Signal Processing for Hyperspectral Data

Pramod K. Varshney[a] , Manoj K. Arora[b] , Raghuveer M. Rao[c]

[a]Electrical Engineering and Computer Science Department, Syracuse University, Syracuse, USA
[b]Department of Civil Engineering, IIT Roorkee, ROORKEE, INDIA
[c]Electrical Engineering Department, Rochester Institute of Technology, Rochester, USA

ABSTRACT

Hyperspectral data form a data-cube consisting of images of an object collected at several hundred, closely spaced wavelengths. They have been found to be of significant potential benefit in areas such as remote sensing of the Earth, medicine, and non-destructive evaluation. Effective extraction of information from the hyperspectral data cube presents several signal processing challenges, some of them unique to hyperspectral data. The problems involved range from registration and enhancement to development of statistical signal processing algorithms and models for object detection and classification. The focus of this paper is to provide an overview of select processing and modeling techniques for hyperspectral data.

1. INTRODUCTION

Since the beginning of remote sensing technology, a number of sensors have been developed to observe the Earth's surface [1]. The observed data form an image, a collection of radiance or reflectance of each pixel at the wavelength of interest. Hyperspectral imaging provides the capability to characterize and quantify the Earth's diverse environments in considerable detail using several hundred contiguous wavelength bands typically ranging from the visible to the infrared regions [2]. Many sensors such as the airborne AVIRIS, HYDICE, HYMAP and the spaceborne HYPERION have been developed for gathering hyperspectral signals. The sheer increase in the volume of available hyperspectral data has created the need for the development of signal processing techniques that can automate information extraction. These techniques need to be objective, reproducible, and feasible to implement within available resources [3].

A hyperspectral data-set, referred to as a *hyperspectral image cube,* can be visualized as a 3-D stack, or sequence of 2-D grayscale images each of which is obtained at a specific wavelength. Since an almost continuous spectrum can be generated for a pixel, hyperspectral imaging is also referred to as imaging spectrometry.

The number and variety of potential applications for hyperspectral imaging is enormous [4], which brings challenges in processing of hyperspectral data that are different from multispectral image processing. The majority of the processing algorithms fall into four primitive application-specific tasks [4]: searching for rare pixels in the hyperspectral cube (target detection), finding changes between two scenes (change detection), assigning a label to each pixel (classification), and estimating the fraction of the pixel area covered by each material type.

In the remainder of the paper, we enumerate the steps typically used in hyperspectral data processing and highlight two problems that are representative of demonstrating how hyperspectral data analysis problems can be cast into frameworks based on signal processing models for their solution.

2. HYPERSPECTRAL DATA PROCESSING

The whole process of hyperspectral imaging may be divided into three steps: preprocessing, radiance to reflectance transformation and data analysis. Preprocessing is required for the conversion of raw radiance into at-sensor radiance. This is generally performed by the data acquisition agencies and the user is supplied with the at-sensor radiance data. The processing steps involve operations such as spectral calibration, geometric calibration and geocoding, signal to noise adjustment, de-striping etc. Further, due to topographical and atmospheric effects, various spectral and spatial variations may occur in at-sensor radiance. Therefore, the at-sensor data need to be normalized in the second step for accurate determination of the reflectance values in different bands. A number of atmospheric models and correction methods have been developed to perform this operation. A detailed overview of the first two steps can be found in refs. [5] and [6]

Data analysis is aimed at extracting meaningful information from the hyperspectral images. A limited number of image analysis algorithms were developed earlier to exploit the extensive information contained in hyperspectral signals for the applications stated above. While some have demonstrated significant success, the combination of physical and mathematical modeling that optimally extracts information from hyperspectral signals is still to be determined [7].

1-4244-0469-X/06/$20.00 ©2006 IEEE

3. LINEAR MIXTURE SEPARATION AND INDEPENDENT COMPONENT ANALYSIS

The ground footprint of each hyperspectral pixel can cover a relatively large area. Thus, each pixel can include several different types of cover such as, for example, grass, soil and asphalt. The hyperspectral vector corresponding to each pixel is, therefore, modeled as a linear combination of "endmember" spectra, that is, the spectra of known basic material:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{1}$$

Here $\mathbf{x}$ is the pixel vector, the columns of $\mathbf{A}$ are the endmembers and $\mathbf{s}$ is the mixing vector which provides the proportionate abundance of each endmember component. We require

$$s_i \geq 0 \; and \sum_i s_i = 1 \tag{2}$$

for the components s_i of $\mathbf{s}$.

Given $\mathbf{A}$, one needs to estimate $\mathbf{s}$. Often both $\mathbf{A}$ and $\mathbf{s}$ are unknown, in which case the problem resembles Blind Source Separation (BSS). Each component of $\mathbf{s}$ is a source value, $\mathbf{A}$ is a mixing matrix and $\mathbf{x}$ is the vector of observations. Typically there are many more spectral components than endmembers.

Independent component analysis (ICA) [8] has been applied to the problem of recovering $\mathbf{A}$ and $\mathbf{s}$ from $\mathbf{x}$. One proposed approach [9] models each source as non-Gaussian. In this approach, the observation is first whitened to obtain

$$\mathbf{y} = \tilde{\mathbf{H}}\mathbf{s} \tag{3}$$

with

$$\mathbf{y} = \left(\mathbf{E}\boldsymbol{\Lambda}^{-\frac{1}{2}}\mathbf{E}^{\mathbf{T}}\right)\mathbf{x},\ \tilde{\mathbf{H}} = \left(\mathbf{E}\boldsymbol{\Lambda}^{-\frac{1}{2}}\mathbf{E}^{\mathbf{T}}\right)\mathbf{H}$$

where $\mathbf{E}$ is the matrix of eigenvectors of $E\left\{\mathbf{x}\mathbf{x}^{\mathbf{T}}\right\}$ and Λ is the diagonal matrix of corresponding eigenvalues. The following optimization is then performed:

$$\begin{aligned} &\mathbf{w} = \arg\left(\max\left(\text{kurtosis}\left(\mathbf{w}^{\mathbf{T}}\mathbf{y}\right)\right)\right) \\ &\text{subject to} \left\|\mathbf{w}\right\| = 1. \end{aligned} \tag{4}$$

The basis for maximizing a non-Gaussianity measure such as the kurtosis is provided by the reasoning that a linear combination of non-Gaussian random variables is closer to Gaussian than each individual random variable. Thus, each maximum of the objective function likely corresponds to a value of $\mathbf{w}$ for which the row vector $\mathbf{w}^{\mathbf{T}}\tilde{\mathbf{H}}$ has a high value in one of its components and values close to zero in the rest. As many values of $\mathbf{w}$ are computed at maxima as the number of endmembers and equations are solved to determine $\tilde{\mathbf{H}}$ and $\mathbf{s}$ using (2) and (3). Matrix $\mathbf{H}$ is then determined from $\tilde{\mathbf{H}}$. References [10]-[13] provide a more detailed treatment of ICA.

The mixture model concept of ICA can be extended further to perform unsupervised classification of remote sensing images, where the distribution of the entire data is modeled as a weighted sum of the class-component densities [14]. When the class-component densities are assumed to be multivariate Gaussian, the mixture model is known as the Gaussian mixture model. However, if a class happens to be multimodal, it is no longer appropriate to model the class with a multivariate Gaussian distribution and, as we have seen, ICA exploits higher order statistics in multivariate data. The ICA mixture model (ICAMM) algorithm [15], derived from ICA, can be implemented for unsupervised classification of non-Gaussian classes from hyperspectral data [16].

4. SUB-PIXEL OR SUPER RESOLUTION MAPPING AND MARKOV RANDOM FIELDS

Occurrence of mixed pixels in remote sensing images is a major problem particularly at coarse spatial resolutions such as those obtained, for example, from the hyperspectral MODIS sensor. By mixed pixels we mean, as in the previous section, the situation where the spectral data of a pixel correspond not just to one material but to a mixture of more than one. Therefore, sub-pixel classification is often preferred, where a pixel is resolved into various class components (also called class proportions or fractions). In sub-pixel classification, a pixel is decomposed into a number of component classes by assigning membership grades to each class within the pixel [17]. These membership grades or values reflect the proportions of classes in a mixed pixel. Some of the prevalent techniques used for sub-pixel classification are fuzzy *c*-means clustering, LMM and artificial neural networks.

Under most circumstances, classification at the sub-pixel level is meaningful and informative. However, it fails to account for the spatial distribution of class proportions within the pixel [18]. An alternative approach is to consider the spatial distribution of class proportions within and between pixels to perform sub-pixel super resolution mapping (i.e. mapping at a spatial resolution finer than the size of the pixel of the image). Tatem *et al.*, [19] provide an excellent review on this subject. A range of algorithms based upon knowledge-based procedures, Hopfield neural networks, linear optimization, genetic algorithms and neural network predicted wavelet coefficients, have been proposed for super-resolution mapping. Markov random field (MRF) models are also well suited to represent the spatial dependence within and between pixels.

Under an MRF model, a statistical correlation of intensity levels among neighboring pixels can be exploited. MRF has long been recognized as a useful model to describe a variety of image characteristics such as texture. Under this model, the configuration (intensity level) of a site (pixel) is assumed to be statistically independent of configurations of all remaining sites excluding itself and its neighboring sites when configurations of its neighboring sites are given. In other words, the configuration of a pixel given the configurations of the rest of the image is the same as the configuration of a pixel given the configurations of its neighboring pixels. This can mathematically be represented as:

$$\Pr\left(x(t)\middle|X\left(\mathcal{T}-\left\{t\right\}\right)\right)=\Pr\left(x(t)\middle|X(N_t)\right) \quad \textbf{(5)}$$

where $\mathcal{T}-\left\{t\right\}$ is the set of all the pixels in $\mathcal{T}$ excluding the pixel t, and N_t is the set of pixels in the neighborhood of pixel t. For example, in the context of classification of remotely sensed images, this property implies that the same class is more likely to occur in connected regions than at isolated pixels. Hence, the conditional probability density functions (PDFs) in Eq. (5) have a higher value if the configuration of a pixel t is similar to the configurations of its neighboring pixels than the cases when it is not. From [20] and [21], the marginal PDF of X takes the form of Gibbs distribution,

$$\Pr(X)=\frac{1}{Z}\exp\left[-\sum_{C\subset \mathrm{T}} V_C(X)\right] \quad \textbf{(6)}$$

This model may fit in a variety of remote sensing applications. For instance, in a classification problem, the spatial structure is usually in the form of homogenous regions of classes. As a result, an MRF model based approach assigns higher weights to these regions than to the isolated pixels thereby accounting for spatial dependence in the dataset. The approach is based on an optimization whereby raw coarse resolution images are first used to generate an initial sub-pixel classification, which is then iteratively refined to accurately characterize the spatial dependence between the class proportions of the neighboring pixels. Thus, spatial relations within and between pixels are considered throughout the generation process of the super resolution map. The implementation of MRF model for the generation of sub-pixel maps from HYMAP sensor can be seen in Kasetkasem *et al*. [22].

4. CONCLUSION

Hyperspectral imaging is a fast growing area in remote sensing for a variety of applications related to the Earth's environment. The field poses various problems that fit naturally into a signal processing framework. It is hoped that the papers in this special session will motivate more members of the signal processing community to investigate the development of solutions to these problems.

Acknowledgements: This work was supported in part by NASA under grant NAG5-11227 and by a grant from NYSTAR/CEIS with corporate matching from ITT Corporation.

5. REFERENCES

[1] Lucas, R., Rowlands, A., Niemann, O., and Merton, R., "Hyperspectral semsors and applications," In *Advanced Image Processing Techniques for Remotely Sensed Hyperspectral Data,* (Eds.) P. K. Varshney and M. K. Arora, Springer, Berlin-New York, 2004.

[2] Richard J. Aspinall, W. Andrew Marcus, Joseph W. Boardman, "Considerations in collecting, processing, and analyzing high spatial resolution hyperspectral data for environmental investigations," *Journal of Geographical Systems, Vol. 4, No. 1,* pp. 15-29, March 2002.

[3] DeFries RS, Chan JC, "Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data." *Remote Sensing of Environment* 74: 503-515, 2000.

[4] Shaw, G. and Manolakis, D, "Signal processing for hyperspectral image exploitation," IEEE Signal Processing Magazine, pp. 12-16, January 2002.

[5] Van der Meer F, "Imaging spectrometry for geologic remote sensing," *Geologie en Mijnbouw* Vol. 77, pp.137-151, 1999.

[6] Rao, R. and Arora, M. K., 2004, Overview of Image Processing, In *Advanced Image Processing Techniques for Remotely Sensed Hyperspectral Data*, (Eds.) P. K. Varshney and M. K. Arora, Springer Verlag Heidelberg, Germany pp. 51-85.

[7] Keshav, N. and Mustard, J, F.,, "Spectral unmixing," *IEEE Signal Processing Magazine*, pp. 44-57, January 2002.

[8] Hyvärinen A, Karhunen J, Oja E *Independent component analysis.* John Wiley and Sons, New York, 2001

[9] R.S. Rand, H. Chen and P.K. Varshney, "Separating patterns and finding the independent components of mixed signals based on non-Gaussian distribution properties," in *Proceedings of SPIE -- Volume 5909 --Applications of Digital Image Processing XXVIII*, Andrew G. Tescher, Editor, pp. 1-12, September 2005.

[10] P. Comon, "Independent component analysis: a new concept?" *Signal Processing*, pp. 287-314, April 1994.

[11] Chang C-I, Chiang S-S, Smith JA, Ginsberg, "Linear spectral random mixture analysis for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing* 40(2): 375-392, 2002.

[12] Robila SA, Haaland P, Achalakul T, Taylor S, "Exploring independent component analysis for remote sensing," *Proceedings of the Workshop on Multi/Hyperspectral Sensors, Measurements, Modeling and Simulation,* Redstone Arsenal,

Alabama, U.S. Army Aviation and Missile Command, CD, 2000.

[13] Robila S.A., and Varshney P.K., "Target detection in hyperspectral images based on independent component analysis," in *Proceedings of SPIE Automatic Target Recognition XII*, 4726, F. Sadjadi, Editor, pp 173-182, 2002.

[14] Duda RO, Hart PE, Stork DG, *Pattern classification,* 2nd edition. John Wiley and Sons, New York, 2000.

[15] Lee T-W, Lewicki MS, Sejnowski TJ "ICA mixture models for unsupervised classification of non-Gaussian classes and automatic context switching in blind signal separation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22: pp. 1078 – 1089, 2000.

[16] Shah CA, Arora MK, Varshney PK, "Unsupervised classification of hyperspectral data: an ICA mixture model based approach," *International Journal of Remote Sensing* 25: 481 – 487, 2004.

[17] Foody GM and Cox DP, "Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions," *International Journal of Remote Sensing* 15: 619-631, 1994.

[18] Verhoeye J and Wulf RD, "Land cover mapping at sub-pixel scales using linear optimization techniques," *Remote Sensing of Environment* 79: 96-104, 2002.

[19] Tatem AJ, Hugh G, Atkinson PM, Nixon MS, "Super-resolution land cover pattern prediction using a Hopfield neural network." *Remote Sensing of Environment,* 79: 1-14, 2002.

[20] Winkler G, *Image analysis random fields and dynamic Monte Carlo methods.* Springer Verlag, New York, 1995.

[21] Bremaud P, *Markov chains Gibbs field, Monte Carlo simulation and queues.* Springer Verlag, New York, 1999.

[22] Kasetkasem, K. Arora, M. K. and Varshney, P. K., "An MRF model based Approach for Sub-pixel Mapping from Hyperspectral Data," In *Advanced Image Processing Techniques for Remotely Sensed Hyperspectral Data,* (Eds.) P. K. Varshney and M. K. Arora, Springer Verlag, 2004.