

Deforestation Identification Model using Satellite Images by Color Based Segmentation

*N. Suresh Kumar, Arun Manoharan

School of Computing Science and Engineering, VIT University,
Vellore, Tamil Nadu, India.

Abstract: Primarily due to the progress in spatial resolution of satellite imagery, the methods of segment based image analysis for generating and updating geographical information are becoming more and more important. This work presents a novel image segmentation based on color features with K-means clustering algorithm. The entire work is divided into two stages. First enhancement of field separation of satellite image using de-correlation stretching is carried out and then the regions are grouped into a set of five classes using K-means clustering algorithm. Using this two step process, it is possible to reduce the computational cost and to avoid feature calculation for every pixel in the image.

Keywords: Deforestation, Image Segmentation, K-Means, Satellite Image Processing, GIS.

1. Introduction

This research work is based on the real time problem of estimating the natural resources in a given district. The process of image segmentation is defined as: “the search for homogenous regions in an image and later the classification of these regions”. It also means the partitioning of an image into meaningful regions based on homogeneity or heterogeneity criteria. Image segmentation techniques can be differentiated into the following basic concepts: pixel oriented, Contour-oriented, region-oriented, model-oriented, colour oriented and hybrid. The performance of colour segmentation may significantly affect the quality of an image understanding system. The most common features used in image segmentation include shape and colour.

For comparison purpose, This work proceed with the segmentation process and then the classification, the images that need to be classified are segmented into many homogeneous areas with similar spectrum information firstly, and the image segments features are extracted based on the specific requirements of ground features classification^[1]. The color homogeneity is based on the spectral colors, while the shape homogeneity is based on the compactness and smoothness of shape.

This work presents image segmentation based on k-means clustering algorithm. The work is divided into two stages. The present work is organized as follows: Section 2 describes the data resources and software used. Section 3 describes the K-means clustering method. Section 4 describes de-correlation stretching. In section 5 the proposed method of segmentation of image based on color with K-means clustering is presented and discussed. Experimental results obtained with suggested method are shown in section 6. Finally, section 7 concludes with some final remarks. This can be very useful for municipal corporations to get an estimate of a given resources. To implement the processing techniques MATLAB is used.

2. Architecture

The algorithm is easy to implement and can be done on any platform.. The implementation is easy and can be modified for future work. The Google map image was used to produce images that were directly given as input to the MATLAB software. No modification of type conversion is required.



Figure - 1: Architecture diagram of Deforestation identification model

In earlier days the satellite has been used for finding the vegetation area. But nowadays as the technology improved the Google map helps us to capture the satellite images.

Satellite Image: The satellite image has been used for our further implementation and the image is taken from the Google Map. **Image Preprocessing:** MATLAB image processing techniques are used to detect the edges in an image and to make the image as uniform background^[6]. Subtract the background image from the original image to view the edges. The techniques are used to clearly identify the vegetation field area, wasteland, water areas and building areas. The techniques like filter process, Noise removal are used here.

Filter technique: The noises are removed from the image using noise removal techniques.

Noise removal: Noise is the result of errors in the image acquisition process that result in pixel values that do not reflect the true intensities of the real scene. There are many types of noise in MATLAB. Electronic transmission of image data can introduce noise. The noise is removed from the image using noise removal techniques.

Deforestation Identification Model: The Deforestation identification model is to find out the areas in an aerial view image by differentiating the various fields.

Earth observation: It is the gathering of information about planet Earth's physical, chemical and biological systems. It is used to monitor and assess the status of, and changes in, the natural environment and the built environment. In recent years, Earth observation has become technologically more and more sophisticated. It has also become more important due to the dramatic impact that modern human civilization is having on the Earth, the need to minimize the negative impact and the opportunities. Earth observation provides to improve social and economic well-being.

3. Data Resources and Software Used

The images are taken from the region of VIT, Vellore city, India. The entire work is carried out using MATLAB v7.1, and MS-Office.

MATLAB Image Processing Toolbox™ provides a comprehensive set of reference-standard algorithms and graphical tools for image processing, analysis, visualization, and algorithm development. You can perform image enhancement, image deblurring, feature detection, noise reduction, image segmentation, geometric transformations, and image registration. Many toolbox functions are multithreaded to take advantage of multicore and multiprocessor computers^[2].

Image Processing Toolbox supports a diverse set of image types, including high dynamic range, gigapixel resolution, embedded ICC profile, and topographic. Graphical tools let you explore an image, examine a region of pixels, adjust the contrast, create contours or histograms, and manipulate regions of interest (ROIs). With toolbox algorithms you can restore degraded images, detect and measure features, analyze shapes and textures, and adjust color balance.

4. Methodology:

4.1. K-Means clustering

K-means clustering can best be described as a partitioning method. There is many methods of clustering developed for a wide variety of purposes. K-means is the clustering algorithm used to determine the natural

spectral groupings present in a data set. This accepts from analyst the number of clusters to be located in the data.

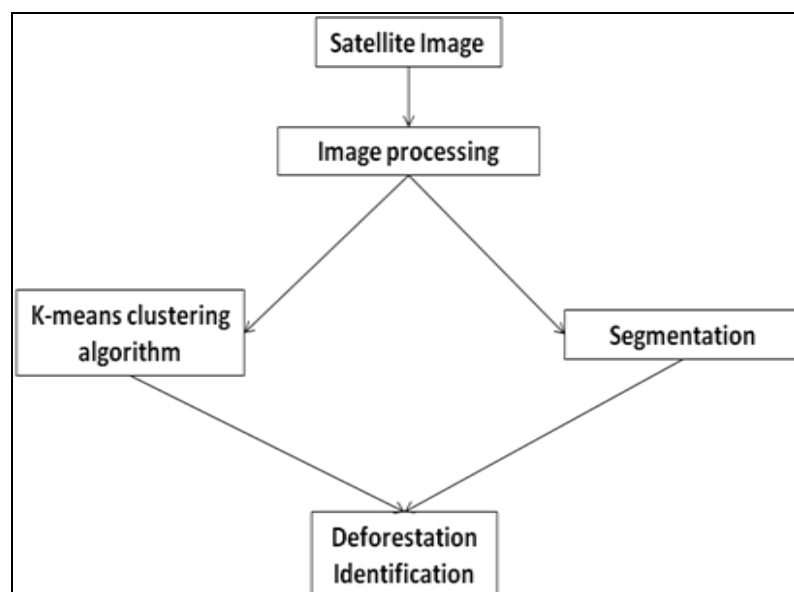


Figure - 1(a): Frame work of Deforestation Identification model

That is, the function k-means partitions the observations in your data into K mutually exclusive clusters, and returns a vector of indices indicating to which of the k clusters it has assigned each observation. K-means does not create a tree structure to describe groupings in your data, but rather creates a single level of clusters. K-means clustering uses the actual observations of objects or individuals in your data. It is more suitable for clustering large amounts of data. K-means uses an iterative algorithm that minimizes the sum of distances from each object to cluster centroid, over all clusters. This algorithm moves objects between clusters until the sum cannot be decreased further.

K-means uses an iterative algorithm that minimizes the sum of distances from each object to its cluster centroid, over all clusters. This algorithm moves objects between clusters until the sum cannot be decreased further. The result is a set of clusters that are as compact and well-separated as possible. You can control the details of the minimization using several optional input parameters to k-means, including ones for the initial values of the cluster centroids, and for the maximum number of iterations ^[3].

4.2. De-Correlation Stretching

De-correlation stretching enhances the color separation of an image with significant band-to-band correlation. The exaggerated colors improve visual interpretation and make feature discrimination easier. De-correlation stretching with the de-correlation stretch function is applied. The number of color bands, NBANDS, in the image is taken three. This is implemented using the de-correlation stretching regardless of the number of color bands. The original color values of the image are mapped to a new set of color values with a wider range. The color intensities of each pixel are transformed into the color eigenspace of the NBANDS-by-NBANDS covariance or correlation matrix, stretched to equalize the band variances, and then transformed back to the original color bands. To define the band wise statistics, this can use the entire original image, with the subset option, or any selected subset of it ^[4].

4.3. Color-Based Segmentation using K-Means Clustering

The basic aim is to segment colors in an automated fashion using the $L^*a^*b^*$ color space and K-means clustering. To describe a region of colors, for specific values of L , a specific 2D field must be defined by A and B . This is not very convenient. The $L^*a^*b^*$ color space is derived from the CIE XYZ tri stimulus values. The $L^*a^*b^*$ space consists of a luminosity ' L^* ' or brightness layer, chromaticity layer ' a^* ' indicating where color falls along the red-green axis, and chromaticity layer ' b^* ' indicating where the color falls along the blue-yellow axis ^[7]. The entire process can be summarized in following steps

Step 1: Read the image in .JPEG format, which is a fused image of part of Vellore city of Tamilnadu, India.

Step 2: Convert Image from RGB Color Space to L*a*b* Color Space.

Step 3: It creates the color transformation structure C that defines the color space conversion specified by type. To perform the transformation, pass the color transformation structure as an argument to the apply form function.

Step 4: Classify the fields Using K-Means Clustering is a way to separate groups of objects. K-means clustering treats each object as having a location in space. It finds partitions such that objects within each cluster are as close to each other as possible, and as far from objects in other clusters as possible. K-means clustering requires that you specify the number of clusters to be partitioned and a distance metric to quantify how close two objects are to each other. Use K-means to cluster the objects into three clusters Using the Euclidean distance metric and replicates.

Step 5: Label Every Pixel in the Image Using the Results from K-MEANS for every object in our input, K-means returns an index corresponding to a cluster. Label every pixel in the image with its cluster index.

Step 6: Create Images in each window which segmented clustered objects.

Using pixel labels, separate objects in image has to be done, which will result in five images.

Step 7: Then programmatically determine the index of the cluster containing the blue objects because K-means will not return the same cluster_idx value every time. This is implemented using the cluster center value, which contains the mean 'a*' and 'b*' value for each cluster.

Step 8: For color separation of an image apply the De-correlation stretching.

Algorithm:

The K-means clustering algorithm has been widely used in segmentation of color images. It takes large amount of computation time for large size image. Many algorithms use an iterative refinement technique due to its ubiquity, which is called the k-means algorithm^[5].

Given an initial set of k means $m_1^{(1)}, \dots, m_k^{(1)}$, the algorithm proceeds by alternating between two steps.

Assignment step: Assign each observation to the cluster with the closest mean.

$$S_i^{(t)} = \left\{ x_p : \|x_p - m_i^{(t)}\| \leq \|x_p - m_j^{(t)}\| \forall 1 \leq j \leq k \right\} \quad (\text{Eq 1})$$

Where each x_p goes into exactly one $S_i^{(t)}$, even if it could go in two of them.

Update step: Calculate the new means to be the centroid of the observations in the cluster.

$$m_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{x_j \in S_i^{(t)}} x_j \quad (\text{Eq 2})$$

This k-means algorithm commonly used initialization methods are Forgy and Random Partition. The Forgy method randomly chooses k observations from the data set and uses these as the initial means. The Random Partition method first randomly assigns a cluster to each observation and then proceeds to the Update step. The Random Partition method is generally preferable for algorithms such as the k-harmonic means and fuzzy k-means^[8]. Demonstration of the standard algorithm is shown below.

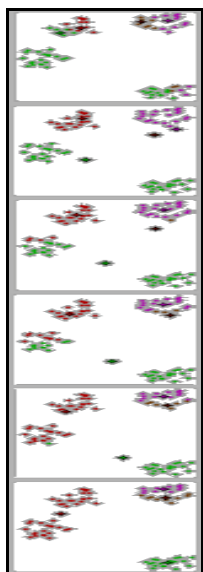


Figure – 2: Concept of K-means clustering technique

In this example, the result of k-means clustering (last image) contradicts the obvious cluster structure of the data set. The small circles are the example of data points, the four ray stars are the centroids (means). The initial configuration is on the left figure. The algorithm converges after five iterations presented on the figures, from the left to the right.

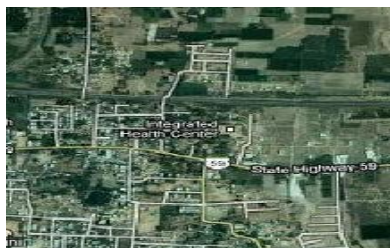


Figure - 3: Original Image



Figure - 4: Clustered image 1(Populated Areas)

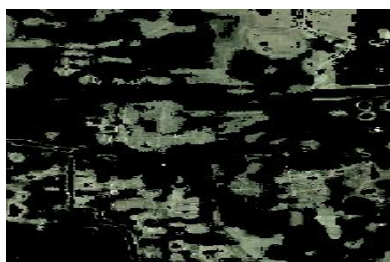


Figure - 5: Clustered image 2 (Sandy Areas with unused lands)



Figure - 6: Clustered image 3 (Vegetation agricultural ongoing areas – Forestation)



Figure - 7: De-correlated Image of Vellore District

5. Conclusion

Thus the clustered images show the status of Deforestation in Vellore district. Continuous evaluation of clustered images can give accurate status of Deforestation of particular area. This kind of image segmentation may be used for mapping the changes in land use land cover taken over temporal period in general. In future this work can be extended for the LANDSAT and other specific satellite images to get the accurate results. This color based segmentation also used for identification of soil type and detection of suitable cultivation for the soil type.

6. References

1. Ahmed Darwish, et al, "Image Segmentation for the Purpose Of Object-Based Classification", 2003 IEEE, pp 2039-2041
2. Brianh.Hahn, Daniel, Valentine, "Essential Mat lab for engineers and scientists".
3. Darren MacDonald, et al; "Evaluation of color image segmentation hierarchies", proceeding of the 3rd Canadian conference on computer and Robot Vision, IEEE, 2006.
4. H C Chen et al, "Visible color difference-based quantitative evaluation of color Segmentation", IEEE proceeding, Vis image signal process vol.153 No.5 Oct 2006 pp 598-609.
5. Jun Tang, "A color image segmentation algorithm based on region growing", 978-1-4244-6349-7, IEEE, vol 6, 2010, pp 634-637
6. Kristina Sandberg, "Introduction to image processing in Matlab", department of applied Mathematics, university of Colorado to boulder.
7. Dae-Jung Kwak et al, "color image segmentation using edge and adaptive threshold value Based on the image characteristics", IEEE proceeding 0-7803-8639-6, 2004, pp 555-558.
8. Pal N.R.et al; "A review on image segmentation techniques", Pattern Recognition 26(9), 1993, Pp1277-1294.
