# Classification of Plant Species using Hyperspectral Remote Sensing Data

**Amarsinh B. Varpe*, K.V. Kale**, Amol D. Vibhute***, Rupali R. Surase****, Sandeep V. Gaikwad*****, Dhananjay B. Nalawade******, Yogesh D. Rajendra*******

Geospatial Technology Research Lab, * Department of Computer Science & I.T,
Dr.Babasaheb.Ambedkar.Marathwada. University Aurangabad. (M.S) India

***Abstract-*** *Hyperspectral remote sensing has been widely used for identification of plant species. Hyperion hyperspectral satellite data with 10nm spectral resolution of Aurangabad region of Maharashtra (India) were studied for classification. Three various plants types have been selected from the hyperion image. SVM has achieved overall accuracy of 82.75 % with Kappa coefficient 0.80. The Binary encoding classification has given better the result with an accuracy of 66.21% with kappa coefficient 0.61. The result is of significance for plant analysis of very complex region.*

***Keywords-*** Hyperspectral Data, Hyperion, QUAC, Classification Techniques.

## I. INTRODUCTION

Recently, Hyperspectral remote sensing imaging is spectral imaging technique that is able to find materials, identify and distinguish spectrally unique materials. This is done by collecting and processing hundards of contiguous narrow wavebands from the scene, which provide spectral information [1]. The information is acquired by using an airborne or satellite sensor at a short, medium or long distance from the scene [2].The main advantage of hyperspectral image is its potential to provide more accurate results than any other type of remotely sensed techniques, because they commonly collect more than 200 spectral bands to perform a detailed information extraction in order to classify, identify, and detect objects [3][4].

The small range of spectral band is a primary disadvantage to multispectral sensors [5]. Some of the practical applications for hyperspectral image classification are vegetation, soil classification, mineral identification, rock type classification etc. The main usage of hyperspectral imaging is for vegetation and minerals extraction. The objective of hyperspectral image classification is to identify the features occurring in an image in terms of the object or different features related to the land represented in the ground as shown in Fig.1. It is difficult to find an appropriate and optimum classifier for all situations as the characteristics of each data set and the circumstances for each study vary so differently [6]. The present study reports results obtained by two supervised classifiers, using SVM and BE classifiers.

## II. STUDY AREA

The region of Aurangabad distraction located on. $19^0$28’43.27”-$20^0$24’52.19” N latitude and $75^0$13’10.75”– $75^0$30’14.87” E longitude “EO1H1460462015358110KV_MTL_L1T” has been selected for our study. The total area covered about 241 $km^2$. The estimated area related to agriculture, urban, water was about 1.41.1 $km^2$. One example of the hyperspectral image is given in Fig.1. It has a usual semiarid climate with an annual average temperature of 05 % C to 46 % C and Humidity was 15 % annual average rainfall of 734 nm.

**Figure.1** The acquired hyperspectral original image covering the studied area RGB-VNIR (29, 20, 11)

## III. DATASETS

The satellite data for the study was taken from United States Geological Survey (USGS) with NASA Earth Observing-1 (EO-1) satellite of Hyperion sensor hyperspectral imagery. The Hyperion space borne sensors used in this study have information like different spectral, spatial, and radiometric characteristics. Hyperion sensor on board the Earth Observing-1 (EO-1) mission is unique as it collects high spectral resolution data in the range of VIS/NIR and SWIR wavelengths, with 220 contiguous spectral bands of these, 196 are well calibrated, whereas 24 bands were considered uncalibrated because they did not meet desired performance requirements or were noisy bands. Hyperion is a push-broom instrument providing from a 705 km orbit, 30 m spatial resolution imagery over a 7.5 km wide swath perpendicular to the satellite motion. Reference data collected during an extensive field survey were used in the fuel models development and the later accuracy assessment of the captured maps. The information about types of band was collected ground points in the study area. The geographical coordinate were acquired through google software and the types were determined by observation such 400 points were sampled in the month of Dec 2015 which was about two months later of acquiring image by the satellite. The details about Hyperion parameters used for research study with spectral, spatial coverage information. EO-1 Hyperion acquisition date was 24 December 2015 collected from USGS site with hyperion parameters shown in Table.1.

**Table.1.** Hyperion hyperspectral data parameters used for research study

| **Spacecraft/Instrument** | **USGS EO-1/Hyperion** |
|---|---|
| Entity ID | EO1H1460462015358110Kv-AK3-01 |
| Product Type | Level T(Geo TTFF) |
| Acquisition Date | 2015-12-24 |
| Orbit Path/Target Path | 147/46 |
| Orbit Row/Target Row | 146/46 |
| Spectral Range | 355.59-2577.08 nm |
| Visible Bands/Near Infrared Bands/Short Wave Infrared | 35/35/172 |
| Spatial Resolution/ Number of Bands/Swath Width | 30 m/242/7.5 km |
| Spectral Coverage | Continuous |
| Temporal Resolution | 200 days |

## IV. METHODOLOGY

The present study was performed through Environment for Visualizing Imaging (ENVI-5.3) Hyperspectral image analysis software for data processing. Details are discussed in the following sections.

### *A. Hyperspectral data pre-processing*

For using Hyperion data in ENVI software, Hyperion tools plug-in was used to convert level 1R HDF and level 1 G/1T HDF with GeoTIFF file formats into ENVI format. Hyperion toolkit provides particular parameters to each input dataset that will be helpful for Hyperion data within ENVI.

4.1 Transformation of DN to Pre-Processing

EO-1 Hyperion hyperspectral image consists of number of continuous spectral bands, each pixel of which stored the energy as a digital number (DN). Stacked image is used to convert DN into Radiance values. The digital numbers were stored as 16 bit signed integer. Image was converted into absolute radiance and each band of NIR 1 to 70 and SWIR 71 to 242 was divided by equation.1 [7].

$$\frac{VNIR}{40} and \frac{SWIR}{80} \qquad 1$$

Where, Visual Near Infrard (VNIR)
Short Wave Infrard (SWIR)

4.2 Transformation of DN to Radiance Conversion

To convert the radiance into reflectance, equation.2 was used on specific band and stacked for further processing step.

$$\frac{\pi . L_\lambda . d^2}{\cos\theta_s . ESUN_\lambda} \qquad 2$$

Where, $L_\lambda$ is spectral radiance of sensor ($W.m^{-2}.sr^{-1}.\mu m^{-1}$), $d$ is Earth-Sun distance at Astronomical unit, $\theta_s$ is solar zenith angle, $ESUN_\lambda$ is mean solar Exo-atmophersic irradiances at band $\lambda$ ($W.m^{-2}.\mu m^{-1}$).

4.3 Quick Atmospheric Correction (QUAC)

It is a scene empirical approach used for the removal of atmospheric effects. It is based on the radiance values of the image. The QUAC algorithm is scence based on empirical approach which is based on radiance values of a view for the removal of atmospheric effects. It works on multispectral as well as hyperspectral data in VNIR to SWIR spectral wavelength array. As compared to other methods, it has used atmospheric recompense factors directly from the information contained within the image scene ancillary information. It has achieved atmospheric retrieval components directly through data contained within the view i.e. observed pixel spectrum [7].
The QUAC algorithm has three basic steps,

Endmember selection or pure objects composing the scene
Baseline estimation
Reflectance estimation and can be solved through the equation.3

$$\rho = (\rho_1 + \rho_2 + ... + \rho_n)/n \qquad 3$$

Where, n specifies the endmember number.

*B. Hyperspectral Image Classification Algorithm*

This section contains classification methods followed in this work.

1. Binary Encoding

Binary encoding (BE) is a standard technique in classifying hyperspectral images. The basic idea is to reduce the large amount of data while preserving as much information as possible. Standard binary encoding reduces the information of a pixel into one or two bits per channel only [8]. We consider regions rather than individual pixels and our code for an image region is,

$$2L+28 \text{ bits} \qquad 4$$

Where, L is the number of spectral channels of the hyperspectral image. The code consists of four parts, i.e., spectrum, size, shape, and height. The spectral amplitude and slop are represented by 2L bits, the size and shape of the segment is coded by 25 bits and the relative height of a segment is represented by 3 bits.

2. Support Vector Machine

Support Vector Machine (SVM) is a supervised classification method derived from statistical learning theory that often yields good classification results from complex and noisy data. It separates the classes with a decision surface that maximizes the margin between the classes. The surface is often called the optimal hyperplane and the data points closest to the hyperplane are called support vectors. The support vectors are the critical elements of the training set [9][10].

The kernel simulated the ridge of the preliminary data in a feature space with high dimension.
To make data linearly separable, the dot product (x, xi) is replicated by the kernel function.
The polynomial kernel performs well in machine learning approach by performing activation of nonlinear models. Gamma parameter 2 was selected for SVM classification with X as number of end members extracted from pure spectral responses as mentioned in equation.4.

$$K(x, xi) = (gxTxi + r)\ d,\ g > 0 \qquad 4$$

Where, g indicates gamma parameter, x, xi are input variable to the function.

## V. EXPERIMENTAL RESULT AND DISCUSSION

The Figure.2 shows spectral signature of vegetation after the atmospheric correction by QUAC. The original Hyperion image has spectral profile of various objects. In our study we have classified objects in eight categories which are Synygium Cumini, Azadirachta Indica, Mango, Hill Area, and Hill Area with rock, Water, Urban, Agricultural Crop.

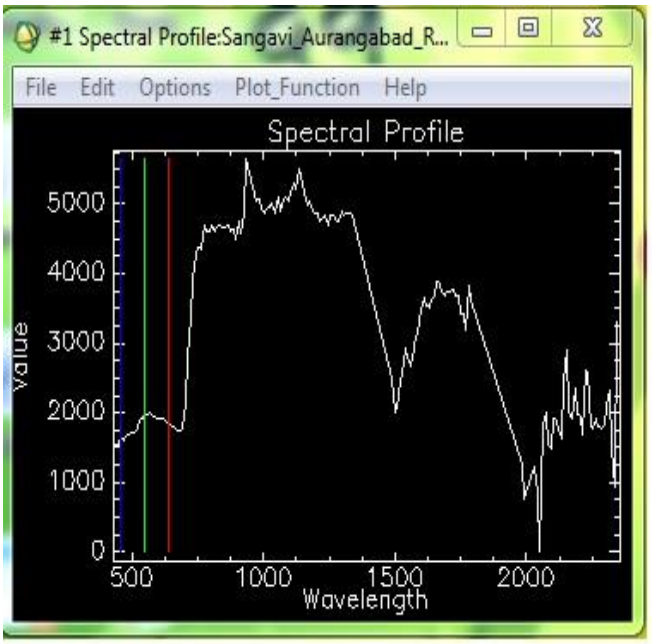


**Figure.2** Spectral signature of vegetation

The region of interest (ROI) of satellite Hyperion image were considered with minimum 50 pixels for training the samples for each class and for validation overall 400 pixels were used. Two supervised classifiers were performed to extract the different features in the Hyperion image from the trained samples. Figure.3 and Figure.4 shows the final classified (BE), (SVM) classifiers, respectively**.**

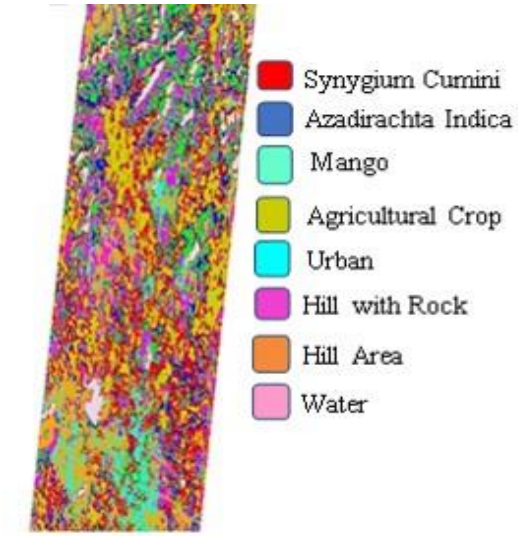


**Figure.3** Classified image using Binary Encoding classification approach

The resulted areas derived from classifications are listed in Table.2. Results of classified area shows that the synygium cumini area was identified with 38% plants. The most part of the studied area was identified by Mango i.e. 74%.
Azadirachta Indica area without large plants compared to mango plant. The hill with rock region are less in the studied region. Similarly hill area, water and Urban, Agricultural crop is very less in study region. The overall accuracy is 66.17 %, and Kappa coefficient are 0.61of Binary Encoding Classifier.

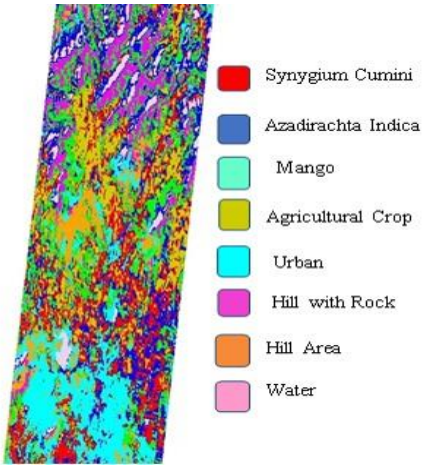


**Figure.4** Classified image using Support Vector Machine classification approach

Table.3. resulted of classified area shows that the synygium cumini area was identified with 49.02 %, mango are 20 % and Azadirachta Indica 74.00 % plants. The accuracy assessment of support vector machines are overall accuracy 82.75 % and kappa coefficient 0.80 respectively.

## VI. CONCLUSION

The study has provided useful information about classification of plants through hyperspectral signatures at Aurangabad region. This information may be used by the planner for better agricultural development. The present work has verified the use of binary encoding, support vector machine techniques for identification, mapping and classification of various types of plant species using high spectral resolution hyperspectral data. The binary encoding classifier have given 66 % accuracy where as SVM has an accuracy 82 % which indicates the correctness of SVM algorithm.

## ACKNOWLEDGMENT

Authors would like to acknowledge partial technical support SAP, DST-FIST under UGC SAP (II) DRS, partial support DST-NISA project to Department of Computer Science & I.T, Dr.B.A.M.University, Aurangabad, Maharashtra, India and financial assistance under UGC-BSR research fellowship for this work.

## REFERENCES

[1] Lamyaa G. Taha, Atia A. Shahin, Assessment of cartographic potential of airborne hyperspectral data for large scale mapping, Recent Advances in Image, Audio and Signal Processing, pp 143-153, 2013.
[2] Sahar A. Rahman, Wareen A. Aliady, Nada I.Alrashed, Supervised Classification Approaches to Analyze Hyperspectral Dataset, I.J Image Graphics and Signal Processing,pp.42-48, 2015.
[3] Stefan A. Robila, Andrew Gershman, Spectral Matching Accuracy in Processing Hyperspectral Data, IEEE, 0-7803-9029-6/05, pp.163-166, 2005.
[4] Chein-I Chang, Spectral Information Divergence for Hyperspectral Image Analysis, IEEE, pp.509-511, 1999.
[5] Amol D.Vibhute, Rajesh K. Dhumal, Ajay D. Nagne, Yogesh D. Rajendra, K. V. Kale and S. C. Mehrotra, Analysis Classification and Estimation of pattern for Land of Aurangabad Region using High-Resolution Satellite Image, Second International Conference on Computer and Communication Technologies, Advances in Intelligent Systems and Computing 380.DOI 10.1007/978-81-322-2523-2-40, Springer India, pp. 415-421, 2016.
[6] Belkacem Baassou, Mingyi He, Shaohui Mei, Yifan Zhang, Unsupervised Hyperspectral Image Classification Algorithm by Intergrating Spatial Spectral Information, ICALIP, IEEE, pp.978-982, 2012.
[7] Amol D. Vibhute, Rajesh K. Dhumal, S. C. Mehrotra, K. V. Kale, Hyperspectral Imaging Data Atmospheric Correction Challenges and Solutions using QUAC and FLAASH Algorithms, International Conference on Man and Machine Interfacing (MAMI), pp. 400-404, 2015.
[8] Huan Xie, Shanghai, China, Christian Heipke, Peter Lohmann, A New Binary Encoding Algorithm for the Simultaneous Region based Classification of Hyperspectral Data and Digital Surface Models, Photogrammetrie Fernerkundung, Geoinformation, pp. 800-805, 2011.
[9] Amarsinh B. Varpe, Yogesh D. Rajendra, Amol D. Vibhute, Sandeep V.Gaikwad, K.V.Kale, Identification of plant species using non-imaging hyperspectral data, Man and Machine Interfacing (MAMI), International Conference, IEEE, pp.1-2, 2015.

**Table.2.** Error matrix accuracy assessment of Binary Encoding classification

| Ground Truth (Percent) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Class** | **Synygium Cumini** | **Mango** | **Azadirachta** | **Hill with Rock** | **Hill Area** | **Water** | **Urban** | **Agriculture** | **Total** | **User Accuracy** | **Prod Accuracy** |
| Synygium Cumini | 38.00 | 2.00 | 0.00 | 4.00 | 0.00 | 4.00 | 4.00 | 22.00 | 9.14 | 51.35 | 38.00 |
| Mango | 16.00 | 74.00 | 20.00 | 2.00 | 0.00 | 8.00 | 0.00 | 34.00 | 19.01 | 48.05 | 74.00 |
| Azadirachta Indica | 6.00 | 14.00 | 10.00 | 0.00 | 0.00 | 2.00 | 0.00 | 28.00 | 6.17 | 56.00 | 28.00 |
| Hill with Rock | 4.00 | 8.00 | 74.00 | 6.00 | 0.00 | 14.00 | 0.00 | 6.00 | 13.83 | 66.07 | 74.00 |
| Hill Area | 0.00 | 0.00 | 4.00 | 78.00 | 0.00 | 32.00 | 0.00 | 2.00 | 14.32 | 67.24 | 78.00 |
| Water | 0.00 | 0.00 | 0.00 | 2.00 | 98.18 | 0.00 | 0.00 | 2.00 | 13.83 | 96.43 | 98.18 |
| Urban | 0.00 | 2.00 | 2.00 | 8.00 | 0.00 | 40.00 | 0.00 | 4.00 | 6.91 | 71.43 | 40.00 |
| Agricultural Crop | 36.00 | 0.00 | 0.00 | 0.00 | 1.82 | 0.00 | 96.00 | 2.00 | 16.79 | 70.59 | 96.00 |
| Total | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |

**Table.3** Error matrix accuracy assessment of Support Vector Machine classification

| Ground Truth (Percent) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Class** | **Synygium Cumini** | **Mango** | **Azadirachta Indica** | **Hill with Rock** | **Hill Area** | **Water** | **Urban** | **Agricultural Crop** | **Total** | **User Accuracy** | **Prod Accuracy** |
| Synygium Cumini | 49.02 | 2.00 | 0.00 | 0.00 | 0.00 | 2.00 | 16.00 | 10.00 | 9.85 | 62.50 | 49.02 |
| Mango | 5.88 | 72.00 | 4.00 | 0.00 | 0.00 | 8.00 | 0.00 | 14.00 | 12.81 | 69.23 | 72.00 |
| Azadirachta Indica | 27.45 | 16.00 | 10.00 | 0.00 | 0.00 | 0.00 | 0.00 | 74.00 | 14.53 | 62.71 | 74.00 |
| Hill with Rock | 0.00 | 4.00 | 96.00 | 2.00 | 0.00 | 0.00 | 0.00 | 0.00 | 12.56 | 94.12 | 96.00 |
| Hill Area | 0.00 | 0.00 | 0.00 | 96.00 | 0.00 | 0.00 | 0.00 | 0.00 | 11.82 | 100.00 | 96.00 |
| Water | 0.00 | 0.00 | 0.00 | 0.00 | 100.00 | 0.00 | 0.00 | 0.00 | 13.55 | 100.00 | 100.00 |
| Urban | 3.92 | 6.00 | 0.00 | 2.00 | 0.00 | 90.00 | 0.00 | 2.00 | 12.81 | 86.54 | 90.00 |
| Agricultural Crop | 13.73 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 84.00 | 0.00 | 12.07 | 85.71 | 84.00 |
| total | 100.0 | 100.00 | 100.00 | 100.0 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |