IJCTA, 9(36), 2016, pp. 473-479

Various Clustering Techniques for Segmentation of Satellite Image

Ankayarkanni. B* and Ezil Sam Leni. A**

Abstract : With the help of high resolution, multi spectral imagery,minor differences that occur based on time,natural calamities,destruction and advancement in infrastructure could be analysed. The feature extraction of the differences in the urban environment helps in the design of urban maps from images that are high in resolution. A multi level object based classification of shape descriptors and hierarchical segmentation. Due to variation in geometric accuracy of spatial resolution,the data processing is being done pixel by pixel. The remote sensing imagery has to be segmented initially and then classified so as to convert it into some meaningful data. This paper proposes methods namely, K-means, Kernel-FCM and Moving -KFCM for the classification of the segments.

Keywords : High resolution, contrast image, K-means, KFCM, Moving KFCM, histogram.

1. INTRODUCTION

So as to improve the accuracy in classification,there have been various research that have been done. Image classification based on remote sensing have been used in many environmental and socio economic applications. There always exist complexities in studying the landscape, sensing data from remote areas,processing the images and classifying them. An appropriate classification technique has to come into effect so as to decrease the challenges involved in classifying remote data. Satellite images have been used to update the GIS database and digital maps by building extraction from them. When carrying out building extraction,there are multiple steps to be performed which could be carried out by an object based image analysis method[1]. This method is divided into two main procedures, segmentation and classification. In urban areas there is a need to list the changes in climate patterns , town planning, environmental protection and topography analysis. To enable speed and to update the GIS database,the complexities and the time required are high. The object oriented approach has to include the form,spectral information and textures. Grouping of neighboring pixels into meaningful areas is done initially in order to extract homogeneous image objects. The various levels could be differentiated based on object categories as they allow multi resolution segmentation. Automatic recognition and segmentation of roads,water bodies ,vegetation , houses and buildings from high resolution images have been carried out. The images are being classified into four classes - water, agriculture, forest and built ups. The eCognition software is being used so as to help in the procedure for classification of the images. Panchromatic images of the interested area are being used for the classification method.

The image classification is mainly divided into various modules which include the determination of the required classification system, selecting the required samples for training, processing of the image, extraction of the features, choosing the required classification method, processing post classification and accessing the accuracy. There are low resolution,medium resolution,high resolution images that are classified based on pixel size,as referred in figure 1. The scale of study,budget,analyzing skill and

* Research Scholar, Department of Computer Science and Engineering, Sathyabama University, Chennai-119, India.

** Professor, Department of Computer Science and Engineering, SRR Engineering College, Chennai-119, India

the requirement greatly influence the way in which remotely sensed data are selected,quality of classification and design of classification. Training samples and a classification system is required prior to the classification process. The software being used,the sources of data and spatial resolution of remote sensing data are factors to be considered ion the selection of classification system.

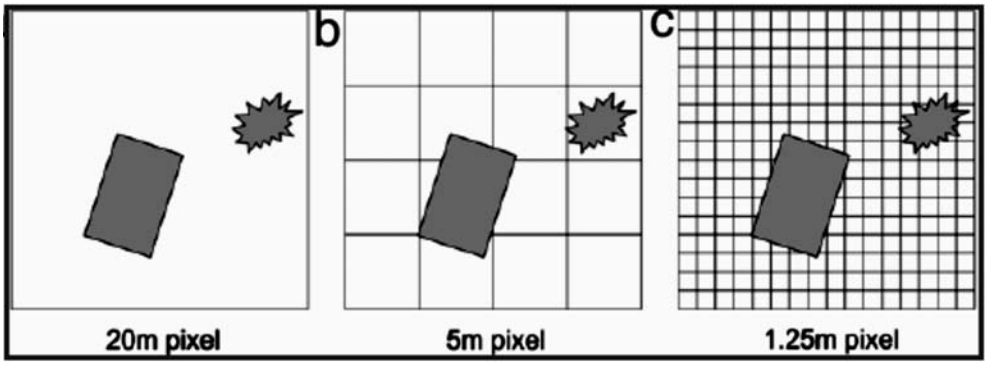

Figure 1: Pixel and object based comparison

2. OBJECTIVE

Research has been carried out in remote sensing image processing as they allow various real life application. In remote sensing research,automatic feature extraction is a promising field of study. High resolution images are needed in order to extract the features of roads,building, vegetation etc. Changes occurring after natural disasters, and other destruction need to be analyzed,hence developing an ideal technique is the work to be done, that will help in updating the GIS and required in urban planning.

Problem definition

Color adjustment technique is used to remove the areas of noise in the pre-processing section. Using the clustering methods as a segmentation technique to segment out the areas of noise in the image into meaningful regions. Clustering the data points as objects.

3. EXPERIMENTAL SETUP

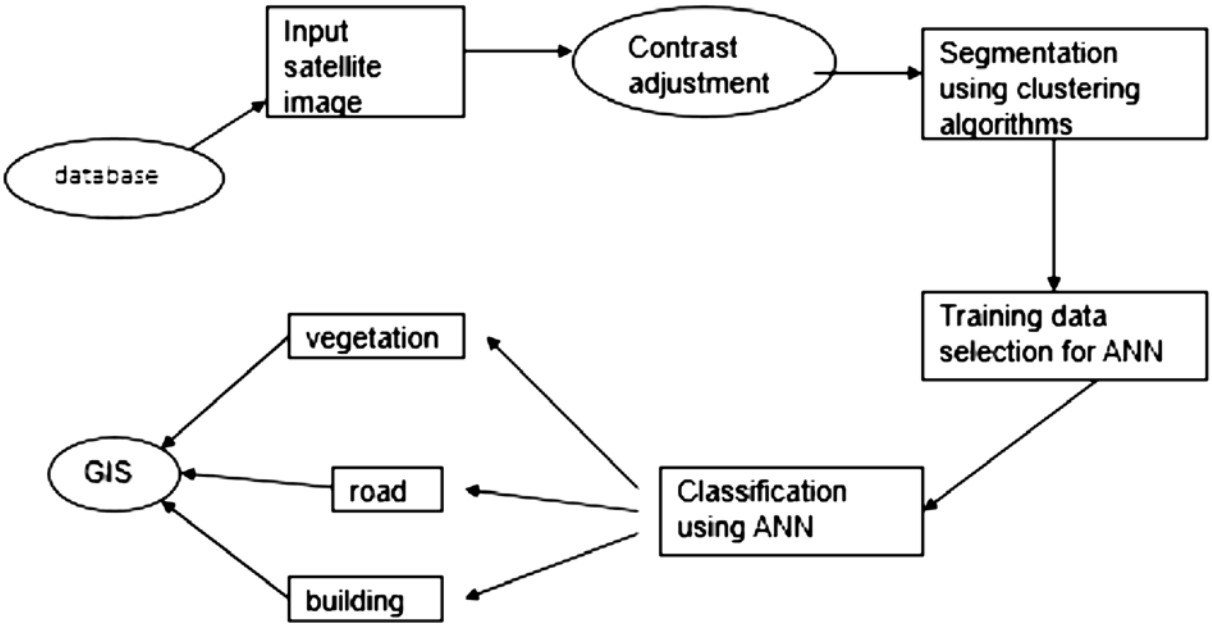

Figure 2: System Architecture

The project makes use of MySQL Server Management 2008 Visual Studio 2010 as the back-end and front-end respectively. In order to save the input image with the identified information such as name of the image and its id, MySQL is being used for creating a database while image pre-processing,clustering method usage and finally performing the classification method post segmentation is done on an windows application formed by Visual Studio.

The system architecture is visualized in figure 2. The classification technique is furnished by the efficient segmentation. The process of segmentation being divided into the major clustering algorithms(K-Means,KFCM, MKFCM) requires pre-processing of the data that involves formation of image objects depending on colour,shape and scale parameters and thus adjustments being done accordingly. Final classification is done by drawing the histogram.

The complete processing is done into 3 steps :

1. Creation of the database where image with its characteristic information like name of the image, its ID etc are stored. As the high resolution images consume more size, they are changed to byte configuration and stored as output image.
2. Proper adjustments of the contrast along with noise reduction and enhancements by using pre processing of data discussed earlier are being done.
3. Image is converted into 6 colour space pixels grouped into clusters of building, vegetation etc which is followed by actuating the histogram to find the final classified image and to perform this,the principle of nearest neighborhood is taken into consideration where pixels of the neighbors are found out and taken for mean calculation.

K-means algorithm

This K-Means algorithm follows the easy classification of the data into clusters of k number having k centers. The positioning of the center needs to be precise as displacing of center yields different output. The determination of center of each cluster continues till the k^{th} cluster. K-means algorithm is unsupervised, iterative technique used to segment image into various clusters based image features. This simple method due to its fast convergence. The simple iterative procedure is as follows:

1. Initialize the k clusters and the cluster center.
2. Calculate the Euclidean distance for each pixel, using the following equation:

$$D = ||p(x, y) - c(k)||$$

   Where, $p(x, y)$ – pixel of an image
   $c(k)$ – cluster center

3. Based on the distance D, assign the pixels to the closest cluster.
4. After assigning all pixels to clusters, calculate the new cluster center, using,

$$C(k) = \frac{1}{k} \sum_{y \in c(k)} \sum_{x \in c(k)} p(x, y)$$

5. Iterate until a tolerable error value.
6. Reshape the pixels of each cluster into image.

Although k-means is easy in implementing, it has few limitations. The cluster center has to be initialized on random basis. Different results are achieved for different selection of cluster centers; hence the quality of the final result varies[2]. Careful selection of initial cluster center should be done. The computational complexity depends on number of clusters formed and number of iterations involved.

KFCM Algorithm

The KFCM algorithm overrides the limitations of that FCM algorithm which cannot handle even the small differences between clusters that are created. This method adds kernel information to fuzzy c-means

algorithm. In this method,non-linear mapping of the input data space into a high dimensional feature space is carried out[3]. The KFCM[4] minimizes the objective function,

$$J_m(U, V) = \sum_{i-1}^{c}\sum_{k-1}^{n} v_{ik}^{m} \left\|\phi(x_i) - \phi(v_i)\right\|^2$$

where n – number of data points.

c – number of cluster centers

v_{ik}^{m} – fuzzy membership of pixels

m – fuzzification exponent

MKFCM Algorithm

Improvement of object segmentation in satellite images is being done with the help of moving KFCM. The distance between the centre and the cluster is being constantly verified and if the centre does not satisfy a particular criterion, the centre is then moved to a region with most active centres. The properties of the designed algorithm are as follows:

- Due to similar fitness in terms of fitness criteria, there is the absence of dead centers.
- Though more centers are allocated in heavily populated data area, in order to get data at an acceptable distance, few data will be assigned to the remaining data.
- Sensitivity to the initial centers is reduced by the algorithm, therefore,the algorithm overcomes local minima situation.

Classification Technique

The pixel wise feature vector based classification consumes valuable time in the classification process[5]. As a solution, statistical measures such as mean and variance and histogram based maximum intensity are used in the classification process. The overall classification is shown in the figure 3.

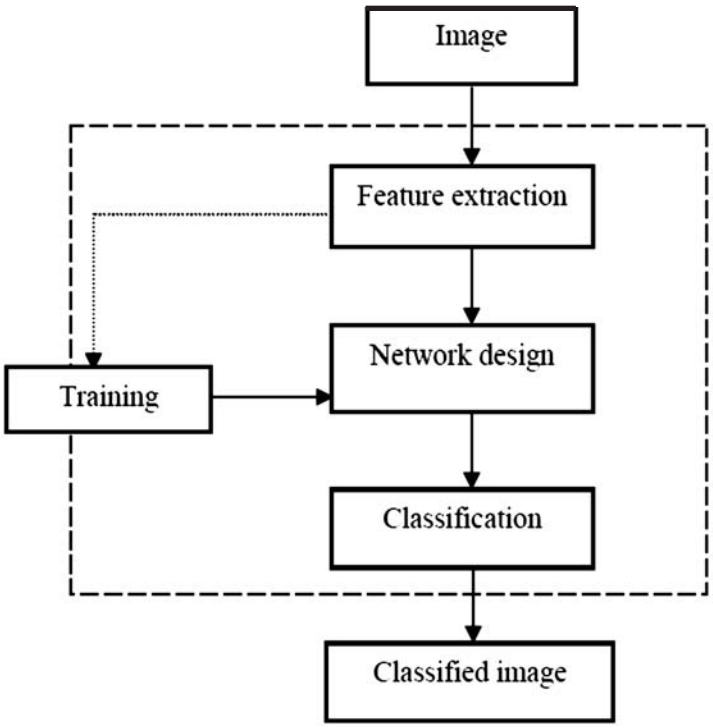

Figure 3: Overall System Classification

The parameters used in classification :

- Pixel wise calculation of the two statistical measures, mean and variance is being done.

Mean $$M = \sum_{i=1}^{N} \frac{p_i}{N}$$

where

p_i is the N pixel

N is the number of pixel

The images are converted into six color space namely

- RGB ,HSV , YIQ color channels, XYZ, LAB color system and LUV components

The features of the six color spaces such as mean (18 features), variance (18 features) and histogram based maximum intensity feature (18 features) so 54 features are given into the artificial neural network.

Calculation of the histogram is being done as follows:

$$f_{set} = \begin{cases} M^R, V^R, H^R, M^G, V^G, H^G, M^B, V^B, H^B, M^H, V^H, H^H, M^S, V^S, H^S, M^V, V^V, H^V, \\ M^Y, V^Y, H^Y, M^I, V^I, H^I, M^Q, V^Q, H^Q, M^X, V^X, H^X, M^Y, V^Y, H^Y, M^Z, V^Z, H^Z, \\ M^L, V^L, H^L, M^A, V^A, H^A, M^B, V^B, H^B, M^L, V^L, H^L, M^U, V^U, H^U, M^V, V^V, H^V \end{cases}$$

Neural Network Classifiers are used to classify input image into roads, vegetation and buildings.

4. RESULTS AND DISCUSSION

Comparison of the output image is being done based on the pixel length in the image and the image with the smallest pixel length is considered the best. The input and output image is shown in figure 4 and 5.

Figure 4: Sample Input Image

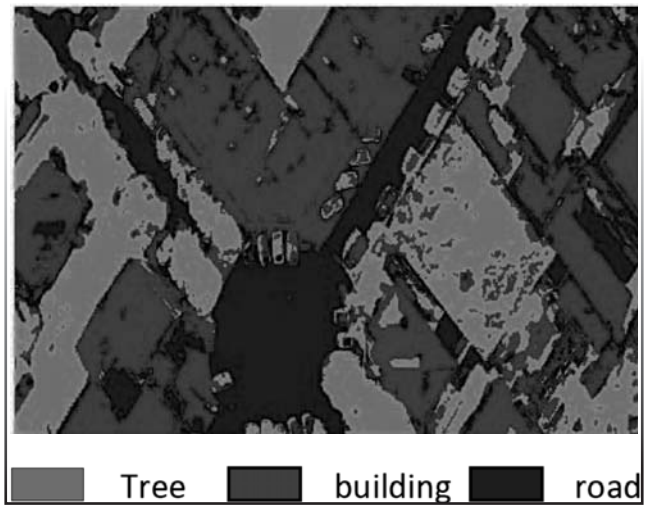

Figure 5: Output Image

The average of the results was calculated and moving KFCM is found to be the best of classification.

5. SUMMARY AND CONCLUSION

The number of similar pixels are being divided into land-use,tree,road and building regions. The performance of the proposed technique is compared with the k-means and KFCM using different kernels like kernel-smoother normal, kernel smoother-triangle, kernel smoother-box and kernel smoother-epanechnikov. The results are depicted in the graph shown in figure 6 and 7. In tree,road,shade and building region classification using all four kernels,the accuracy value is approximately 0.9 for Moving KFCM and has a better results compared to k-means and KFCM.

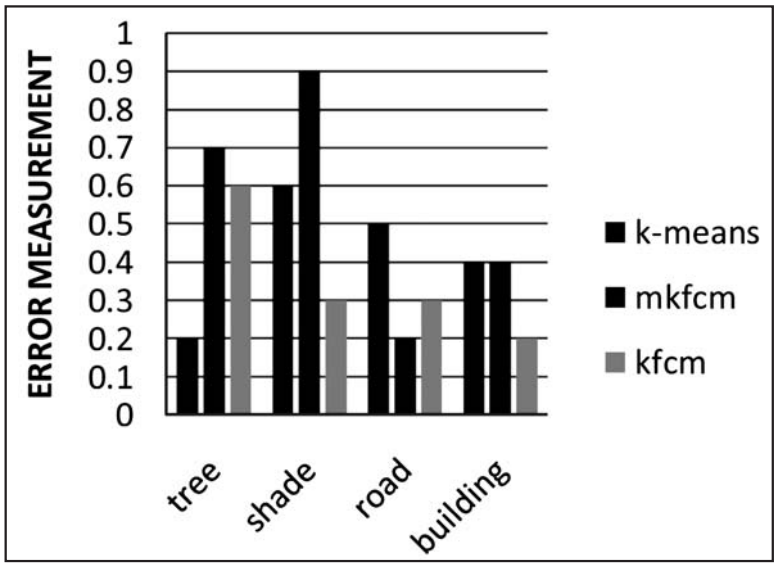

Figure 6: Accuracy graph-tree, shade, road and building region classification using kernel Smoother- box

Synoptic monitoring of the growth patterns and their specific dynamics is achieved through satellite remote sensing imagery. Due to the increase in population and the rise in urban area,the urbanized areas develop more and more into metropolitan regions.

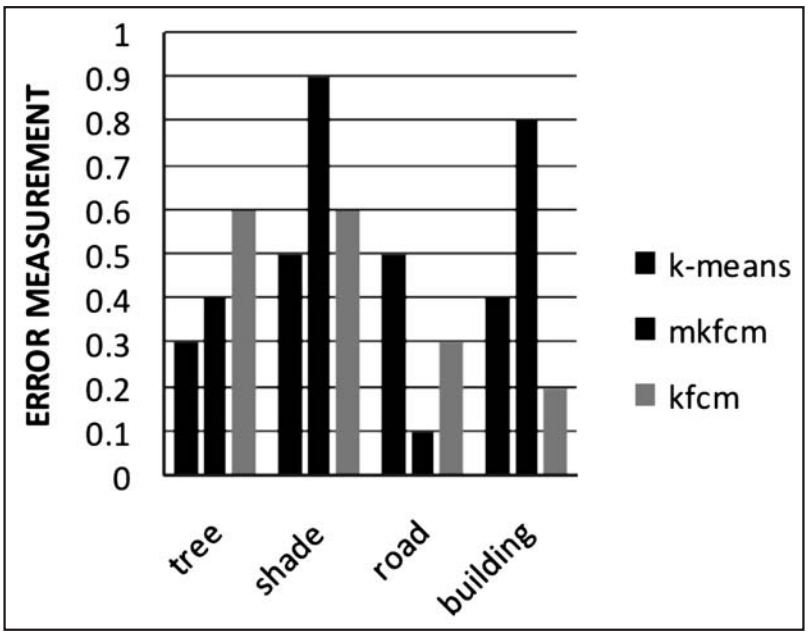

Figure 7: Accuracy graph-tree, shade, road and building region classification using kernel Smoother- epanechnikov

The increase of building and roads are criteria to analyse the development process and so has a direct impact on area consumption. Metropolitan areas with a limited expansion area helpful in converting the inner urban lands into an efficient use.

6. REFERENCES

1. Dezso,B,Giachetta R.,Laszlo I., Fekete I., “Experimental study on graph based image segmentation methods in the classification of satellite images”, EARSeL Proceedings (ISSN 1729-3782), 2012, pp no12–24.
2. Nameirakpam Dhanachandra, Khumanthem Manglem, Yambem Jina Chanu,” Image Segmentation using K-means Clustering Algorithm and Subtractive Clustering Algorithm”, Eleventh International Multi-Conference on Information Processing-2015 (IMCIP-2015),Procedia Computer Science 54 (2015) 764 – 771 .
3. Ameera P., Saritha A., “Image segmentation based on kernel fuzzy C means clustering using edge detection method on noisy images, International Journal of Advanced Research in Computer Engineering & Technology ,Vol 2,Issue 2**,** 2013, pp 399-406.
4. Z.Wu,W.Xie, and J. Yu,”C-means clustering algorithm based on kernel method,” International conference on Computational Intelligence and Multimedia Applications, 2003(ICCIMA 2003),Proceedings Fifth International Conference on. IEEE, 2003, pp. 4954.
5. Blaschke.T, ”Object Based Image Analysis for Remote Sensing”,ISPRS Journal of Photogrammetry and Remote Sensing 65, 2010, pp. 2-16.
6. Andras Benczur, “Object based image analysis in remote sensing application using various segmentation techniques,” Annales Univ. Sci. Budapest., Sect. Comp. 37 (2012) 103–120.
7. Christophe E.,Inglada J.,Osman J., “Interactive Object Segmentation In High Resolution Satellite Images”, Geoscience and Remote Sensing Symposium,2009 IEEE International, IGARSS.
8. Cluster based Key Management Authentication in Wireless Bio Sensor Network “, ,International Journal of pharma and bio sciences, Impact Factor = 5.121(Scopus Indexed).