present in these pixels and their endmembers are extracted by using PPI algorithm.

By using the LMM and utilizing the JSTV function, the abundance maps of the five minerals in a pixel are obtained.

6. ACKNOWLEDGEMENT

I also acknowledge my gratitude to all other faculty members of the department of Electronics and Communication Department.

7. REFERENCES

[1] Somdatta Chakravortty and Devdatta Sinha, “Performance of Pure Pixel Extraction Algorithms on Hyperspectral Data for Species-Level Classification of mangroves”,http://ieeexplore.ieee.org/document/7052047/
[2] Chein-I Chang, Chao-Cheng Wu, and Hsian-Min Chen, “Random Pixel Purity Index”, http://ieeexplore.ieee.org/document/5345838/
[3] Rob Heylen and Paul Scheunders, “Multidimensional Pixel Purity Index for Convex
Hull Estimation and Endmember Extraction”, http://ieeexplore.ieee.org/document/6392932/
[4] Chein-I Chang, Chao-Cheng Wu, Eliza Yingzi Du, and Hsian-Min Chen, “Component Analysis-Based Unsupervised Linear Spectral Mixture Analysis for Hyperspectral Imagery”
[5] Hemant Kumar Aggarwal and Angshul Majumdar
“Hyperspectral Unmixing in the Presence of Mixed
Noise Using Joint-Sparsity and Total Variation” http://ieeexplore.ieee.org/document/7414394/

[6] Chang Li, Yong Ma, Yuan Gao, Zhongyuan Wang, and Jiayi Ma, "Sparse Unmixing of Hyperspectral Data based on Robust Linear Mixing Model"
[7] David Gillis, Jeffrey Bowles, Emmett J. Ientilucci and David W. Messinger, "A generalized linear mixing model for hyperspectral imagery"
[8] Marian-Daniel Iordache and Marian-Daniel Iordache "On The Use Of Spectral Libraries To Perform Sparse Unmixing Of Hyperspectral Data"
[9] J. Boardman, F. Kruse, and R. Green. Mapping target signatures via partial unmixing of AVIRIS data. In R. Green, editor, *Summaries of the 5th Annu. JPL Airborne Geoscience Workshop*, Volume 1, pages 23–26, Pasadena, CA, JPL Publ. (1995)
[10] Rogge D. M., Rivard B., Zhang J., Sanchez A., Harris J., and Feng J.: Integration of spatial-spectral information for the improved extraction of endmembers. *Remote Sensing of Environment*, 110:287–303,(2007).
[11]. Mustard J. F. and Sunshine J. M., "Spectral analysis for earth science investigation," in *Remote Sensing for the Earth Sciences*, pp. 509–564, John Wiley & Sons Inc., New
York (1999).
[12] Yang C., Everitt J. H., and Bradford J. M., "Using Multispectral Imagery and Linear Spectral Unmixing Techniques for Estimating Crop Yield Variability," *Trans. ASABE* 50(2), 667–674(2007).
[13] Beck R., *EO-1 User Guide Version 2.3*, p. 74,
The University of Cincinnati, Ohio(2003)
[14] Jensen J. R., *Introductory Digital Image Processing: A Remote Sensing Perspective*, 3rd ed., p. 526, Prentice- Hall, Upper Saddle River(2005).
[15] Maaten L., Postma E., and Herik J., "Dimensionality Reduction: A Comparative Review,"*Tilburg University Technical Report*, TiCC-TR-2009-005(2009).
[16] Du Q.et al., "Variants of an N-FINDR algorithm for endmember extraction," *Proc. SPIE* 7109, 71090G, http://dx.doi.org/10.1117/12.799361(2008).
[17]. R. Hsuang and C. Chang, "Automatic spectral target recognition in hyperspectral imagery," *IEEE Trans. Aerosp. Electron. Syst*. 39(4), 1232–1249,
http://dx.doi.org/10.1109/TAES.2003.1261124(2003)
[18] Berman M., Kiveri H., Lagerstrom R., Ernst A., Donne R., and Huntington J. F.: A Statistical Approach to Identifying Endmembers in Hyperspectral Images. *IEEE Transactions on Geoscience and Remote Sensing*, 42:2085–2095 (2004).
[19] Chang C.-I, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, Kluwer
Academic/Plenum Publishers (2003).
[20] Harsanyi J.C. and Chang C.-I, "Hyperspectral Image Classification and Dimensionality Reduction: An Orthogonal Subspace Projection Approach," *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 32, No. 4, pp. 779-785, (1994)
[21] Keshava N. and Mustard J., "Spectral unmixing," *IEEE Signal Process.* Mag.19 (1), 44–57,
http://dx.doi.org/10.1109/79.974727 (2002). 214

