International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 3 Issue 12, December-2014

Change Detection in VI for Identification of Cotton Crop using Fuzzy C-Mean Clustering A Review

Ankush V. Anpat[1,]
M.E Student,
Department of Computer Science & Engineering,
Everest Educational Society's Group of Institutions
College of Engineering & Technology,
Aurangabad (M.S), INDIA-431003 [1]

Rajesh. A. Auti[2]
Professor & Head,
Department of Computer Science & Engineering,
Everest Educational Society's Group of Institutions
College of Engineering & Technology,
Aurangabad (M.S), INDIA-431003 [2]

Seema Singh Solanki[3]
Professor,
Department of Computer Science & Engineering,
Everest Educational Society's Group of Institutions
College of Engineering & Technology,
Aurangabad (M.S), INDIA-431003 [3]

Abstract **:The classification of cotton crop is an important task because the automated identification of cotton crop is use full for the identification of exact cultivation area of cotton crop in particular area because the cotton crop is one of the one of the important cash crop so the identification is important, Due its economic significant advantages change detection in the different vegetation indices such as SR(simple ration), NDVI (normalized vegetation indices),TNDVI (Transferred normalised Vegetation indices),SAVI(soil adjusted Vegetation Index), and TVI (Triangular Vegetation Indices) is challenging problem, In this paper we are represented the different vegetation indices to improve the effectiveness but in remotely sensed dataset to find out the accurate cotton crop is very challenging task . In proposed paper change detection to be investigated to extract cotton crop area from LISS-III dataset. In proposed methodology, fuzzy c-mean clustering was used to identify the change detection in vegetation indices to experiment this five LISS-III satellite image was taken and experiment done with six vegetation indices. In the method firstly vegetation indices are computed, then c-mean clustering is applied to the vegetation indices to segment the image to identify the cotton crop, c-mean clustering find the particular pixels belonging to the cotton crop. Along these green leaf area also considered for investigation, the general objective of this study is to Comparison and accuracy assessment of different vegetation indices for identification of cotton crop using fuzzy c-mean clustering for classification of image**

Keywords: **Vegetation indices, K-mean classifier, White balance colour, Principal component analysis (PCA),Fuzzy c-mean clustering.**

I. INTRODUCTION

The ratio of leaf surface area to unit ground surface area, called as leaf area index (LAI)(Breda,2003)[1],describe the potential surface area available for leaf gas exchange between the atmosphere and the terrestrial biosphere(cowling and field,2003). Therefore it is important for much biological and environmental change detection and for the crop and forest condition is agriculture production and forest area is in good condition or any kind of loss has been occurred in the forest area of in agriculture area we can say by calculating the various vegetation indices and GreenLAI. The different kind of vegetation indices area available for identifying the different vegetation like agriculture area, forest area, grass area, by using classification approach we can able to identify the required area on which we want to make processing for identifying the crop condition of forest condition that specific area we can capture by using the classification approach. Image change detection is a process that analyse LAI can be separated into its photosynthetic and non-photosynthetic components. The portion of LAI composed of green leaf area (i.e. Green LAI) is the photo synthetically functional component.

For estimating the GreenLAI one main approach has been used in this paper i.e. nothing but: 1) empirical relationships between Green LAI and spectral vegetation indices (chen and cihlar, 1996; Curran, 1983a,b; Jordan 1963). Spectral vegetation indices are mathematical combination of different spectral bands mostly in the visible and near infrared regions of the electromagnetic spectrum. These numerical transformations are semi-analytical measures of vegetation activity and have been widely shown to vary not only with the seasonal variability of green foliage. The main purpose of vegetation indices is nothing but the to enhance the information contained in spectral reflectance data by extracting the variability due to vegetation characteristic e.g GreenLAI and vegetation indices the vegetation indices provide the most convenient approach for extracting the information from the remote data for different kind of application like agriculture production, forest condition or climate change detection using the different vegetation indices of GreenLAI. Application of vegetation indices has ranged from leaf to global level and in the case of Green LAI some successes have been obtained for different crops (Boegh et al., 2002; Broge and Mortensen, 2002; Clevers, 1989). Classification of remote sensing imagery has always been a challenge for the analyst as it involves various aspects which are to be considered, such as type of information, spatial resolution, spectral resolution, temporal resolution etc. Different techniques have been evolved to cater the varying needs and demands. The simplest scenario is in assigning a unique class to pixel and this is readily available to the analyst. However, with the improvement in spatial and spectral resolution, the demands have changed significantly. The requires more intensive classification algorithms and production of tactical and strategic decision making by all stakeholders in agriculture, such as producer, processors, resource managers, marketing , finance and the government is well known. Suppose any of user interested in specific crop then by using the vegetation indices or by classification method we can extract the area in which user is interested and by classifying that area of interest he can make sale and purchase of crop or whatever interested crop using these technique we working on the principal of these idea to extract the area of interested crop in agriculture area for identifying the crop condition like the crop growth is in good condition or not by using the different the classification approach in which K-mean classifier, nearest neighbour classifier ,principal component analysis are the some classifier are available for identifying he crop area in agriculture field or for finding crop in good manner we used the different vegetation indices and GreenLAI for prediction of crop condition the growth of crop is going in good manner or not or rather than crop we consider the forest area is forest area is in good manner or not or the foliage area of the forest is in good condition or not we can calculate using different vegetation indices and using GreenLAI these two approach are important for calculating the crop condition.

Te-Ming et al., 2009 has evaluated the NDVI approach which has been widely utilized with various sensors to enhance and extract vegetation information. However, the characteristics from different sensors would affect the desired results. Meanwhile, threshold selection is the key issue for extracting vegetation information from various scenes. However, the saturation effects of the proposed (Vegetation index) VI with respect to the vegetation abundance, the level of VI correlation with leaf area index, biomass, and other vegetation quantification field methods up to a threshold are to be investigated[2]. In this paper we study the different vegetation indices which available for extracting the feature from the spectral image of satellite and comparing the result of different vegetation indices with GreenLAI for concluding the crop condition we also have to identify the crop using different classification approach we have to first identify the study area using classification method and then calculate the different vegetation indices on which and compare the GreenLAI that different vegetation index.

This paper is divided into four sections. Section-II describes the different vegetation indices and Green leaf area index. Section-III describes the various classification approach. The last section presents the conclusion.

B.

II. VEGETATION INDICES

A. *Vegetation indices:*

To enhance the vegetation signal in remotely sensed data and provide an approximate measure of green vegetation amount, a number of spectral vegetation indices have been proposed. By combining the data from various bands into single band because they cover the land they correlate the biophysical characteristics of the vegetation of the land cover from the satellite spectral signal [2].

The remote sensing is the use of albedo band ratio for the vegetation calculation Jordan (1969) the different vegetation indices has been used from that first normalised difference vegetation index(NDVI), ratio vegetation index(RVI), for calculation the high biomass using NDVI is that NDVI is saturated towards high biomass and perturbing factures another vegetation index is soil adjusted vegetation index(SAVI),triangular vegetation index(TVI), which describe the radioactive energy absorbed by the pigments as a function of the relative difference between red and near-infrared reflectance in conjunction with the magnitude of reflectance in the green region, i.e. for green leaf area index calculation of the vegetation indices transformed normalised vegetation index(TNDVI) ,chlorophyll green (CLgreen), vegetation index green (VIgreen), green normalised vegetation index(GNDVI) simple ration, green ratio vegetation index (GRVI).

B. *Green Leaf area index (GreenLAI):*

Leaf surface area to ground surface area is called as green leaf area index (GreenLAI). There are two different approach are available for measure green leaf area index that are 1) Direct method 2) Indirect method the purpose of Green leaf area index is nothing but t the photosynthetic ally active and inactive component using the vegetation indices and by using the Green leaf area index. Direct method: In case of direct method the man power required for calculation of leaf area index in this case total leaf surface area is to be measured by measuring the leaf which are spread on ground of below canopy of tree or below the crop by measuring the no of leaf under the tree we can calculate the leaf area index of that particular crop or the tree[3].Indirect method: indirect method leaf area index is calculated using transmission of radiation through canopy making use of radioactive transfer theory. For calculation of Green leaf area index of any crop or any tree we use an device like LAI 2200 such kind of devices are available by making use of these devices we can able to calculate the leaf area index of any crop or tree using inverse canopy method. Table 1 show the various indices studied in this work.

Vegetation Index	Equation	References
Simple Ration (SR)	$SR = \frac{P_{nir}}{P_{red}}$	Birth and McVey, 1968; Colombo et al.,2003
Normalized Difference Vegetation Index (NDVI)	$NDVI = \frac{P_{nir} - P_{red}}{P_{nir} + P_{red}}$	Rouse et al.,1973; Deering et al.,1975; Huete et al., 2002
Soil Adjusted Vegetation Index (SAVI)	$SAVI = \frac{(1+L)(P_{nir} - P_{red})}{P_{nir} + P_{red} + L}$	Huete and Liu, 1994; Running et al.,1994
Triangular Vegetation Index (TVI)	$TVI = 0.5(120(P_{nir} - P_{green})) - 200(P_{red} - P_{green})$	Broge and Leblanc, 2000
TNDVI	$\left[\left(\frac{P_{nir} - P_{red}}{P_{nir} + P_{red}}\right) - 0.5\right]^{\frac{1}{2}}$	Tucker, 1979

Table 1: Various indices studied in this work

III. GENERAL PRINCIPAL FOR RECOGNIZING VEGETATION:

A) Fuzzy C-Mean

Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. This method (developed by Dunn in 1973 and improved by Bezdek in 1981) is frequently used in pattern recognition

This algorithm works by assigning membership to each data point corresponding to each cluster center on the basis of distance between the cluster center and the data point. More the data is near to the cluster center more is its membership towards the particular cluster center. Clearly, summation of membership of each data point should be equal to one. After each iteration membership and cluster centers are updated accordingtotheform

$$\mu_{ij} = 1 / \sum_{k=1}^{c} (d_{ij} / d_{ik})^{(2/m-1)}$$

ula:

$$v_j = \left(\sum_{i=1}^{n} (\mu_{ij})^m x_i\right) / \left(\sum_{i=1}^{n} (\mu_{ij})^m\right), \forall j = 1, 2, \ldots.c$$

where,

'n' is the number of data points.

'vj' represents the j^{th} cluster center.

'm' is the fuzziness index m € [1, ∞]. *'c'* represents the number of cluster center. *'µij'* represents the membership of i^{th} data to j^{th} cluster center. *'dij'* represents the Euclidean distance between i^{th} data and j^{th} cluster center.

Main objective of fuzzy c-means algorithm is to minimize:

$$J(U,V) = \sum_{i=1}^{n} \sum_{j=1}^{c} (\mu_{ij})^m \left\| x_i - v_j \right\|^2$$

where,

'$||x_i - v_j||$' is the Euclidean distance between i^{th} data and j^{th} cluster center.

Algorithmic steps for Fuzzy c-means clustering

Let $X = \{x_1, x_2, x_3 \ldots, x_n\}$ be the set of data points and $V = \{v_1, v_2, v_3 \ldots, v_c\}$ be the set of centers.

1) Randomly select *'c'* cluster centers.

2) Calculate the fuzzy membership 'μ_{ij}' using:

$$\mu_{ij} = 1 / \sum_{k=1}^{c} (d_{ij} / d_{ik})^{(2/m-1)}$$

3) Compute the fuzzy centers 'v_j' using:

$$v_j = \left(\sum_{i=1}^{n} (\mu_{ij})^m x_i\right) / \left(\sum_{i=1}^{n} (\mu_{ij})^m\right), \forall j = 1, 2, \ldots.c$$

4) Repeat step 2) and 3) until the minimum *'J'* value is achieved or $||U^{(k+1)} - U^{(k)}|| < \beta$.

where,

'k' is the iteration step.

'β' is the termination criterion between [0,

'U = $(\mu_{ij})_{n*c}$' is the fuzzy membership matrix.

'J' is the objective function.

ISSN: 2278-0181
Vol. 3 Issue 12, December-2014

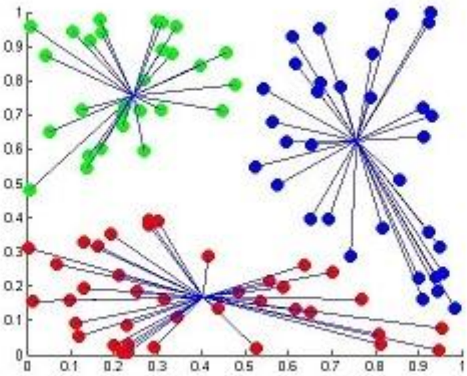

Fig 1: Result of Fuzzy c-means clustering

B) K-mean classifier:

In statistics and machine learning, ***k*-means clustering** is a method of cluster analysis is which aims to partition observations into *k* clusters in which each observation belongs to the cluster with the nearest mean. It is similar to the expectation-maximization algorithm for mixtures of Gaussians in that they both attempt to find the centers of natural clusters in the data as well as in the iterative refinement approach employed by both algorithms. In this paper for classification of vegetation components k-means is used. It is found that the output image has more good number of identified class of vegetations; it is shown in figure 1 and in table 1. The *k*-means clustering algorithm is commonly used in computer vision as a form of image segmentation. The results of the segmentation are used to aid border detection and object recognition. In this context, the standard Euclidean distance is usually insufficient in forming the clusters. Instead, a weighted distance measure utilizing pixel coordinates, RGB pixel color and/or intensity, and image texture is commonly used. In image analysis finding groups in data is very useful. We can find pixels with similar intensities i.e. automatically find regions in images. We can also find images that are similar i.e. can automatically find classes/clusters of images.[8]-[14]Various steps in the algorithm are as follows:

1. Compute the intensity distribution (also called the histogram) of the intensities.
2. Initialize the centroids with k random intensities.
3. Repeat the following steps until the cluster labels of the image does not change anymore.
4. Cluster the points based on distance of their intensities from the centroid intensities.
5. Compute the new centroid for each of the clusters.[3]

$$J = \sum_{j=1}^{k}\sum_{i=1}^{n}\left\|x_i^{(j)} - c_j\right\|^2$$

C. Principal component analysis(PCA):

Steps for PCA are as follows

1. Let an image band X(x, y) be a two dimensional mXn array (8-bit Gray Scale) of intensity values. An image may also be considering the vector of dimension mn, so that a typical image of size 112x92 becomes a vector of dimension 10304. Let the input set of image bands {X1, X2, X3… XN}. The average face of the set is defined by

$$W_i = V_i^T (X_i - \overline{X}) \quad \ldots.. \quad (1)$$

2. Calculate the covariance matrix to represent the scatter degree of all feature vectors related to the average vector. The covariance matrix C is defined by

$$C = \frac{1}{N}\sum_{i=1}^{N}(Xi - \overline{X})(Xi - \overline{X})^T \quad \ldots.. \quad (2)$$

3. The Eigenvectors and corresponding eigen values are computed by using

$$CV = \lambda V \quad \ldots\ldots (3)$$

Where V is the set of eigenvectors associated with its eigen value λ.
4. Sort the eigenvector according to their corresponding eigen values from high to low.
5. Each of the mean centered image project into eigen space using

$$W_i = V_i^T (X_i - \overline{X}) \quad \ldots\ldots \quad (4)$$

D. White balance color:

In shooting film, color balance is typically achieved by using color correction filters over the lights or on the camera lens. Color balance changes the overall mixture of colors in an image and is used for color correction; generalized versions of color balance are used to get colors other than neutrals to appear correct or pleasing.[6][7] The image has operated through white balancing and it is found that due to it the segmentation has given very good resultant image, shown in figure 4 and in table 1.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 255/R'_w & 0 & 0 \\ 0 & 255/G'_w & 0 \\ 0 & 0 & 255/B'_w \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix}$$

IV. CONCLUSION:

We have represent methodology of change detection in vegetation indices to identify the cotton crop using the fuzzy c-mean clustering ,by using c-mean clustering we have segmented the image from multispectral image of LISS-III dataset then to classify to identify the cotton crop , we have assess the accuracy of cotton crop to further identify the cotton crop disease .

REFERENCES

[1] Musande, V. Kumar, A. Kale, K., 2012. Cotton Crop Discrimination Using Fuzzy Classification Approach, Journal of Indian Society and Remote Sensing, vol. 40, no. 4, pp. 589-597.

[2] Musande, V. Kumar, A. Roy, P.S. Kale, K., 2013. Evaluation of fuzzy-based classifiers for cotton crop identification, Geocarto International, vol. 28, no. 3.

[3] Omar, M., 2011, Classification of Cotton and Sugarcane Plants on the Basis of their Spectral Behavior, Pak. J. Bot. vol. 43, no. 4, pp. 2119-2125.

[4] Ozdogan, M. Yang, Y. Allez, G. Cervantes C., 2010. Remote sensing of irrigated agriculture: Opportunities and challenges, Remote Sensing, vol. 2, no. 9, pp. 2274-2304.

[5] Rajesh, K, D. Rajendra, Y. Kale, K.V. Mehrotra, S.C., 2013. Classification of Crops from remotely sensed Images: An Overview", International Journal of Engineering Research and Applications (IJERA), Vol. 3, no. 3, pp.758-761.

[6] Rao, N. R. P. K. Garg, S. K. Ghosh,. 2007, Evaluation of radiometric resolution on land use/land cover mapping in an agricultural area, International Journal of Remote Sensing, vol. 28, no. 2, pp. 443-450.

[7] Rouse, J. W., Haas, R. H., Jr., Schell, J. A., Deering, D. W., 1974. Monitoring vegetation systems in the Great Plains with ERTS, in proceedings of Third ERTS-1 Symposium, pp. 309–317, Washington, DC: NASA.

[8] Tansey, K. Chambers, I. Anstee, I. Denniss, A. Lamb, A., 2009. Object-oriented classification of very high resolution airborne imagery for the extraction of hedgerows and field margin cover in agricultural areas. Applied Geography, Vol. 29, No. 2, pp. 145-157. .

[9] Vina, A. Anatoly, A. G. Anthony, L. Nguy, R., Yi, P., 2011. Comparison of differ area index of crops, Remote Sensing of Environment, vol. 115, pp. 3468–3478.

[10] Gitelson, A. A., Kaufman, Y. J., Merzlyak, M. N., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS, Remote Sensing of Environment, vol. 58, pp. 289–298.

[11] Gitelson, A. A., Viña, A., Arkebauer, T. J., Rundquist, D. C., Keydan, G., Leavitt, B., 2003. Remote estimation of leaf area index and green leaf biomass in maize canopies", Geophysical Research Letters, vol. 30, pp. 1248.

[12] Gitelson, A. A., Viña, A., Ciganda, V., Rundquist, D. C., & Arkebauer, T. J., 2005. Remote estimation of canopy chlorophyll content in crops, Geophysical Research Letters, vol. 32, L08403.

[13] Govender, M. C. Naiken, V. Bulcock, H., 2008. A comparison of satellite hyperspectral and multispectral remote sensing imagery for improved classification and mapping of vegetation. Water SA, Vol. 34, No. 2, pp. 147-154.

[14] Huete, A. R., Liu, H. Q., Batchily, K., vanLeeuwen, W., 1997. A comparison of vegetation indices global set of TM images for EOS-MODIS, Remote Sensing of Environment, vol. 59, no. 3, pp. 440–451.

[15] James A. S. Daniel B. C., 2002. A Comparison of Classification Methods for Large Imagery Data Sets. JSM 2002 Statistics in an ERA of Technological Change-Statistical computing section, New York City, pp. 3205-3207.

[16] Jia, K. Wu, B. Li, Q., 2013. Crop classification using HJ satellite multispectral data in the North China Plain, J. Appl. Remote Sens., vol. 7, no.1.

[17] Jordan, C. F., 1969. Derivation of leaf-area index from quality of light on forest floor", Ecology, vol. 50, pp. 663.