ISSN:0975-9646

Sandip A. Kharat et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (5) , 2015, 4381-4384

Cotton Crop Discrimination Using Landsat-8 Data

Sandip A. Kharat, Vijaya B. Musande

Dept. of Computer Science & Engineering
MGM's Jawaharlal Nehru Engineering College, N-6 CIDCO, Aurangabad 431003
Dr. Babasaheb Ambedkar Marathwada University, Aurangabad (MH), India

Abstract- The classification and recognition of agricultural crop types is an important application of remote sensing. This paper includes the approach and technique of Remote Sensing (RS) based single crop identification, based on multispectral temporal data. In this study multispectral time series images of Landsat-8 has been used for identification of cotton crop for Aurangabad region (MH) in India. The pixel based Unsupervised K-Means classification technique is used for discriminate the land cover distributions. Accuracy and efficiency of the pixel based classification technique is compared using kappa statistics and confusion matrix.

Keywords: Crops Classification, Multispectral, Confusion matrix, Kappa Coefficient.

I. INTRODUCTION

Agriculture is backbone of Indian economy providing livelihood to 67% population and contributing approx 35% to Gross National Product. So keeping track of agricultural information is essential, remote sensing systems with their synoptic viewing capability and variety of temporal and spatial resolution helps in the same. Remote sensing methods are superior to conventional methods since it is fast and economic. Remote sensing plays significant role in agriculture and crop management applications such as crop inventory, crop production forecasts, drought, flood damage assessment and crop classification [1]. We have mainly emphasized on crop classification using multispectral temporal data. In crop classification geographic area, crop diversity, field size, crop phenology and soil condition plays important role.

Cotton crop is main cash crop in India. For an efficient production and management of cotton crop up to date information is needed. An early detection of crop plays an important role and helps policy maker in finding acreage, crop yield production if further goes in depth it helps in stress detection, crop disease identification etc.

II. LITERATURE REVIEW

Literature exploited supervised as well as unsupervised classification of multispectral images. The multispectral airborne as well as satellite remote sensing technologies have been utilized as a widespread source for the purpose of remote classification of vegetation. During the literature it is found that for crop classification mostly exploited vegetative indices are NDVI and TNDVI [2]. These VI's has its effect on temporal data. Vyas et al. recently proposed a multi temporal crop type classification that successfully classified crops in India [3].They have worked on multiple crops with multi date data and compare the result with single date. Multi temporal data improve the accuracy of classification. Sujay Datta et al. have used LISS-I Data for wheat crop classification by combining two dates data using PCA (Principle Component Analysis) & derived there first two principle components they have got 94% classification accuracy [4]. V.B.Musande et.al has worked on cotton crop classification using fuzzy approach; the image to image maximum classification accuracy observed was 96.5%, data used for this study was AWIFS for soft classification and LISS-III data for soft testing [5]. Md. Rejaur Rahman et.al used IRS LISS II digital data and NDVI to identify the sugarcane area and its condition assessment, the accuracy achieved was 85.25% [6]. The literature also emphasized on finding best classification approach for image segmentation. Xiaofang Liu and Xiaowen Li used dot density weighted fuzzy c-means clustering (WFCM) to overcome the limitation of FCM (i.e. equal partition trend for data set.) [7]. Comparisons of different classification algorithms in the multi-date classification category have been extensively studied. For example, Chan et al (2001) compared four classifiers, namely Multi-Layer Perceptron (MLP), Learning Vector Quantization (LVQ), Decision Tree (DT) and Maximum-Likelihood Classifier (MLC). Seto and Liu (2003) compared ARTMAP neural network with MLC and observed that ARTMAP neural network classifiers were more accurate than MLC classifiers. It is hard to make a conclusion that some classifiers are always better than the rests when multiple criteria are used to evaluate the suitability of algorithms. The proposed work uses unsupervised K-Means classifier for crop discrimination.

III. PROPOSED METHOD

This section provides proposed crop classification system using pixel based K-Means unsupervised Classifier. The use of this technique for specific crop identification follows a sequence of steps which are explained as below. The flow chart of the proposed methodology is given in figure 1.

The overall sequence of the proposed methodology is given as follows:

1) Acquisition of remote sensing image from multispectral source
2) Image processing
3) Sample set selection
4) Image classification for various land cover features
5) Accuracy assessment of classified image using ground truth data/Field information.

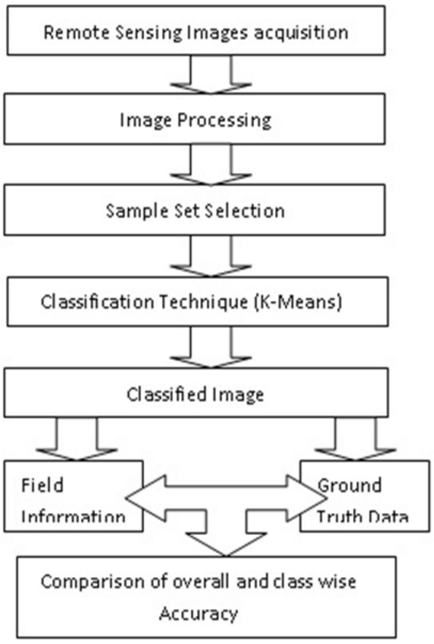

Figure 1: Flow Chart of Proposed Methodology

A. Multispectral Image Acquisition:

For the crop identification using remote sensing, remote sensing images are needed to acquire. In our case we are only concerned with multispectral sensors hence we can consider the images from available satellite sensors like AWIFS, LISS (IRS series), SPOT 5 and also LANDSAT, MODIS which are good sources of multispectral data. But here in this study we have used Landsat-8 images of Aurangabad region.

B. Image Processing :

In this phase we can enhance and restore the image. Image enhancement may include contrast stretching, edge enhancement, etc. While in image restoration we can consider the geometric correction, radiometric correction as per the need of acquired image. Once the data is made ready and georeferenced, it would be pre-processed using various indices and then applied to the classifier.

C. Sample Set Selection:

Here in this phase we have considered the temporal data of specific region (Mali Sagaj) and nearby villages related with cotton crop and evaluated the variation in accuracy over a period of time.

D. Classification Technique:

K-MEANS CLUSTERING ALGORITHM

K-means is one of the basic clustering methods introduced by Hartigan [8]. This method is applied to segment the remote sensing image in recent years. The K-means clustering algorithm for classification of remote sensing image is summarized as follows:

Algorithm K-means(x, n, c)

Input:

N: number of pixels to be clustered; x={x1, x2, x3… xN}: pixels of remote sensing image

c = {c1, c2, c3… cj}: clusters respectively.

Output:

Cl: cluster of pixels

Begin

Step 1: cluster centroids are initialized.

Step 2: compute the closest cluster for each pixel and classify it to that cluster, i.e. the objective is to minimize the sum of squares of the distances given by the following:

$$\Delta ij = \| xi - cj \| .\ \arg\min \sum_{i=1}^{N} \sum_{j=1}^{C} \Delta ij^{2} \qquad (1)$$

Step 3: Compute new centroids after all the pixels are clustered. The new centroids of a cluster is calculated by the following

$$cj = -\sum xi \text{ , where xi belongs to cj} \qquad (2)$$

Step 4: Repeat steps 2-3 till the sum of squares given in equation is minimized.

End

E. Accuracy Assessment:

An accuracy assessment of classification is undertaken using confusion matrices and Kappa statistics. The accuracy of the classified image was the assessed using a range of reference data including field data collected in the study area during the seasonal period. Producer and user accuracies for each class were calculated along with the overall accuracies and Kappa statistics.

IV. STUDY AREA AND DATA USED

The study area chosen for the present study is located between 19^{0}31'57.03" N - 19^{0}57'45.74"N latitude and 74^{0}49'44.46" E - 75^{0}16'9.37" E longitude in the state of Maharashtra, India. The important places in study are Mali Sagaj, Vaijapur and Aurangabad as shown in Figure 2. Bajara, Jawar, Soyabin crops are grown in this region in Kharip Season. Cotton is the main crop in this region having homogeneous field.

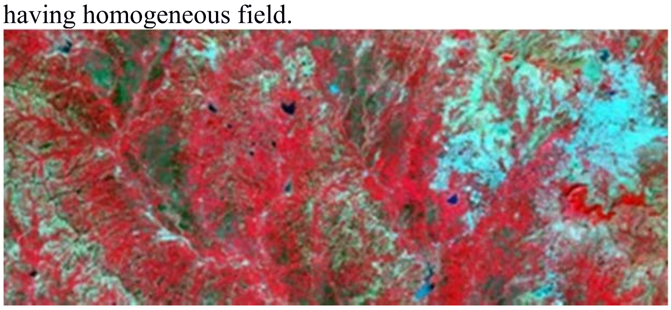

Figure 2: Location of study Area

In this study remotely sensed multispectral images from Landsat-8 are considered for cotton crop identification. Landsat-8 has 11-bands with different resolutions given as follows:

Band (1 to 7) and Band 9 =30m.

Band 8 (Panchromatic) =15m.

Band 10 and Band 11(TIRS) =100m.

But study has considered only three bands namely Green (B3), Red (B4) and NIR (B5) for classification purpose which has spatial resolution of 30 m. The details of datasets used for the study are shown in Table I.

TABLE I: Data Set Used

Date	Date Number	Phenological Stage
21-Sep-2014	1	Leaves formation
07-Oct-2014	2	Flowering Stage.

In this study the field visit is carried out on 16-17, November 2014 to collect ground truth data so that the exact positions of training and testing Cotton fields could identify.

V. RESULT AND DISCUSSION

Color composite image extracted from the Landsat-8 reveals the crops and other cover types in this area Figure 3(a). Figure 3(b), (c) is classification map of study image on two different dates.

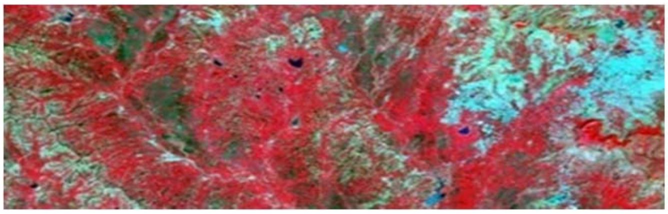

Figure 3(a): Various land cover types in study Area.

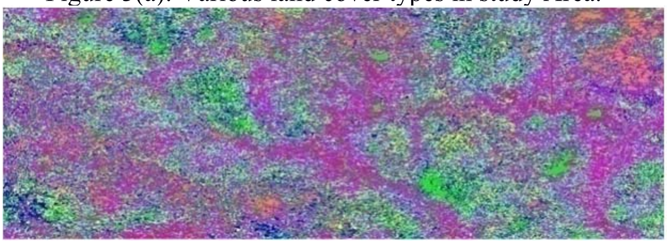

Figure 3(b): classification map image Dated 07-10-14 (at K=10)

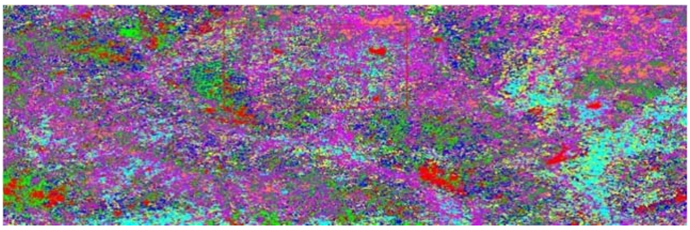

Figure 3(c): classification map image dated 08-11-14 (at K=10)

Visual comparison with color composites indicates classification maps provided very little separation between various cover types due to pixel based classification approach. The image is classified using K-means clustering for different values of K. According to D.T. Pham et.al for value K>9 results converges [9]. So result obtained in the work are consider for (K=10, 11).

The accuracy assessment of classified images is mapped with ground truth images and classification accuracy statistics including overall accuracy, producer's accuracy and user's accuracy were calculated based on the confusion matrices.

Among the classified temporal images we have got both accuracies (i.e. Producer's & User's) good (>82 %). The image taken on 07-10-2014 provided overall 98.81% accuracy with Kappa coefficient 0.9801 and image taken on 08-11-2014 has provided overall 96.18% accuracy with Kappa coefficient 0.9436 for K=10. The confusion matrices shows number of pixels classified by classifier with respect to ground truth data as shown in Table II for both the dates at different values of K.

TABLE II: ERROR MATRICES AND ACCURACY MEASURES FOR CLASSIFICATION MAPS FOR STUDY SITE

(a) Result of image 07-10-2014 (For K=10)

Actual category (Ground Truth)

Classified Category	*Water Body*	*Cotton*	*Residential Area*	*Dry Veg.*	*Total*	*User's Accuracy (%)*
Water Body	95360	0	0	0	95360	100
Cotton	0	186025	0	39344	225369	82.54
Residential A.	2179	0	33925	0	36104	93.96
Dry Veg	0	0	0	222234	222234	100
Total	97539	186025	33925	261578	579067	--
Producer's Accuracy (%)	97.77	100	100	84.96	Overall Accuracy (%)	92.82

(b) Result of image 07-10-2014 (For K=11)

Actual category (Ground Truth)

Classified Category	*Water Body*	*Cotton*	*Residential Area*	*Dry Veg.*	*Total*	*User's Accuracy (%)*
Water Body	116463	0	0	0	116463	100
Cotton	0	108386	0	20045	128431	84.39
Residential A.	2275	0	26144	0	28419	91.99
Dry Veg	0	0	0	205126	205126	100
Total	118738	108386	26144	225171	478439	--
Producer's Accuracy (%)	98.08	100	100	91.10	Overall Accuracy (%)	95.33

Result of image 08-11-2014 (For K=10)

Actual category (Ground Truth)

Classified Category	*Water Body*	*Cotton*	*Residential Area*	*Dry Veg.*	*Total*	*User's Accuracy (%)*
Water Body	107376	0	0	0	107376	100
Cotton	0	57692	0	0	57692	100
Residential A.	0	0	144431	22272	166703	86.64
Dry Veg	0	0	467	263249	263716	99.82
Total	107376	57692	144898	285521	595487	--
Producer's Accuracy (%)	100	100	99.68	92.20	Overall Accuracy (%)	96.18

(c) Result of image 08-11-2014 (For K=11)

Actual category (Ground Truth)

Classified Category	*Water Body*	*Cotton*	*Residential Area*	*Dry Veg.*	*Total*	*User's Accuracy (%)*
Water Body	96187	0	0	0	96187	100
Cotton	3	36829	0	0	36832	99.99
Residential A.	0	0	106338	53857	160195	66.38
Dry Veg	0	0	0	224742	224742	100
Total	96190	36829	106338	278599	517956	--
Producer's Accuracy (%)	100	100	100	80.67	Overall Accuracy (%)	89.60

REFERENCES

[1] Vijaya Musande, Anil Kumar and Karbhari Kale, "Effects of Indices on Temporal Data for Specific Crop Identification by using Possibilistic-*c* means" November, 2011.

[2] Sujay Dutta, N K Patel, T T Medhavy, S K Srivastava, Naveen Mishra and K R P Singh, "Wheat Crop Classification Using Multidate IRS LISS-I Data", Journal of the Indian Society of Remote Sensing, Vol. 26, No. l & 2, 1998.

[5] Zhang Miao, Li Qiangzi, Wu Bingfang, "Investigating the capability of multi-temporal Landsat images for crop identification in high farmland fragmentation regions". *IEEE 2012 First International Conference on Agro-Geoinformatics* - Shanghai, China.

[6] Erik Zillmann, Horst Weichelt, "Crop identification by means of seasonal statistics of RapidEye time series", *IEEE 2014 Third International Conference on Agro-Geoinformatics* - Beijing, China.

[7] Rabindra K. Panigrahy, S. S. Ray, S. Panigrahy, "Study on the Utility of IRS-P6 AWiFS SWIR Band for Crop Discrimination and

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

[4] Sujit Kumar BALA, Mohammad ALI, A.K.M. Saiful ISLAM, "Estimation of Potato Yield In and Around Munshigonj Using Remote Sensing and Gis Techniques", *International Conference on Water & Flood Management (ICWFM-2007).*12 - 14 March 2007, Dhaka, Bangladesh.

Classification", *J. Indian Soc. Remote Sensing.* (June 2009) 37:325–333.

[8] Babawuro Usman,"Satellite Imagery Land Cover Classification using K-Means Clustering Algorithm: Computer Vision for Environmental Information Extraction", *Elixir Comp. Sci. & Engg. 63 (2013) 18671-18675.*

[9] D T Pham, S S Dimov, and C D Nguyen," Selection of K in K-means clustering", *Proc. IMechE Vol. 219 Part C: J. Mechanical Engineering Science.*