ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 2, February 2013

Wavelet Based Classification of Multispectral Satellite Image Using Fuzzy Incorporated Hierarchical Clustering With SVM Classifier

S.Sindhu[1], Dr.S.Vasuki[2].

Abstract— **Multispectral satellite images are more efficient and a suitable method of obtaining information about land, because it can captures an image at specific frequency across the spectrum. This spectral image can allow extraction of further information about ground survey than the other traditional image. Classification of multispectral image consists of image processing and classification method. Here, an efficient technique is proposed for classifying the multispectral images using fuzzy incorporated hierarchical clustering with SVM classifier. In the proposed technique, first the multispectral satellite image is subjected to set of pre-processing steps, which are used to transform an image into suitable form that is easier for segmentation and classification. Subsequently, the pre-processed image is segmented using fuzzy incorporated hierarchical clustering. Then, the training data for SVM is chosen from clustered output. Finally the multispectral image is classified into land use and land cover sector based on the training data. The classification is used in the application of land degradation studies, environmental damage, resource management and other environmental application.**

Index Terms— **Classification, Multispectral satellite image, SVM, clustering.**

I. Introduction

Multispectral satellite image produces detail information about environmental factors. It is more economical than other traditional method of ground survey and aerial photography [1]. These details are utilized in a number of applications like making of mapping that was used in military and civil use, estimating of environmental degradation, soil test and crop outcome increment.

Multispectral satellite image classification is done by image processing and classification methods. Multispectral image classification increases efficiency from an efficient method capable of arranging the spectral information in Multispectral dat. usually, image classification refers to creating a group of similar pixels that are belongs to same class. Many classification techniques are introduced in remotely sensing data like KNN, MLP, etc. Comparing to SVM classifier KNN classifier increases storage requirements and complexity [2]. The MLP classifier has a lower accuracy than SVM classifier. Both KNN and MLP produce 67% of accuracy in classification of remotely sensing data [3]. But SVM produces 76% of accuracy in classification field. The time complexity is also reduced in SVM. It is suitable for lower data comparison compared with FNN algorithm.

The main focus of research is to classify the images into land use and land cover region. Land covers characterize the features of land surfaces like vegetation area, soil, mud and crop. Land use is a statement of how people utilize the land like building, roof, etc. This land cover mapping is an important parameter for environmental and land use planning at local and national level. Here, a new technique is introduced for increase the classification efficiency with SVM.

In the proposed techniques consists of four phases of pre-processing by wavelet transform, segmentation by fuzzy incorporated hierarchical clustering, training data selection for SVM, classification using trained SVM. The multispectral satellite image is applied to Gaussian filter. The image quality is enhanced by reducing noise component in an image. The better image can be obtained from Gaussian filtering. Then, it is applied to wavelet transform. It can provide detail information about multispectral data which are used to increase the classification accuracy. Then, fuzzy incorporated hierarchical clustering is used for segmentation of the image into clusters. The training data are subjected to SVM and final classification of the Multispectral data is obtained.

II. PROPOSED WORK

In the proposed technique, the multispectral image is classified by using SVM classifier with clustering. Initially pre-processing is done on the input multispectral image where the image is subjected to set of pre-processing steps such as Gaussian filtering and wavelet transform so that image is transformed suitably for segmentation. The pre-processed image is segmented using fuzzy incorporated hierarchical clustering. Then the training data is selected from clustered output which is the input of SVM classifier. Finally classification of Multispectral image is done which is based on the training data given to the classifier. The proposed technique is illustrated by figure (1).

Manuscript received Feb, 2013.

***S.Sindhu**, PG student, Department of ECE, Velammal college of engineering and technology,Madurai Tamilnadu,India*

***Dr.S.Vasuki**, Professor & Head, Department of ECE, Velammal college of engineering and technology,Madurai , Tamilnadu,India*

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 2, February 2013

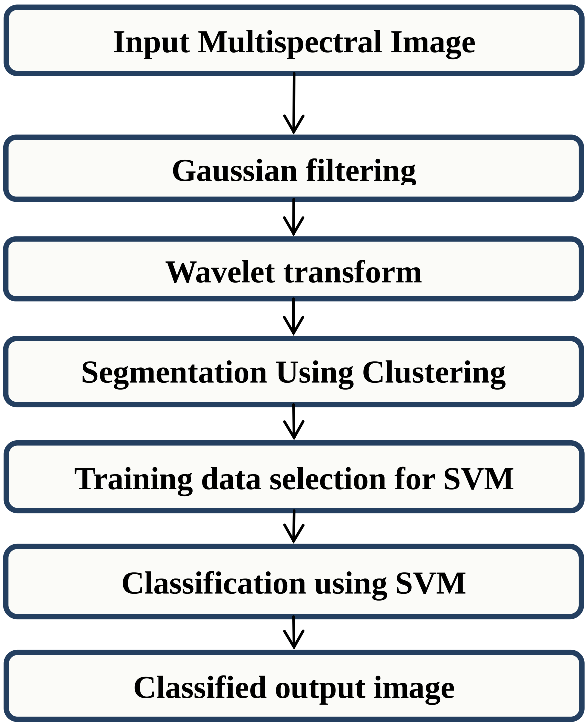

Fig.1. Block Diagram of the Proposed Work

III. GAUSSIAN FILTERING

A Gaussian function is the response of Gaussian filtering. Gaussian filtering is used to avoid the overshoot of step function while reducing the rise and fall time [4]. Because of this characteristics Gaussian filtering has a minimum group delay. Gaussian filter performing convolution between input and Gaussian factors. The result is called Weierstrass transform.

The 1D Gaussian filter is given by,

$$g((x)) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}} \qquad (1)$$

The impulse response of the 1D Gaussian filter is given by,

$$g((x)) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-\sigma^2 u^2}{2}} \qquad (2)$$

Here in the pre-processing the unwanted noise components are reduced from the input image while passing through Gaussian filter. It is mainly used to enhance the image quality than the original image.

IV. WAVELET TRANSFORM

Wavelet transform is based on wavelet theory. It was initially developed for signal transformation. In discrete wavelet transform (DWT) the signal is applied to series of low pass and high pass filter to analyse the low frequency and high frequency component [5]. The amount of detail information is measured from filtering operation and scale is changed by up sampling and down sampling. The DWT analyse the signal at different frequency and different resolution by decomposing signal into approximation and detail information. The DWT is developed by Mallet (1989) for image decomposition. The DWT decompose an image into one approximation and three detail information sub bands. The image is subjected into bank of low pass and high pass filter. It filters an image in all direction. The filtered images are down sampled at every pixel, it producing four sub band images of original image. Figure illustrates the process of first level decomposition. For a one level decomposition wavelet transform, the approximation sub image maintain the maximum portion of the original image, while the detail sub image represents the differences between approximation sub image and original image in different direction. In second level decomposition the filtering process is applied to the approximation sub image, resulting is four additional sub bands in approximation. These details are given in FIGURE (2).

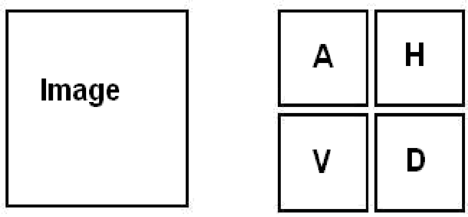

Fig2.(a). Input image and sub images position in output image

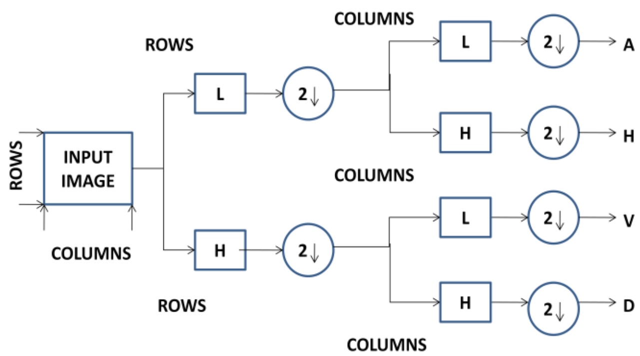

Fig.2. (b). Convolution Of Image.

The wavelet decomposition is otherwise called as filter convolution. In the first level decomposition, the image can be obtained by convolving the image with appropriate filter. The wavelet transform provides set of pixel information in each level of decomposition. From the approximation and detail information, the fine information is obtained for classification. This transform leads to reduce the misclassification of process.

Multispectral satellite images is used as input image which is to be subjected to the Gaussian filtering for noise reduction and applied to the wavelet transform. The three (RED, GREEN and BLUE) planes are extracted from the Gaussian filtered output. The DWT is applied into the separate planes. So four sub bands are obtained for separate planes. An approximation sub band is enough for the future process. Hence, the approximation sub bands of three planes are concatenated to get the approximated output image. This approximated output image is applied to the segmentation using Fuzzy incorporated hierarchical clustering. The training data for SVM are taken from the clustering output. The SVM classifier is classified the image as a land use and land cover based on the training data

V. SEGMENTATION USING PROPOSED CLUSTERING

After performing the pre-processing steps, a suitable image is obtained for segmentation. The processed image consists of thousand of pixels and to classify the image based on of pixels is time consuming. Hence, he input image is clustered based on the centroid pixel value. This is due to the fact, that all pixels are differed only a small amount from centroid pixel value. Hence, the centroid pixel value represents all pixels in the cluster. This centroid value clustering will reduce the input to SVM classifier and reduce the complexity of design. In the proposed clustering, the hierarchical clustering is combined with the Fuzzy C means. This will increase the accuracy of Fuzzy C means and reduce the time complexity of Hierarchical clustering. Hierarchical clustering has two types. 1. Agglomerative 2. Divisive. However, is has disadvantage of high complexity and minor variation in dataset also greatly changes the hierarchical dendrogram structure [6]. The traditional FCM has a disadvantages of it does not yield the accurate result. That is, every time Fuzzy clustered the same data but producing different results. These two drawbacks are overcome by the proposed Fuzzy incorporated hierarchical clustering [7]. The clustering process is explained below:1. Let the image have M number of pixels. Initially all the pixels are act as a different cluster and hence it forms M different cluster. Let the M cluster of an image is represented as C_i, where 0< i<M. Calculate pixel difference matrix λ.

$$\lambda = \begin{bmatrix} \partial_{11} & \partial_{12} & \partial_{13} & \partial_{14} & .. & \partial_{1M} \\ \partial_{21} & \partial_{22} & \partial_{23} & \partial_{24} & .. & \partial_{2M} \\ \partial_{31} & \partial_{32} & \partial_{33} & \partial_{34} & .. & \partial_{3M} \\ \partial_{41} & \partial_{42} & \partial_{43} & \partial_{44} & .. & \partial_{4M} \\ .. & .. & .. & .. & .. & .. \\ \partial_{M1} & \partial_{M2} & \partial_{M3} & \partial_{M4} & .. & \partial_{MM} \end{bmatrix}$$

Where ∂_{ij} = pixel difference value between i th and j th cluster.

From the pixel difference matrix find out two clusters which have a minimum differences and merge together to form a new cluster C_{ij}.
3. The centroid value O_{ij} is calculated for new cluster.

$$O_{IJ} = \frac{C_I + C_J}{2} \tag{3}$$

4. Subsequently, find out the original centroid pixel value by using incorporated Fuzzy C means algorithm. Find membership value and then calculate the original centroid value using membership function.

$$M_{IJ} = \frac{1}{\sum_{k=1}^{c}\left(\frac{\|P_i - O_{ij}\|}{\|P_i - o_k\|}\right)^{\frac{2}{m-1}}} \tag{4}$$

Where,
O_{ij} - approximated centroid pixel value
O_k - centroid pixels of other clusters excluding the newly formed cluster.
m – Positive integer
Modified centroid pixel value for newly formed cluster is given by,

$$C_{ij} \text{ is given by } O_j = \frac{\sum_{i=1}^{N} \mu_{ij}^{m} x_i}{\sum_{i=1}^{N} \mu_{ij}^{m}} \tag{5}$$

5. The above result is the formation of two individual clusters which have greatest similarities. Hence it decreases the total number of clusters by one, after every iteration.
6. The pixel difference matrix dimension is also reduced by replacing the value of C_i and C_j value by Cij. It will lead to decrease the matrix dimension of M X M to (M - K) X (M – K) after K loops.
7. Go back to step 2 until the desired number of cluster is obtained.

VI. SVM CLASSIFIER

Now a day, so many techniques are developed for satellite image classification. The earlier day maximum likelihood classifier is used for classification [8]. In recent times, artificial intelligence techniques are developed for remotely-sensed image classification application. Support vector machines (SVM) is a statistical learning based classification system, which classifies the data based on the decision boundary [9].SVM classifier has a margin between two different classes. That plane is known as optimal hyperplane. SVM classifier is mainly implemented for risk minimization. It has a high margin between two classes, which minimizes the misclassification of data. The adaptability of SVM classifiers are (1) The SVM classifier can become a non linear by employment of kernel function. (2) The multiclass classifier is made by clubbing large number of binary classifier [10]. The result of SVM classifier is grouping of similar pixels for each class of data.

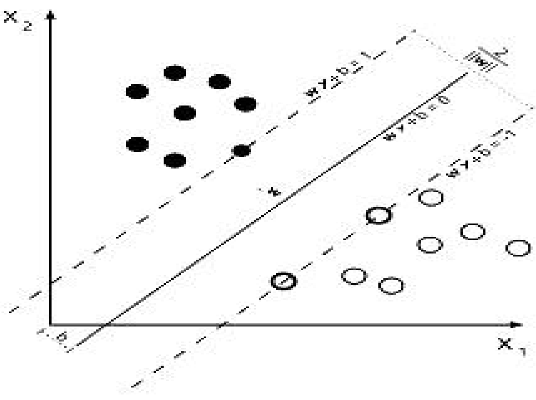

Fig.3.Margins for an SVM Trained with samples from two classes

In the two class classification, assume the set of training data set $f = \{x_k, d_k\}$ where, k = 1, 2 …Q. $x_k \in R^n$, $d_k \in \{-1, 1\}$. Let the classifier is given as,

$F(X, W, w_0) = sign\ (W.\ X + w_0)$

Where,
F = bipolar signum function which will permit a mapping from input point X_k to appropriate point d_k.
W = set of weights
w_0 = bias (which separate positive data points from negative one)

Figure (3) shows the two class classification with high margin hyperplane. Class c1 consists of positive data, which is indicated by +1 and class c2 consists of negative data, which is indicated by -1. The hyperplane is defined by $W.\ X + w_o = 0$. The SVM classify the data based on the quantity of W. If $W.\ X + w_0 > o$, then the datas are belongs to class 1 and W. X + w0 < 0, then the datas are belongs to class2. The data close to the hyperplane is called as support vector. The support vectors lies on the two parallel planes $W.\ X + w_0 = \pm 1$, which maximize the SVM classifier margin.

A. Training data selection:

In this technique, the training data selection from the clustered output image is discussed [11]. Our proposed technique is used to classify the multispectral image into land use and land cover. This classification accuracy increased from the colour features extraction in satellite images. Each and every elements in the earth are differentiated by this own colour. In the multispectral image certain colour stands for land use and
certain colour stands for land cover [12]. These colours are identified from the satellite image and given to the input of SVM classifier.

B. Classification using SVM:

In the proposed method, the multispectral image is segmented by Fuzzy incorporated Hierarchical clustering. All clusters almost have a similar number of pixels value and the clusters are differ from the centroid value by small amount only [13]. Hence, the centroid value represents all the pixels value in the cluster. This centroid valve is given to the SVM classifier hence it reduce the classifier complexity and also time incurred. Suppose consider the i^{th} cluster, which having n number of pixels, where each pixel having a value of P_k. Then the centroid value is calculated by

$$o_i = \frac{\sum_{k=1}^{n} P_k}{n} \qquad (6)$$

If our image consists of N number of cluster, then find the centroid set O.O = (O_1, O_2… ON). It will be given to the input of the SVM classifier. The centroid set is classified based on the training data given to the SVM classifier.

VII. RESULTS AND DISCUSSION

The classification and segmentation of multispectral images is implemented by MATLAB. In this approach, the proposed method of classification is discussed.

A. Experimental Results:

In this section, the output of the proposed technique is discussed. Multispectral satellite images is used as a input images which is to be subjected to the Gaussian filtering for noise reduction and applied to the wavelet transform. The three (RED, GREEN and BLUE) planes are extracted from the Gaussian filtered output [14]. The DWT is applied into the separate planes. So four sub bands for separate planes are obtained. An approximation sub band is enough for the future process. Hence, the approximation sub bands of three planes are concatenated to get the approximated output image. This approximated output image is applied to the segmentation using Fuzzy incorporated hierarchical clustering. The training data for SVM are taken from the clustering output. The SVM classifier is classified the image as a land use and land cover based on the training data. Figure(4) shows the input image. Along with Figure(5) shows the output of Clustering. The final classified output of SVM is given in figure (6).

Fig.4. Input image.

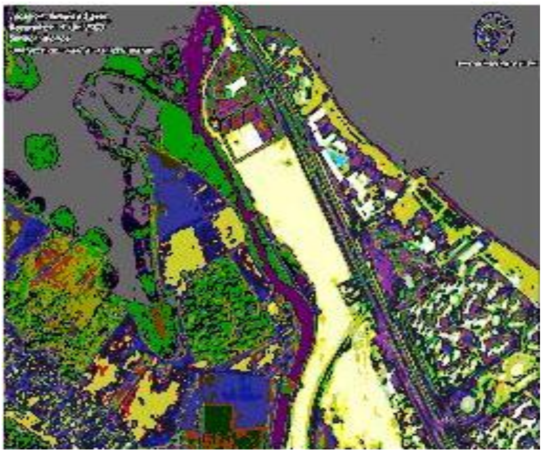

Fig.5. Clustering Output.

ISSN: 2278 – 1323
International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)
Volume 2, Issue 2, February 2013

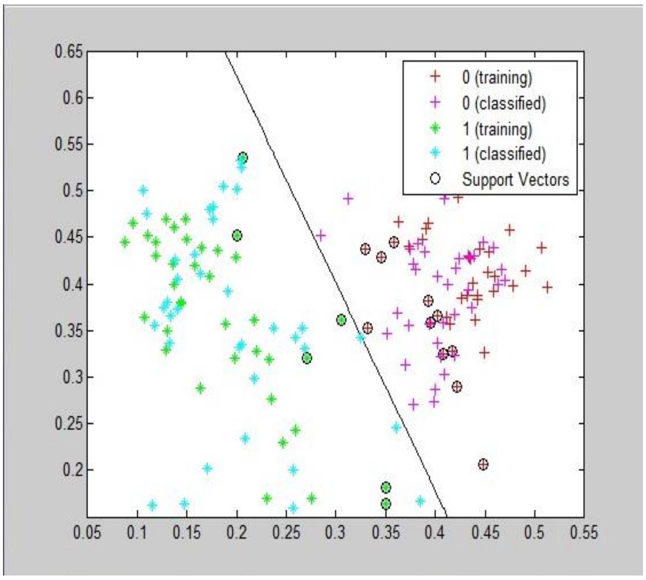

Fig.6. SVM Classifier output

C. Performance Evaluation:

In this section, the performance evaluation of proposed techniques is discussed. Here, the performance of various satellite multispectral images are evaluated. The SVM output of different satellite images are shown in fig.7,8,9,10.

Satellite Images	Accuracy (%)	Error Rate	Sensitivity	Specificity
IKONOS	89.19	0.1081	0.9181	0.8649
QUICK BIRD	91.89	0.0811	0.9459	0.8919
GEO EYE-1	87.84	0.1216	0.8919	0.8649
GEO EYE-2	94.59	0.0547	0.9459	0.9459
RAPID EYE	83.78	0.1622	0.8919	0.7838

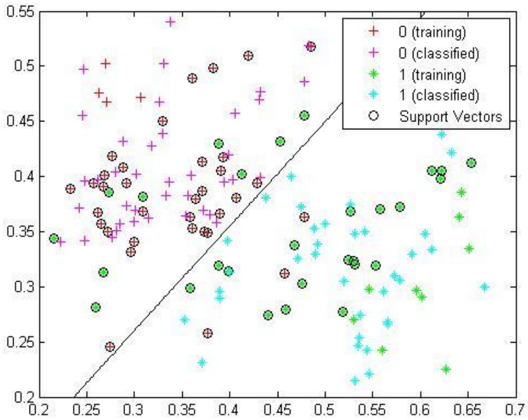

Fig.7. SVM classifier output of **QUICK BIRD**.

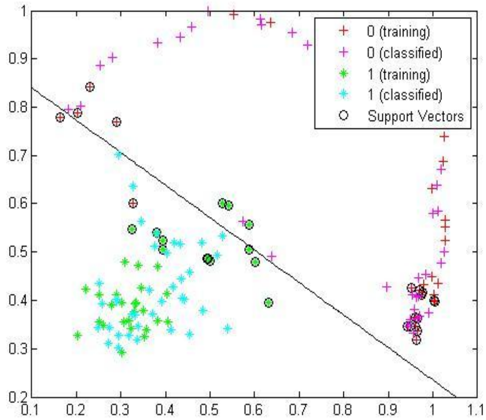

Fig.8. SVM classifier output of **GEO EYE-2**.

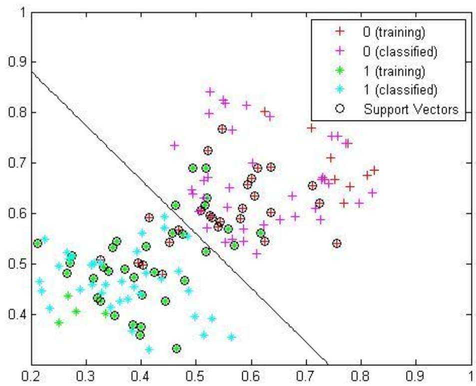

Fig.9. SVM classifier output of **GEO EYE-1**

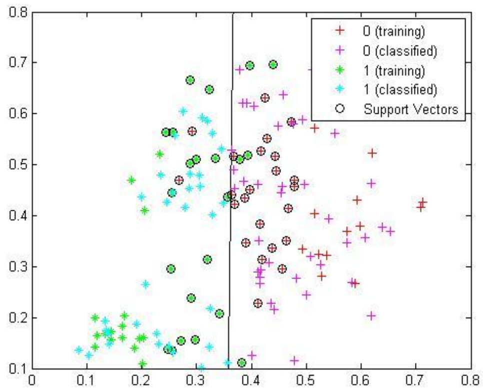

Fig.10. SVM classifier output of **RAPID EYE**

VIII. CONCLUSION

In this paper, an efficient image classification technique is proposed with the help of SVM classifier. Here in our proposed technique is made of four phases namely pre-processing, image segmentation, training data selection and final classification using SVM classifier. First the multispectral satellite image is subjected to set of pre-processing steps, which are used to transform an image into suitable form that is easier for segmentation and classification. Subsequently, the pre-processed image is segmented using fuzzy incorporated hierarchical clustering. This result in the image is segmented into number of clusters. Then, the training data for SVM is chosen from clustered output. The chosen data is given to the input of trained SVM. Finally the multispectral image is classified into land use and land cover sector based on the training data. The experimental results demonstrated the effectiveness of the proposed classification techniques. The analyse ensures that the classification has good accuracy in all type of multispectral satellite images. In the future work, extent our approach into classifying the multispectral satellite images into multiple region rather than land use and land cover. In that case, distinguish between land use and land cover feature are known in a better way and can be more useful.

REFERENCES

[1] . K Perumal and R Bhaskaran , "SVM-Based Effective Land Use Classification System For Multispectral Remote Sensing Images", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 6, No. 2, pp.95-107, 2009.

[2] M. Fernandez-Delgado, P. Carrion, E.Cernadas, J.F. Galvez, Pilar Sa-Otero (2004)-"Improved Classification of Pollen texture Images using SVM and MLP"

[3] Shaomei Yang and Qian Zhu (2008)-"Research on Comparison and Application of SVM and FNN Algorithm".

[4] Dong-Hyuk Shin, Rae-Hong Park, Senior Member, IEEE, Seungjoon Yang, Member, IEEE, and Jae-Han Jung, Member, IEEE (2005)-"Block-Based Noise Estimation Using Adaptive Gaussian".

[5] Chen, D. Chen, and D Blostein (2007)-"Wavelet-Based Classification of Remotely Sensed Images".

[6] Wei Yu, Jason Fritts and Fangting Sun (2000),"A Hierarchical Image Segmentation Algorithm".

[7] Geva,A.B(1999)-"Hierarchical unsupervised fuzzy clustering ".

[8] D. Lu, Q. Weng, "A survey of image classification methods and techniques forimproving classification performance", International Journal of Remote Sensing, Vol. 28, No. 5, pp. 823-870, January 2007.

[9] M. Govender, K. Chetty, V. Naiken and H. Bulcock, "A comparison of satellite hyperspectral and multispectral remote sensing imagery for improved classification and mapping of vegetation", Water SA, Vol. 34, No. 2, April 2008.

[10] B.Sowmya and B Sheelarani , "Land cover classification using reformed fuzzy C-means", Sadhana, Vol. 36, No. 2, pp. 153–165, 2011.

[11] V.K.Panchal, Parminder Singh, Navdeep Kaur and Harish Kundra, "Biogeography based Satellite Image Classification", International Journal of Computer Science and Information Security IJCSIS, Vol. 6, No. 2, pp. 269-274, November 2009.

[12] Huang B, Xie C, Tay R, Wu B, 2009, "Land-use-change modeling using unbalanced support-vector machines" , Environment and Planning B: Planning and Design , Vol.36, No.3, pp.398-416,2009.

[13] Ming-Hseng Tseng, Sheng-Jhe Chen, Gwo- Haur Hwang, Ming-Yu Shen, "A genetic algorithm rule-based approach for land-cover classification", Journal of Photogrammetry and Remote Sensing ,Vol.63, No.2, (3), pp. 202-212, 2008.

[14] Pall Oskar Gislason, Jon Atli Benediktsson, Johannes R. Sveinsson, "Random Forests for land cover classification", Pattern Recognition Letters,Vol.27, No.4, (3), pp. 294-300, 2006.

[15] Cristianini, Nello and Shawe-Taylor, John, "An Introduction to Support Vector Machines and other kernel based learning methods", Cambridge University Press, Cambridge, 2000.

[16] S. C. Johnson, "Hierarchical Clustering Schemes", *Psychometrika*, Vol.2, pp.241-254, 1967.

[17] J. C. Dunn (1973): "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters", *Journal of Cybernetics*, Vol. 3, pp.32-57, 1973.

[18] R.A. Haddad and A.N. Akansu, "A Class of Fast Gaussian Binomial Filters for Speech and Image Processing," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 39, pp 723-727, March 1991.

[19] Hunter and Richard Sewall ,"Accuracy, Precision, and Stability of New Photo-electric Color-Difference Meter", Proceedings of the Thirty-Third Annual Meeting of the Optical Society of America, Vol. 38(12), 1948.

[20] James A. Shine and Daniel B. Carr, "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, 11-15 August 2002.

***Author Profile: Sindhu** is a PG student in the department of Electronic & Communication Engineering at Velammal college of engineering and technology,Madurai,Tamilnadu,India. Her current area of research is Satellite Image Processing.*