ISSN: 2278-0181

Development of Digital Spectral Library and Supervised Classification of Rice Crop Varieties Using Hyperspectral Image Processing

1. B Mahendra Kumar,
Dept of MCA, DSCE,
Bangalore-560078,

2. Dr. S. Karthik
Prof. & Dean
SNSCT, Coimbatore

3. Dr. Shanmukhappa
Secretary & Prof.
SNSCT, Coimbatore

Abstract: **- Hyperspectral Image Processing a promising technique in remote sensing image analysis provides more complete and detailed spectral information about land cover area and has a great potential to discriminate specific plant species using hundred numbers of contiguous bands. This study produces digital spectral library of Rice (Oryza sativa L.) varieties – Ratan (IET-1411), CSR-10(IET-10349/10694), Haryana Basmati-1 (IET-10367), HKR-126 and CSR-13(IET-10348). A supervised classifier called Extreme Learning Machine (ELM) is used to develop the classified image at pixel scale with high classification accuracy. After pre-processing, the classification of rice crop at pixel scale across 155calibrated spectral bands has shown good results.**

Keywords—Hyperspectral Data, Spectral Library, Classification

1. INTRODUCTION

Imaging spectroscopy (Goetz et al., 1985), also known ashypers pectral imaging, is concerned with the measurement, analysis and interpretation of spectra acquired from a given scene (or specific object) at a short, medium or long distance by an airborne or satellite sensor. The concept of imaging spectroscopy originated in the 1980's,when A. F. H. Goetz and his colleagues at NASA's Jet Propulsion Laboratory began a revolution in remote sensing by developing new instruments such as the Airborne Imaging Spectrometer (AIS), then called AVIRIS, for Airborne Visible Infra-Red Imaging Spectrometer(Green, 1998).

The special characteristics of hyperspectral datasets pose different processing problems, which must be necessarily tackled under specific mathematical formalisms, such as classification and segmentation (Jiaetal., 1999) or spectral mixture analysis (Adams et al., 1986; Smith et al.,1990a,b).

For instance, several machine learning and image processing techniques have been applied to extract relevant information from hyper spectral data during the last decade (Varshney & Arora, 2004).Taxonomies of remote sensing data processing algorithms (including hyperspectral analysis methods) have been developed in the literature (King, 2003; Keshava& Mustard, 2002; Richards, 2005). It should be noted, however, that most available hyperspectral data processing techniques focused on analyzing the data without incorporating information on the spatially adjacent data, i.e., hyperspectral data are usually not treated as images, but as unordered listings of spectral measurements with no particular spatial arrangement (Tadjudin&Landgrebe, 1998).

2. RELATED WORKS

Hyperspectral imagery organically includes the spectral information and space information of the ground objects, so it can bring opportunity to ground objects recognition more precisely. Because the performance of many kinds of classifiers can often be dramatically improved by AdaBoost algorithm, in this paper, we introduce the basic procedure of the Discrete AdaBoost algorithm for two-class classification problem, describe the decision stump classifier used as weak learner, and then we bring forward the multiclass Gentle AdaBoost algorithm using hamming loss for hyperspectral imagery classification (Guopeng Yang et al., 2009).

3. AREA OF STUDY AND DATASETS

The study area is Bapauli a town of Panipat district in Haryana state in India and its nearby region. The one side of the study area is surrounded with Yamuna River, Habitation (district Panipat), Yamuna Canal and the boundary of the other province Uttar Pradesh. The study area is lying between 29°21′-29°16′ N and 77°3′- 77°7′ E, and with area of 176.26 ha as shown in Figure 1.

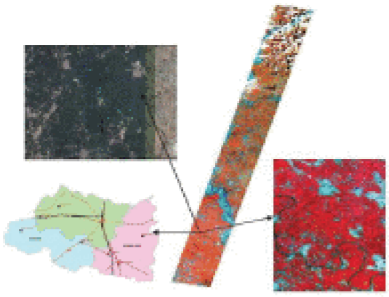


Figure 1: False color image of the study area by EO-1 Hyperion with a spectral resolution of30 meter/pixel

In this research EO (Earth Observing)-1 Hyperion data is used. It is acquired with two spectrometers in push-broom mode over the Western part of Uttar pradesh and Haryana test site on September 2, 2005. The Hyperion L1B GeoTIFF (Geographical Tagged Image File Format) data is provided by United State Geological Survey (USGS). The EO-1 satellite is in a sun-synchronous orbit at 705 km altitude. The product details are summarized in Table 1.

Table 1. EO-1 Hyperion product details

S. No.	Variables	Details
1.	PRODUCT_ID	□470392005245_20001□
2.	PRODUCT_CREATION_TIME	2009 06 30 T 03:052:07Z
3.	GROUND_STATION	□HGS□
4.	PRODUCT_TYPE	□L1GST□
5.	SPACE_CRAFT_ID	□EO1□
6.	SENSOR_ID	□HYPERION□
7.	ACQUISITION_DATE	2005 09 02
8.	START_TIME	2005 05:18:46
9.	END_TIME	2005 05:23:05
10.	PRODUCT_SAMPLES	1021
11.	PRODUCT_LINES	3291
12.	SENSOR_LOOK_ANGLE	16.657
13.	SUN_AZIMUTH	131.214424
14.	SUN_ELEVATION	59.697736
15.	BYTE_ORDER	IEEE
16.	PIXEL_SIZE	30m
17.	FILE_SIZE	1,626,293,724 bytes

4. METHODOLOGY

A. Image Pre-processing

The Pre-processing of satellite data is necessary not only toremove the sensor errors during the acquisition but also display corrections, band selection, reducing data dimensionality and to reduce the computational complexity. In this research the EO-1, Hyperion image is corrected radiometrically, geometrically, abnormal band and pixels are detected and corrected, spectral smoothing to correct the spectral bands and atmospheric correction is implemented for better reflectance.

Abnormal Pixels Detection and Correction

Hyperion acquires data in pushbroom mode with two spectrometers. The maximum abnormal pixels in Hyperion imagery data appear as dark vertical strips in VNIR and SWIR regions. These vertical strips are high frequency errors in form of DN value and divided into four distinct classes, Continuous with atypical DN values, Continuous with constant DN values, Intermittent with atypical DN values and Intermittent with lower DN values (Goodenough D. G.*et al.* 2003).

In order to de-striping to function properly, the data must being the acquired format and cannot be rotated. These bands forbad columns are inspected visually to eliminate the striping errors. In stripping pixels correction 12 bands are visually inspected for de-striping, the first 11-VNIR (band 8, 9, 10,11, 12, 13, 14, 15, 16, 56, 57) and 1-SWIR (band 183).

Spectral Smoothing and Atmospheric correction using AE

AE refers to the combination of EFFORT (Empirical Flat Field Optimal Reflectance Transformation) custom smoothing algorithm with ATREM (Atmospheric Removal).

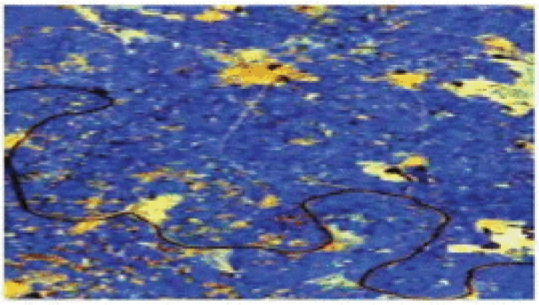


Figure 2: Hyperion FCC for bands 204:150:93 (SWIR) after preprocessing

B. Spectral Digital Library

In satellite image processing the reflectance value of an object can be represented in form of graph or curve called spectral signature. In the present study multiple ROIs (Regions of Interest) with unique color composition for digital library and classification of rice crop varieties are listed in Table 3.

S.NO.	Class Type	No. of pixels	Class Color Composition
1.	RICE 1 Ratna (IET-1411)	31	GREEN
2.	RICE 2 CSR-10(IET-10349/10694)	07	GREEN 1
3.	RICE 3 Haryana Basmati-1(IET-10349/10694)	12	GREEN 2
4.	RICE 4 HKR-126	26	GREEN 3
5.	RICE 5 CSR-13 (IET-10348)	16	SEA GREEN
6.	WATER 1	64	WHITE
7.	HABITATION	66	CYAN 3
8.	POND	04	BLUE
9.	VEG 1(Sugarcane 1)	02	MAGENTA
10.	VEG 2(Sugarcane 2)	01	MAROON
11.	OPEN FIELD	07	CYAN
12.	MOISTURE	03	PURPLE
13.	SAND	08	RED
14.	WATER 2	02	BLUE 2
15.	BUSHES	20	YELLOW
16.	BARON LAND	02	COYRAL
Total Sample Pixels		253	

Table 2: Rice crop variety and color composition details

In hyperspectral image analysis, the digital value of a pixel corresponds to spectral band reflectance in the pixel at nth spectral band. The spectral reflectance value at nth spectra is analyzed for assessing the spectral separability as well as in image classification. The proposed library has the contribution of *photosynthetic pigments* (400-700nm) and *chlorophyll absorption* (600nm) in the first segment of spectra, *red edge-chlorophyll* (700-750nm) in the second phase of spectra, the *liquid water variation* (1080-1170) in the third phase of spectra, the various *leaf waxes and oil*(1700-1780nm) and *cellulose* (2100nm) in the fourth phase of spectra, the *soil properties* (2100-2300nm) and *nitrogen/protein* (2280-2290nm). In this study spectral curve generated for RICE species in the Kharif season, it has shown diversity in reflectance pattern depending on the crop characteristics and growth stage of the crop in the VNIR and SWIR regions shown in Figure 3.

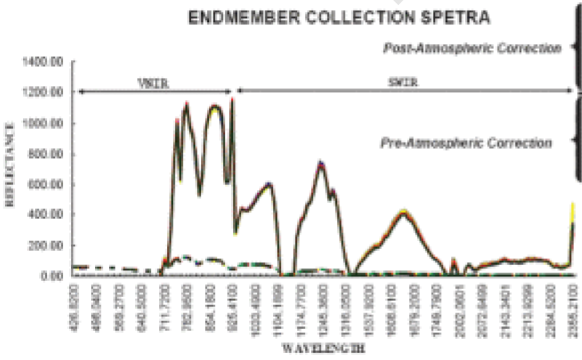


Figure 3. Spectral reflectance curves of various rice verities in VNIR-SWIR bands

i. Reflectance Retrieval

The purpose of atmospheric correction is to decrease the influence of atmosphere on object reflectance. Impact of correction is more visible in VNIR region, for all land cover classes, the reflectance value in visible region is low, while in the NIR-SWIR region it founds high. The pre and post atmospheric spectral library of different rice varieties is shown in Figure 3. The reflectance value is increased in the following spectra:

- 33-96(681.20 – 1104.19nm) • 102-163(1164.68-1780.09nm) • 183-184(1981.86-1991.96nm)
- 187-220(2022.25-2355.21nm) whereas most of the reflectance value is decreased in the following spectra:
- 8-32 (426.82-671.02nm) • 97-101(1114.19-1154.58nm) • 164 (1790.19nm) • 185-186 (2002.06-2012.15nm).

Therefore, the reflectance values of the wavebands 8-32(426.82-671.02nm), 97-101(1114.19-1154.58nm),64(1790.19nm) and 185-186(2002.06-2012.15nm) are found negative and set to zero. In total there are 120 such bands that contained positive DN values, whereas 35 negative and set to zero.

C. ELM Classification for Hyperspectral Image

i. Mask Image

The next step of methodology is to prepare a mask band (image) to mask out non-vegetation (WATER 1, POND,HABITATION, OPEN FIELD, MOISTURE, SAND,WATER 2, SHRUB, BARONLAND) and vegetation (VEG1and VEG2) classes. The purpose of this binary image is to reduce the data for analysis and ensure that the classification performed later will enhance contrast among different varieties of rice crop.

ii. Extreme Learning Machine classifier to classify the image

Extreme Learning Machine (ELM) meant for Single Hidden Layer Feed-Forward Neural Networks(SLFNs) will randomly select the input weights and analytically determine the output weights (S. Karpagachelvi et al., 2011). This algorithm tends to afford the best generalization performance at extremely fast learningspeed.

The structure of ELM network is shown in figure 1. ELM contains an input layer, hidden layerand an output layer.

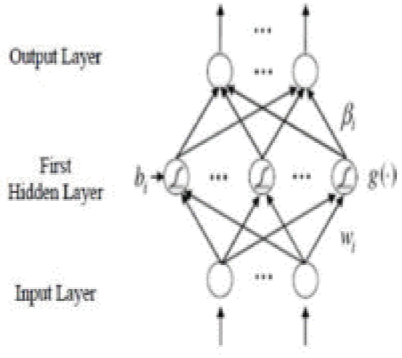


Figure 4: Structure of ELM

The ELM has several interesting and significant features different from traditional popular learning algorithms for feed forward neural networks. These include the following:

• The learning speed of ELM is very quick when compared to other classifier. The learning process of ELM can be performed in seconds or less than seconds for several applications.

• The ELM will attain the results directly without any difficulties. The ELM learning algorithm is much simple than the other learning

Extreme Learning Machine Training Algorithm, it uses a finite number of input-output samples for training. If there are N samples are considered in the lung cavity region as (X_i, t_i), where $X_i = [X_{i1}, X_{i2} \dots\dots X_{in}]^T \in R^n$ and $t_i = [t_{i1}, t_{i2} \dots\dots t_{im}]^T \in R^m$ are the cavity and the area of the cavity region, then the standard SLFNwith N hidden neurons and activation function g(x) isdefined as:

$$\sum_{i=1}^{\tilde{N}} \beta_i g(w_i . X_j + b_j) = o_j, \qquad j = 1, \dots\dots, N$$

Where $w_i = [w_{i1}, w_{i2} \dots\dots w_{in}]^T$represents the weight vector that links the ith hidden neuron and the input neurons,$\beta_i = [\beta_{i1}, \beta_{i2} \dots\dots \beta_{im}]^T$represents weight vector that links theith neuron and the output neurons, and b_i represents the threshold of the ith hidden neuron. The "." in w_i. x_j indicates the inner product of w_i and x_j. The SLFN try to reduce the difference between o_j and t_j. More in a matrix format as H ß = T, where

$$H(a_1, \dots\dots, a_{\tilde{N}}, b_1, \dots\dots b_{\tilde{N}}, X_1, \dots\dots, X_N) = \begin{bmatrix} g(a_1, b_1, X_1) & \cdots & g(a_{\tilde{N}}, b_{\tilde{N}}, X_1) \\ \vdots & \ddots & \vdots \\ g(a_1, b_1, X_N) & \cdots & g(a_{\tilde{N}}, b_{\tilde{N}}, X_N) \end{bmatrix}_{N \times \tilde{N}}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_{\tilde{N}}^T \end{bmatrix}_{\tilde{N} \times m} \text{ and } T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m}$$

The result reduces the norm of this least squares equation is:

$$\hat{\beta} = H^+ T$$

Where $H^+$the Moore-Penrose is generalized inverse (Sluimer IC, et al., 2003) of the hidden layer output matrix H. The ELM algorithm which consists of only three steps, can then be summarized as

Step 1: Given a training set $\aleph = \{(X_i, t_i) | X_i \in R^m, i = 1, \dots\dots, N\}$activation function g(x), and hidden number node $\tilde{N}$,

1) Assign random hidden nodes by randomly generating parameters (a_i, b_i)according to any continuous sampling distribution, $i = 1, \dots\dots, \tilde{N}$

2) Calculate the hidden layer output matrix H.

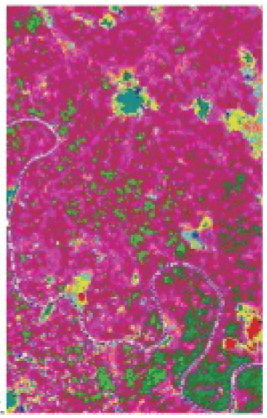
3) Calculate the output weightβ: $\hat{\beta} = H^+ T$

Then find the maximum repeated pixel intensity of the rice region. To discover the maximum repeated pixel, first have to find the intensities of all the pixels in each rice region by implementing histogram and thereafter need to compare all the pixels of rice region with each other. After discovering the maximum repeated pixel of the rice region, have to give the result to the classifier. Similarly, find the maximum repeated pixel of the all rice region is predicted and give the result to the classifier. The classifier classifies the rice region by comparing all the features.

5. EXPERIMENTAL RESULTS

Hyperspectral Image Processing System (HIPS) demonstrate the capacity of rice crop identification. The provided EO-1 Hyperion L1GST is a raw satellite image and not corrected properly. A procession of Pre-processing is employed to get a satisfactory reflectance surface to develop a spectral library and image classification. A specific tedious spectral library of rice crop varieties named *Ratan (IET-1411), CSR- 10(IET-10349/10694), Haryana Basmati-1 (IET-10367), HKR-126* and *CSR-13(IET-10348)* is developed using hyperspectral data. When the sixteen end members are processed by SAM algorithm, an excellent discrimination between the different end members is found in the spectral signatures.

For example, confusion between RICE 1 and VEG 2 classes, RICE 2 and RICE 1 classes, RICE 3 and RICE 2 classes, RICE 4 and RICE 5 classes and RICE 5 and RICE 4 is reduced to 0.00% (from 9.09%), to 14.29% (from 28.57%), to 0.00% (from 8.33%), to 7.69% (from 19.23%) and to 12.50% (from 18.75%) respectively.



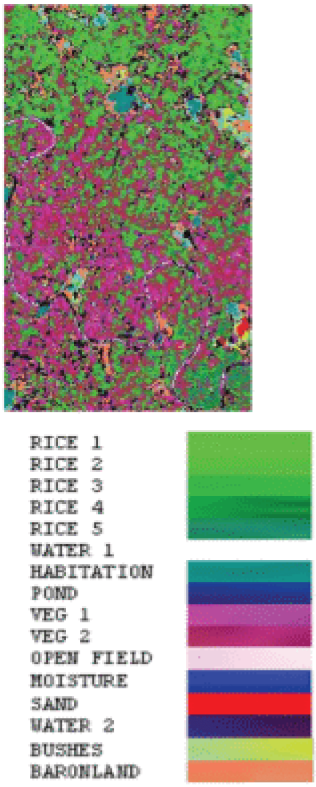


Figure 5: Classified images by SAM classifier of the study area: (a) Before Pre-processing (b) After Pre-processing

Table 3. Before Pre-processing classification accuracy using SAM classifier in VNIR-SWIR

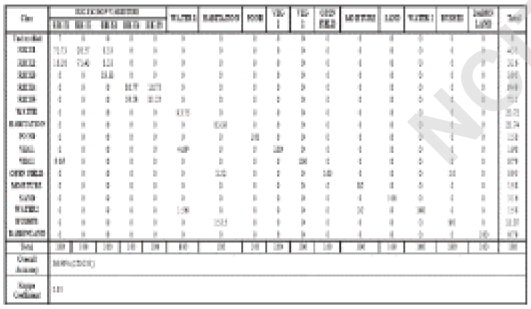
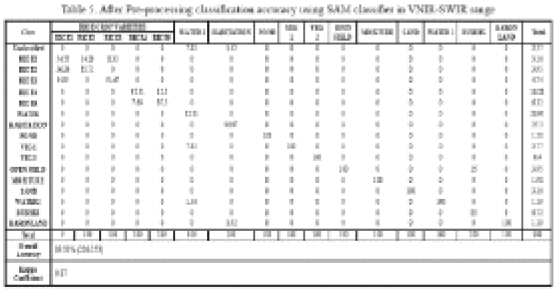


Table 4. After Pre-processing classification accuracy using SAM classifier in VNIR-SWIR range



6. CONCLUSION

During the study a chain of Pre-processing methods to processed EO-1 Hyperion data for developing spectral library and classification. The geometric, radiometric, selection of celebrated bands, abnormal pixel, spectral smoothing and atmospheric corrections are performed before spectral analysis. In pre-processing, each spectral band is examined physically and digitally excluded overlapped and noise bands. The final dataset spans the spectral range from 426.82nm to 2355.20nm with a total of 155 spectra. These methods have been explored to increase the spectral reflectance quality to develop a region spectral library of rice crop and its varieties. It is concluded, on comparing classification accuracy and classification results, the ELM algorithm gives 89.33% overall accuracy and map the rice based agricultural area better than before Pre-processing classification i.e. 86.96%.

REFERENCES

1. Goetz, A. F. H., Vane, G., Solomon, J. E., & Rock, B. N. (1985). Imaging spectrometry for Earth remote sensing. Science, 228, 1147–1153.
2. Green, R. O., Eastwood, M. L., Sarture, C. M., Chrien, T. G., Aronsson, M., Chippendale, B. J.,et al. (1998). Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS). Remote Sensing of Environment, 65, 227–248.
3. Jia, X., Richards, J. A., &Ricken, D. E. (1999). Remote sensing digital image analysis: an introduction. Berlin: Springer-Verlag.
4. Adams, J. B., Smith, M. O., & Johnson, P. E. (1986). Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 site. Journal of Geophysical Research, 91, 8098–8112
5. Chen-guangDai ; Xiao-bo Huang ; Guang-jun Dong," Support Vector Machine for Classification of Hyperspectral Remote Sensing Imagery", IEEE, 2007.
6. LinlinShen ;SenJia, "Three-Dimensional Gabor Wavelets for Pixel-Based Hyperspectral Imagery Classification", Geoscience and Remote Sensing, IEEE Transactions on Volume: 49 , Issue: 12 ,2011.