Crop Identification Using Unsuperviesd ISODATA and K-Means from Multispectral Remote Sensing Imagery

Namita M Kulkarni
Tikaram Jagannath Arts, Commerce and Science college, Khadki, Pune

ABSTRACT
Agriculture is one of the oldest economic practice of human civilization is indeed undergoing a makeover. Remote sensing has played a significant role in crop classification, crop health and yield assessment. Hyper spectral remote sensing has also helped to enhance more detailed analysis of crop classification. This paper focuses the unsupervised classification methods i.e k-means and ISODATA for the crop identification from the remote sensing image.The experimental analysis is perfomed using ENVI tool. The color composite mappings associated with the image is also studied.

I. INTRODUCTION

Agriculture country to be a fundamental instrument for sustainable development and poverty reduction, especially in developing country as they have a significant agrarian component in their economics.

Presently, contribution of agriculture about one third of the national GDP and provides employment to over seventy percent of Indian population in agriculture and allied activities. Therefore, our country's development largely depends upon the development of agriculture. The agricultural production information is very important for planning and allocation of resources to different sectors of agriculture.

The information on crop acreage estimation is backbone of Agricultural statistical system, if area has a strong inter-annual variability while yield remains relatively stable. Reliable and timely information on crop area is of great importance to planners and policy makers for efficient and timely agricultural development and making important decisions with respect to procurement, storage, public distribution, export, import and other related issues. India possesses an excellent administrative setup, which has long standing tradition of generating quality information. However with more emphasis on local area planning, there is further need for crop area with respect to different varieties grown in the area, irrigation availability, the soil type etc. which can go a long way in rapid development of the region.

Since the earliest stages of crop classification with digital remote sensing data, numerous approaches based on applying supervised and unsupervised classification techniques have been used to map geographic distributions of crops and characterize cropping practices. Depending on geographic area, crop diversity, field size, crop phenology, and soil condition, different band ratios of multispectral data and classifications schemes have been applied.

II. LITERATURE SURVEY

Nell is (1986), for example, used a maximum likelihood classification approach with Landsat data to map irrigated crop area in the U.S. High Plains. Price et al. (1997) further refined such approaches, using a multi-date Landsat Thematic Mapper (TM) dataset in southwest Kansas to map crop distribution and USDA Conservation Reserve Program (CRP) lands in an extensive irrigated area. Thenkabail et al. (2004) performed rigorous analysis of hyper spectral sensors (from 400 to 2500 nm) for crop classification based on data mining techniques consisting of principal components analysis, lambda–lambda models, stepwise Discriminant Analysis and derivative greenness vegetation indices. Through these analyses they established 22 optimal bands that best characterize the agricultural crops. By increasing the number of channels beyond 22 bands, accuracies only increased marginally up to 30 bands and became asymptotic beyond that number. In comparison to Landsat Enhanced Thematic Mapper data and other broadband sensors, these hyper spectral approaches increased accuracy for crop classification from 9 to 43%.

Relative to crop condition, some remote sensing studies have focused on individual physical parameters of the crop system, such as nutrient stress or water availability as variables in analyzing crop health and yield. Other research has focused more on synoptic perspectives of regional crop condition using remote sensing indices. At the same time, some researchers (Seidl et al. 2004) have demonstrated that such approaches can be limited for crop yield and health monitoring given satellite over flight timing in the context of the crop calendar.

The normalized difference vegetation index (NDVI), vegetation condition index (VCI), leaf area index (LAI), General Yield Unified Reference Index (GYURI), and temperature crop index (TCI) are all examples of indices that have been used for mapping and monitoring drought and assessment of vegetation health and productivity (Doraiswamy et al. 2003, Ferencz et al. 2004, Prasad et al. 2006). Wang et al. (2005), for example, used satellite remote sensing of NDVI to provide characterizations of landscape level patterns of net primary productivity within the U.S. Great Plains, and Kogan et al. (2005) used vegetation indices from Advanced Very High Resolution Radiometer (AVHRR) data to model corn yield and early drought warning in China.

Hadria et al. (2006) provides an example of developing leaf area indices from four satellite scenarios to estimate distribution of yield and irrigated wheat in semi-arid areas. Zhang et al. (2005) have also modified leaf area indices based on MODIS (MODerate resolution Imaging Spectrometer) using a climate-variability impact index (CVII) related to contributions to monthly anomalies in annual crop growth. Using MODIS the researchers were able to establish the relationship between CVII and LAI to accurately model regional crop forecasts.

In addition, to refine regional crop forecasting, researchers have modified standard NDVI approaches using crop yield masking. This technique involves restricting analysis to a subset region's pixels rather than using all the pixels in the scene. According to work by Kastens et al. (2005) yield correlation masking is shown to have comparable performance to cropland masking across eight major U.S. crop forecasting scenarios. Jensen (2007) provides further examples of the broad range of popular vegetation indices used in remote sensing of agricultural systems.

Recent research has documented radar as a tool for crop monitoring. Chen and Mcnairn(2006) used radar, for example, in rice monitoring within Asia. In their work, they found that backscatter increases significantly during short periods of vegetation growth, which can be used to differentiate rice fields from other land cover.

Recent commercial satellites with fine spatial resolution have also proven of value in mapping crop growth and yield. Yang et al. (2006) used Quick Bird satellite imagery for mapping plant growth and yield patterns within grain sorghum fields as compared with airborne multispectral image data. The results suggest Quick Bird data and airborne spectral data were equally useful. Although satellite remotely sensed data have historically been used for assessing specific crop stress parameters, such as indications of nitrogen stress (Reyniers and Vrindts 2006), more detailed insights regarding crop condition are being gained using hyperspectral remote sensing, thermal radiometers, and related devices. Such approaches have also contributed to effective uses of these data in precision agriculture (Yang et al. 2004). Hyper-spectral sensors and related techniques can also be used toestimate various other crop biophysical and biochemical parameters, such as leaf nitrogen content, leaf chlorophyll content, and associated factors related to soil moisture (Goel et al. 2003). Ye et al. (2006) used hyper-spectral images to predict tree crop yield in citrus groves, Vijaya-Kumar and colleagues (2005) used an infrared thermometer and spectral radiometer for screening germplasm and stress in castor beans, while Nicholas (2004) used visible, near infrared, and thermal sensors to assess crop conditions.

Vegetation stages of development (phenology) are influenced by a variety of factors such as available soil moisture, date of planting, air temperature, day length, and soil condition. These factors therefore also influence plant conditions and their productivity. For example, corn crop yields can be negatively impacted if temperatures are too high at the time of pollination. For this reason, knowing the temperature at the time of corn pollination could help forecasters better predict corn yields.

McMaster (2004) has summarized the importance of phenology, asserting that, 'Understanding crop phenology is fundamental to crop management, where timing of management practices is increasingly based on stages of crop development. Simulating canopy development is also critical for crop growth models, whether to predict the appearance of sources and sinks, determining carbon assimilation and transpiration, partitioning carbohydrates and nutrients, or determining critical life cycle events such as anthesis and maturity.' During the era during which Landsat was temporarily privatized (see also Goward et al., in this volume) after the 1990s, some remote sensing scientists begin experimenting with the use of coarse spatial resolution (1.0 km) imagery (also see Justice et al., in this volume). These data were free, and they allowed for a synoptic view of the conterminous U.S. on a daily to bi-daily temporal frequency. As methods were developed for cloud removal or minimization, new high temporal (weekly to biweekly) cloud-free or near cloud-free datasets were made freely available for a variety of applications and research endeavors. In the early 1980s, Badhwar and Henderson (1981) published a paper in the Agronomy Journal describing the use of spectral data for characterizing terrestrial vegetation development. In 1990, Lloyd described the use of a shortwave vegetation index (the Normalized Difference Vegetation Index (NDVI)) for

characterizing phenological stages of plant development for terrestrial land cover, and in 1994, Reed et al. described how high temporal resolution NDVI datasets could be used to examine variability in interannual phenology at the continental scale. Using coarse spatial resolution and high temporal resolution data, for the first timeplant response to varying growing conditions could be examined at or near a continental scale, for example, the conterminous US. Since the publication of the manuscripts referenced above, many studies that followed used AVHRR NDVI datasets for such tasks as assessing crop relative condition and making yield forecasts (Steven et al. 1983, Quarmby et al. 1993, Groten 1993, Kastens et al. 2005), characterization of Central Great Plains grass life forms (Reed et al. 1996, Tieszen et al. 1997), studying vegetation response to intra- and interannual climatic variation (Yu et al. 2003, Breshears et al. 2005), and for many other agriculturally related purposes.

Crop biophysical characterization Quantitative techniques can be applied to the spectral data, whether acquired from close-range or by aircraft or satellite-based sensors, in order to estimate crop status/condition. The technology is capable of playing an important role in crop management by providing at least the following types of information:

1. fraction of vegetative cover,
2. chlorophyll content,
3. green leaf area index, and,
4. other measurable biophysical parameters
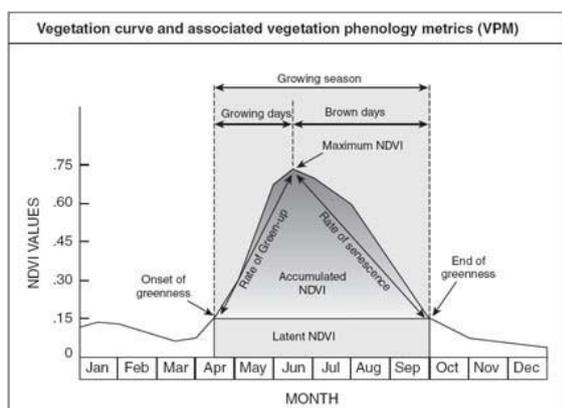
Figure 1. Crop biophysical characterization

A twelve month, hypothetical NDVI temporal response curve for vegetation. Additionally, the vegetation metrics are displayed to show their relation to both NDVI values and time (after Reed et al. 1994). ENVI 4.4 ENVI combines advanced spectral image processing and proven geospatial analysis technology with a modern, user-friendly interface. Whether you use panchromatic, LiDAR, SAR, multispectral or hyperspectral imagery, ENVI has the latest processing and analysis tools to help you extract meaningful information to make better decisions.

ENVI 4.4 introduces the - a robust suite of tools designed to access quickly
And perform common image processing tasks. This reduces the number of steps and produces specific geospatial information products, and easily allows:

➢ Perform change detection
➢ Assess vegetation delineation and stress detection
➢ Produce spectral analogues to find similar materials
➢ Pan-sharpen images
➢ Detect lines of communication
➢ Find watercraft
➢ Measure relative water depth

Classification Methods:
The classification techniques are divided into two types i.e supervised and unsupervised classification.

Supervised classification:
This needs the background knowledge of study area obtained by external sources or field work. The training fields are used to train the computer in recognizing different categories, The selection of training fields is based on optical properties of each category, field observations at known locations, Ariel and ground photographs. Interpreter inputs single pixel, the algorithm adds up the neighbor pixels with similar pixel values with one initially selected.

Unsupervised Classification:
This needs no prior knowledge about study area. It classifies the pixels with spectral classes by examining the pixels of similar DLs. The clusters with similar spectral characteristics are grouped together. It is based on selecting discrimination space and clustering algorithms.

1. K- Means clustering:
1) A set number of cluster centres are positioned randomly through the spectral space.
2) Pixels are assigned to their nearest cluster.
3) The mean location is re-calculated for each cluster.
4) Repeat 2 and 3 until movement of cluster centres is below threshold. Assign class types to spectral clusters

2. ISO- DATA - The Iterative Self-Organizing Data analysis technique (ISODATA):
☐ Extends k-means. Also calculate standard deviation for clusters.
☐ After stage 3 we can either:
– Combine clusters if centres are close.
– Split clusters with large standard deviation in any dimension.

- Delete clusters that are to small.
- Then reclassify each pixel and repeat.
- Stop on max iterations or convergence limit.
- Assign class types to spectral clusters.

III. RESULTS AND DISCUSSION

The crop identification goes through following steps in ENVI tool.

1] Pre- processing step:
This involves the geometric and radiometric correction of available image. Then it is further processed for various filtering methods. The image is processed under various filtering techniques like low pass, high pass, Gaussian, Median and Laplacian. From observations it is seen that Gaussain Equalization provides the best clarity of the original image.
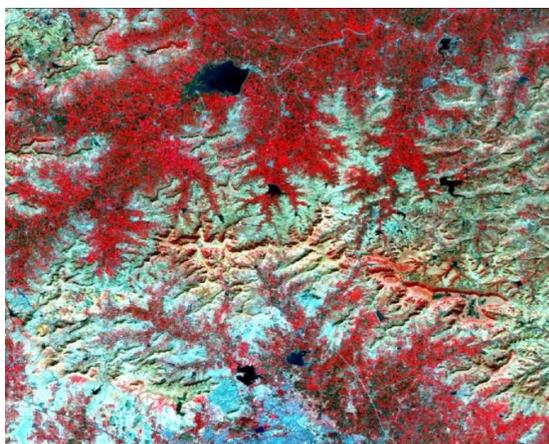
Figure.2 Original Image
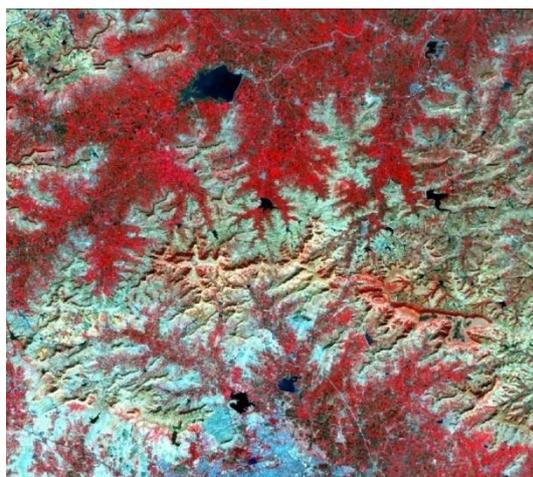
Figure.3 Filtered Image

2] Classification:
The image is processed through ISO-DATA and K-means classification are the results are given as follows:
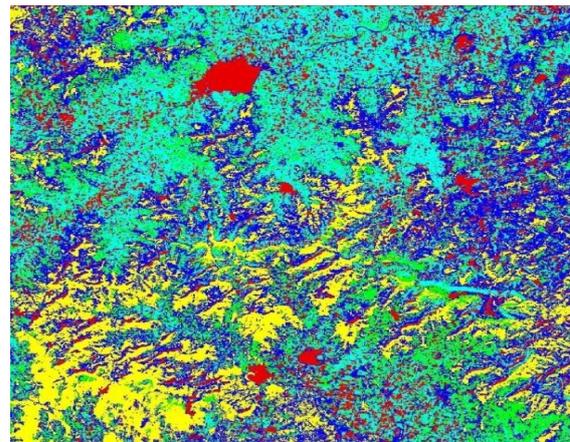
Figure.4 ISO- data classification
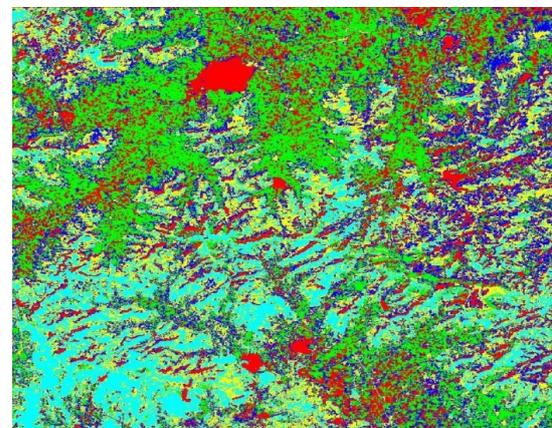
Figure. 5 K- Mean's classification

3] Observations and statistical recordings: The observations and statistical recordings performed using ENVI tool using Unsupervised classification. The table 1. below gives the classification results for the both methods.

Table 1. Class Distribution Summary for the ISODATA and K-Means classification

	ISODATA Classification	**K means**
Water body	241,763 points (13.395%) (96,705,200.0000 Meters²)	286,353 points (15.866%) (114,541,200.0000 Meters²)
crops:	409,378 points (22.683%) (163,751,200.0000 Meters²)	448,654 points (24.859%) (179,461,600.0000 Meters²)
hilly area	246,273 points (13.645%) (98,509,200.0000 Meters²)	405,176 points (22.450%) (162,070,400.0000 Meters²)
barren	370,143 points	387,725 points

land	(20.509%) (148,057,200.0000 Meters²)	(21.483%) (155,090,000.0000 Meters²)
hilly area with crops	**537,253 points (29.768%) (214,901,200.0000 Meters²)**	**276,902 points (15.342%) (110,760,800.0000 Meters²)**

IV. CONCLUSION

The paper studies the area of agriculture with the usage of remote sensing. It reviews the various papers and the techniques available for the specified area.The experimental analysis is performed using ENVI tool. The unsupervised classification method i.e ISO-DATA and K-Means are performed for the crop classification from the image. The statistical recordings of the image shows that the k-means gives the best result for the crop classification from the image.

REFERENCES

[1] Sebastian Mader, Michael Vohland, Thomas Jarmer and Markus Casper. CROP CLASSIFICATION WITH HYPERSPECTRAL DATA OF THE

[2] HYMAP SENSOR USING DIFFERENT FEATURE EXTRACTION

[3] TECHNIQUES. Proceedings of the 2nd Workshop of the EARSeL SIG on Land Use and Land Cover.

[4] Goswami S. B.1, Dr. Aruna Saxena2 and Dr. G. D. Bairagi. Remote Sensing and GIS based wheat crop acreage estimation of Indore district, M.P. International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 3, March 2012)

[5] M. Duane Nellis, Kevin P. Price, and Donald Rundquist. Remote Sensing of Cropland Agriculture. University of Nebraska – Lincoln Papers in Natural Resources School of Natural Resources, 1-1-2009

[6] Md. Rejaur Rahman* A.H.M. Hedayutul Islam** Md. Ataur Rahman. NDVI Derived Sugarcane Area Identification and Crop Condition Assessment

[7] http://www.geospectra.net/lewis_cl/remote/remote.htm

[8] CARLOS ANTONIO OLIVEIRA VIEIRA , PAUL MATHER, PAUL APLIN . AGRICULTURAL CROP CLASSIFICATION USING THE SPECTRAL-TEMPORAL RESPONSE SURFACE. Anais XI SBSR, Belo Horizonte, Brasil, 05-10 abril 2003, INPE, p. 255-262.

[9] Krishna Bahadur K.C. Improving Landsat and IRS Image Classification: Evaluation of Unsupervised and Supervised Classification through Band

[10] Ratios and DEM in a Mountainous Landscape in Nepal.

[11] Remote Sens. 2009, 1, 1257-1272; doi:10.3390/rs1041257