EXCELLENT PUBLISHERS

International Journal of Current Research and Academic Review

ISSN: 2347-3215 Volume 2 Number 6 (June-2014) pp. 56-66

www.ijcrar.com

Histogram Binning and Morphology based Image Classification

K.Kavitha[1], S.Arivazhagan[2], and B.Suriya[3]*

[1]Assistant Professor, [2]Principal, [3]Student, Department of E.C.E, Mepco Schlenk Engineering College, Sivakasi, India

***Corresponding author**

KEYWORDS

Hyperspectral bands, Histogram binning, morphological operations, ELM classifier, statistical classifier

A B S T R A C T

A hyperspectral image is characterized by a large dimensionality data, recorded at very fine spectral resolution in hundreds of narrow frequency bands. These bands provide a wealth of spatial and spectral information of the scene, imaged by the sensors. Histogram binning and morphological operator based Extreme Learning Machine classifier is proposed. Histogram binning and morphological operator is used as a preprocessing step which reduces the computational complexity, which involves in the hyperspectral data processing. Each object have distinct reflectance value based on which the objects are classified by choosing appropriate features and the classifier. Histogram binning represents the image by reduced gray level. Morphological operations envisage the finer inner details of the image. Now from this new version of the processed image, statistical features such as mean, median, standard deviation, mode and variance have been extracted. The extracted features are used for classification. The Extreme Learning Machine is proposed to classify the image. The classification is done with different types of kernels. The performance of each type of kernel is evaluated. The experiment is conducted on the AVIRIS hyperspectral dataset taken over the North-western Indiana's Indian Pine Site.

Introduction

A remarkable increase in spectral resolution has led to imaging sensors that can gather data in hundreds of contiguous narrow spectral bands to generate Hyperspectral images. This ability of imaging sensors, to acquire the reflectance spectrum of a pixel in significant detail, leads to substantial differences in the reflectance values of the pixels belonging to disparate materials on the Earth's surface. Hence, one of the most prominent applications of Hyperspectral imagery is land cover classification.

A generalized procedure for hyperspectral data analysis is suggested and it involves identification and labeling of training samples followed by feature extraction and classification[1]. Classification of heterogeneous classes present in the Hyperspectral image is one of the recent research issues in the field of remote sensing.

Hyperspectral imaging is related to multispectral imaging. The distinction between hyper- and multi-spectral is sometimes based on an arbitrary "number of bands" or on the type of measurement, depending on what is appropriate to the purpose. A Hyperspectral image has hundreds of bands, where as a multi spectral has 4 to 7 bands only [1].Thus every pixel in the hyper spectral image contains values that proportionate to the detailed spectrum of reflected light [2],[3].This rich spectral information in every spatial location lead to the potential classification. So, the hyper spectral images provide more information for within class discrimination, i.e. they can discriminate between different types of rocks and vegetation while multi spectral images can discriminate between rock and vegetation alone. Such hyper spectral images can be used effectively to classify the heterogeneous classes which are present in the image. To classify the multi-classes of information in an image it is necessary to extract the proper set of features and to choose a proper classifier.

Classification of Hyperspectral images requires several factors to be considered. They are 1) the large number of land-cover class to be handled. 2) High number of spectral bands but low number of the availability of training samples. To overcome these problems proper classifier should be selected. The classifier has to support the voluminous data, multi class data and the nonlinear dataset. Significant features among the extracted features must be selected and to be fed to the classifiers to obtain the high accuracy rate. This process is commonly known as feature selection in the literature. An extensive literature is available on the pixel-level processing techniques. i.e., techniques that assign each pixel to one of the classes based on its spectral values [4], Maximum likelihood or Bayesian estimation methods [5], decision trees [6 - 7], genetic algorithm based methods [8-10], genetic algorithms [11] and kernel based techniques [12-15] have been investigated.

Among the large number of extracted features, the recursive features can be eliminated by the method of feature selection [16] and the reduced set of significant features can be used for classification.

As classifiers are concerned, Spectral Angle is a time consuming work. For very complex boundaries of data this shows poor accuracy.

Classifiers are classical one, used for hyper spectral image analysis in early works [17], where the input image spectra are compared with the reference spectra for classification. The classes were decided by calculating angular separation between the input and reference spectra known as spectral signature. So, a classifier which is able to perform classification even with very complex boundaries and to overcome all the above cited problems was needed. It is obvious that the extreme learning machine classifier outrage the performance of all other types of classifiers. The hyper spectral domain is a non-linear one. Non-linear domain can be converted into the linear domain by using kernel trick. Many types of kernels like linear, polynomial, radial Basis Function (RBF), Sigmoid etc., are available. Selection of proper kernel gives proper results. The use of other classifiers in [18],[19],[20],[21] are with certain drawbacks of omitting many pixels.

From the literature, it is evident that the context or the position information of the pixel is also more important along with the spectral information. Extracting features and the use of proper classifier are the general aspects concluded from the literature survey. So in this proposed methodology first the input bands undergo histogram binning which is done as the pre-processing feature reduction step. This reduces the overall dimensionality which is just sufficient for the efficient classification. With this reduced dimensionality the image is pre- processed by morphological operations like dilation. With this image the statistical features are extracted. The Combined Features can be derived. Such Combined Features are expected to yield the better spatially classified image. The robust performance of ELM classifier makes the classification much effective and efficient.

Proposed Work

A remotely sensed hyper spectral image is acquired by sensors which provide ample spectral information to identify and distinguish between spectrally similar materials. This image acquired is in multi bands which are then processed to overcome the complexity but still holding the accuracy. Our paper progress in the following step by step approach

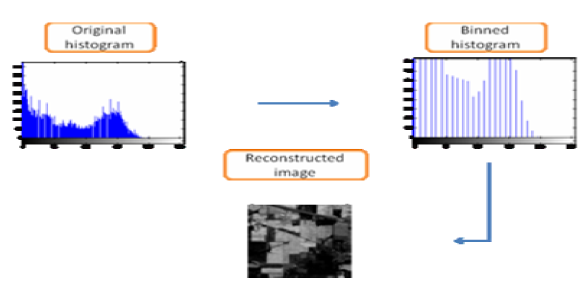
- ✓ Remote sensing
- ✓ Image processing
- ✓ Pre-processing
- ✓ Histogram binning
- ✓ Morphological operations
- ✓ Feature extraction

Classification by classifier

The remotely sensed image which is readily available is processed as the initial stage to overcome the computational complexity in handling the multi bands. The processed bands are then taken as the initial input spectral image. here the input is the Hyperspectral image acquired by the AVIRIS sensor over Indiana pines which consists of 145*145 pixels and 224 spectral bands.

Histogram Binning

The histogram which is the graphical representation of the tonal distribution of the image is first plotted for each separate bands considered. It is a graph in which the horizontal axis represents the tonal variations and the vertical axis represents the number of pixels in that tone. Once the histogram is plotted then the plotted histogram is linearly binned. Binning is the process of combining two or more pixels into a single super pixel. In linear binning the gray level is divided into equally divided into equal sized bins. Thus the overall impact is the image is being represented by reduced gray levels which are just enough to represent the image. From this binned histogram the original image is then again reconstructed. Thus the overall data dimensionality is being reduced.



Morphological Operations

While point and neighborhood operations are generally designed to alter the look or appearance of an image for visual considerations, morphological operations are used to understand the structure or form of an image. This usually means identifying objects or boundaries within an image. Morphological operations play a key role in applications such as machine vision and automatic object detection.

The foundation of morphological process is in the mathematically rigorous field of set theory. The operation is mainly based on shapes that is it process the image based on shapes. A proper structuring element of appropriate shape and size is chosen which slides on the image to be processed. The value is calculated for each and every pixel based on its neighbors which is being decided by the structuring element. Among the several morphological operators available the one suitable for this work is dilation. Here in this process the value of the output pixel is the maximum of the value of the input neighboring pixels. The pixel beyond the image border is assigned the minimum value. The overall impact of this process is that the area of foreground pixels grows in size whereas the holes within the regions become smaller. This is the resulted image after dilation

Feature Extraction

Feature reduction is a special form of dimensionality reduction. Since our input data is too large to be processed in which some of the data may lack much details. In such case the data is transformed to a reduced set of features. Thus the input is transformed to a reduced set of features which represents the relevant information required to be further processed. Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately.

Mean

The *mean* of a data set is simply the arithmetic average of the values in the set, obtained by summing the values and dividing by the number of values. In other words mean is a measure of the *center* of the distribution. In words,

Mean = Sum of X values / N(Number of values)

Mathematically mean is calculated by the formula,

$$\bar{X} = \frac{\sum_{i=1}^{i=n} X_i}{n}$$

Median

The median is the middle value of a set of data containing an odd number of values, or the average of the two middle values of a set of data with an even number of values. The median is especially helpful when separating data into two equal sized bins

Standard Deviation

Standard deviation is a widely used measure of variability or diversity used in statistics and probability theory. It shows how much variation or "dispersion" exists from the average (mean, or expected value).

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{i=n}(X_i - \bar{X})^2} \tag{2}$$

Variance

Variance is the Average of the squared distance from mean. The variance (σ^2) is a measure of how far each value in the data set is from the mean.

Mathematically it is calculated as

$$\sigma^2 = \frac{\sum(X-\mu)^2}{N}$$

Mode

In statistics, the mode is the value that occurs most frequently in a data set or a probability distribution. In some fields, notably education, sample data are often called scores, and the sample mode is known as the modal score. Mode is the most frequently occurring value in frequency distribution

Several statistical features are chosen to be extracted. Some of the features extracted are mean, median, standard deviation, mode and variance. These five feature vectors are then combined into a single super feature vector which is the final input to the classifier.

Classifier

Extreme Learning Machine Classifier

There are numerous aspects that must be considered when conducting a supervised classification.. Training areas must be selected for each of the classes and statistics calculated for them. The appropriate classification algorithm has to be selected, and once each pixel in the image (including the ones used as training areas) are evaluated and assigned to a land cover class, the accuracy of the classification has to be assessed.

There are numerous factors that can affect the training signatures of the land cover classes. Environmental factors such as differences in soil type, varying soil moisture, and health of vegetation, can affect the signature and affect the accuracy of the final thematic map.

The classifier used here is extreme learning machine (ELM) classifier. ELM classifier is a supervised classifier. for a supervised classifier a reference image is needed according to which the classifier is trained and tested. Image classification consists of assigning a label to each pixel of an observed image. Some of the pixels from reference image is taken to be trained and some pixels from original image is taken to be tested. The supervised classifier also includes groundtruthing process where the image acquired by the sensor is compared with the pixels what is actually there in reality. Thus the efficiency of the final classification by the proposed classifier is being determined.

Extreme Learning Machine (ELM) was able to overcome the difficulties of a neural network through a fast learning speed and high performance. There are different types of kernels available and some of which are as follows:

Linear kernel

The Linear kernel is the simplest kernel function. It is given by the inner product <x,y> plus an optional constant **c**. Kernel algorithms using a linear kernel are often equivalent to their non-kernel counterparts, i.e. KPCA with linear kernel is the same as standard PCA.

$$k(x,y) = x^T y + c \tag{4}$$

Sigmoid kernel

The Hyperbolic Tangent Kernel is also known as the Sigmoid Kernel and as the Multilayer Perceptron (MLP) kernel. The Sigmoid Kernel comes from the Neural Networks field, where the bipolar sigmoid function is often used as an activation function for artificial neurons.

$$k(x,y) = \tanh(\alpha x^T y + c) \tag{5}$$

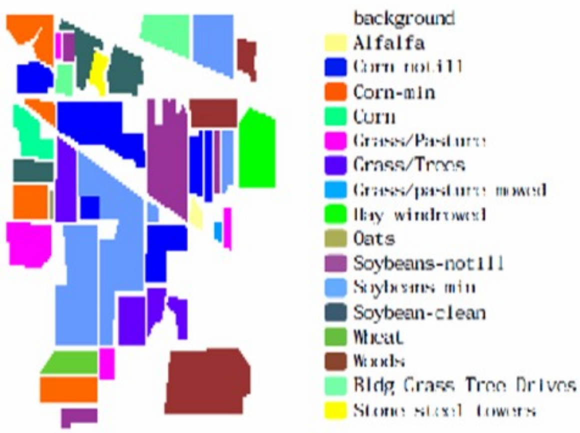
Polynomial kernel

The Polynomial kernel is a non-stationary kernel. Polynomial kernels are well suited for problems where all the training data is normalized. Adjustable parameters are the slope **alpha**, the constant term **c** and the polynomial degree **d**.

$$k(x, y) = (\alpha x^T y + c)^d \quad (6)$$

Experimental Design

The experiment is conducted on the AVIRIS hyperspectral dataset taken over the Northwestern Indiana's Indian Pine Set. The dataset consists of 220 bands and each band consists of 145x145 pixels. The original dataset contains 16 classes and ground truth is available for that.



Results and Discussion

The Feature set derived from the Feature Extraction method is used for classification. Randomly chosen pixels from each class and their corresponding feature vectors are used for training. The classifier produces output based on, whether the particular pixel under test belongs to the interested trained class or not. Thus, the pixels under the same class are separated from whole dataset.

Similarly the process is performed with other class each separating it from the rest. Thus individual classes have been separated and classified. The following is the combination of classes classified by using three different kernels. The overall accuracy for the individual classes for each kernel is 94.14, 95.70, 97.2 for sigmoid, linear and the sigmoid+ linear kernels respectively.

The following are the classes and their discussion with respect to the kernels employed. For the class alfalfa, the total number of pixels in this class is 54. The pixels which are omitted is either null or one. The commission error is less when employing sigmoid kernel than the other. This may be due to its orientation which by direct observation is not linear.

For the Class hay, there are totally 489 pixels belonging to this class. Here more or less the accuracy is same with respect to both linear and sigmoid kernels separately. So while employing the combination of these kernels the performance is still enhanced. But compared to linear kernel sigmoid is better. Here also the commission error is more whereas there is no pixels omitted or a very few of the pixels omitted.
For the Class corn, there are totally 234 pixels belonging to this class. Here also more or less the performance rate is same for both linear and sigmoid kernels. Whereas the performance is still enhanced when used with the combination of two kernels. Comparing both the individual classes employed linear kernel is little more in performance than the sigmoid kernel.

For the class bldg grass tree drives, there are totally 380 pixels belonging to this class. Here the commission error is more in case of sigmoid kernel compared to the linear kernel. So the performance is still higher when using the combination of the both. There is no omission error when using linear or the combination of linear or sigmoid kernel. The performance is robust when using the linear kernel and worst for the case of sigmoid kernel. This shows that its orientation supports only linear type of kernel and doest support sigmoid kernel.

For the class steel, the total number of pixels in this class is 95. Here the performance is almost similar for linear and sigmoid kernels. the classification is best for linear than the sigmoid case. Also when using the combination of both linear and sigmoid kernels the performance is still higher and robust.

For the class corn, here the total number of pixels belonging to this class is 234. The performance is better for Linear kernel than the sigmoid type. Also when using their combination the accuracy is still enhanced. when employing sigmoid kernel than the linear type. Also when the combination of kernels is employed the performance is still enhanced.

For the class wheat, the total number of pixels in this class is 212. Here the performance is almost same for all the three kernels that have been employed. However the accuracy is more for the combination of two kernels than with the individual kernels. For the class soynot, there are totally 968 number of pixels belonging to this class. The accuracy rate is higher for the sigmoid type than for the linear type. However when employing the combination of both the kernels the accuracy rate is still enhanced.

For the Class soymin there are totally 2468 number of pixels belonging to this class. The accuracy of classification is higher for linear type than the sigmoid kernel. Also when employing their combination the accuracy of the classification process is still higher. For the class Soyclean this class has total pixels of 968. The accuracy of classification is more for the linear type than the sigmoid or using their combination.

For the class Cornmin there are totally 824 numbers of pixels belonging to this class. The accuracy of classification is higher for the case of sigmoid type than the linear type of kernel. Also when the combination of both the kernels is employed the performance rate and hence the accuracy is still increased.

Linear kernel is good for corn, bgtd, steel, grass pasture, grass trees, soymin, and soycln. Whereas sigmoid for the remaining classes except oats and grass pasture mowed for which both the kernel performance rate is same.

The classification process is done by three different kernels and the performance of each kernel and their accuracies are tabulated and compared as above: By using three different kernels the classification process has been done for a simple comparison process. Thus the combination of linear and sigmoid kernels is robust among the rest. The classification of similar type of elements is based on their reflectance value which in turn is obtained by the spectral features extracted. It is from the classified results it is inferred that the performance is robust for the combination of two or more kernels than with the individual kernel. It is because some of the classes have been well classified in the certain type of kernel while some of the classes are well classified in other kernels. So when different kernels are combined the performance is enhanced compared to that of the individual kernels.

So it is predicted that the performance of the classifier depends upon the type of kernel employed which in turn depends upon the orientation of the classes in that image.

Hence depending upon the orientation of the classes the type of kernel must be used. For ex: the classes which are linearly oriented are suitable to use linear type kernel. And similarly for other kernels. The kernel type is also chosen according to the number For the class grass pasture, the total number of pixels belonging to this class is 294. The accuracy of classification is worst for sigmoid kernel and average in case of linear kernel. Here the combination of kernels does not support the classification process. For the class grass pasture mowed, here the total number of pixels belonging to this class is 26. Here the accuracy is more for the sigmoid type compared to the other two types of kernels.

For the class oats, the total number of pixels in this class is 20. Here the performance is similar for almost all the three kernels employed. This may be due to its very less number of pixels in this class. For the class grass trees, there are totally 747 pixels belonging to this class. The performance is average in case of sigmoid kernel which is slightly enhanced in linear type. So when the combination of both is employed the accuracy rate is still higher than employing individual classes. For the class woods the total number of pixels in this class is 1294. The classification accuracy is more of pixels present in that class. Also in ELM classifier the commission error is greater and mostly none of the pixels are omitted from classification. This is contrary to SVM classifier.

Here linear, sigmoid and a combination of both kernels are used. Each class is classified with certain accuracies. Also certain classes are well classified in particular type of kernels whereas with little accuracies in others. This bring a conclusion that the classes are classified based upon their orientation also. So the proper selection of kernel type is needed such that the chosen kernel should match the orientation of the class.

The performance of various kernels has been plotted with respect to its accuracies

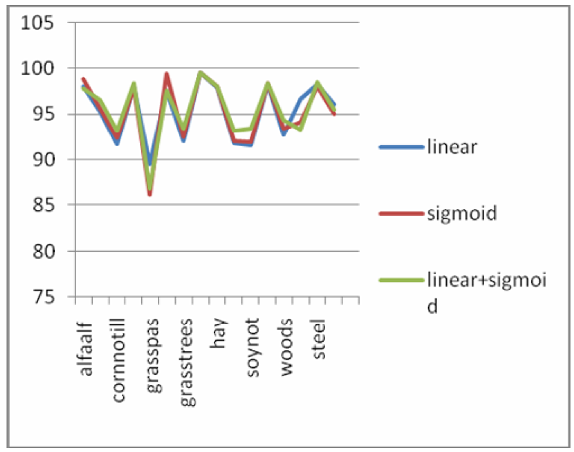


Fig 4: Performance of Various Kernel

From the above table graph is plotted and the following observation is made. It is obvious from the graph that Some of the classes are well classified for certain types of kernels and not in others .thus the kernel type must be chosen by considering the orientation of the classes. Also when combination of two or more kernels is used the accuracy rate is higher than the individual performance rate for most cases. Also the kernel type depends on the number of pixels belonging to particular class.

Table I: Accuracy Table

	SIGMOID KERNEL				LINEAR KERNEL				LINEAR + SIGMOID KERNEL			
Class	**Accuracy**	**Misclassified Pixels**	**Commission error**	**Omission error**	**Accuracy**	**Mis-classified Pixels**	**Commission error**	**Omission error**	**Accuracy**	**Mis-classified Pixels**	**Commission error**	**Omission error**
Alfalfa	98.8	264	233	1	98.02	416	416	0	97.9	444	444	0
Corn min	95.6	917	891	26	95.3	989	985	4	96.6	719	715	4
Cornn otill	92.3	1602	1545	57	91.7	1742	1666	76	93.23	1423	1377	46
Hay	98.0	415	413	2	97.9	422	416	6	98	406	406	0
Soynot	91.9	1693	1665	28	91.5	1777	1728	49	93.4	1382	1352	30
Soymi n	92.0	1674	1527	147	91.8	1720	1576	144	93.2	1424	1334	90
Soycln	95.0	1040	996	44	96	827	813	14	95.5	950	940	10
Wheat	98.4	337	336	1	98.2	367	367	0	98.5	326	325	1
Woods	93.	1364	1347	17	92.8	1522	1501	21	94.4	1178	1172	6
Steel	97.9	422	422	0	98.3	350	350	0	98.5	320	320	0
Oats	99.5	90	90	0	99.57	90	90	0	99.5	90	90	0
Grass pas	86.2	2902	2878	24	89.46	2217	2204	13	86.9	2762	2748	14
Grass pasm	99.4	117	117	1	97.6	495	495	0	97.6	495	495	0
Grass trees	92.5	1560	1540	20	92	1671	1655	16	93.4	1383	1364	19
Bgtd	91.1	1863	1859	4	96.7	704	704	0	93.4	1399	1399	0
corn	98.0	421	418	3	98	407	407	0	98.5	326	326	0

Conclusion

The proposed work developed a classification scheme that uses statistical features alone while classifying an image. It is predicted that those features alone gave. The accuracy of calculation depends upon the type of kernels used also depends on the orientation of objects. certain objects are well classified in certain kernels. So a combination of two or more kernels is found to give greater accuracies. Here in this case the combination of linear and sigmoid kernel produces more accuracies than classified by individual kernels.

Acknowledgment

The authors are grateful to the Management and Principal of our college for their constant support and encouragement.

References

[1] David Landgrebe,"Some Fundamentals and Methods for Hyper spectral Image data Analysis", SPIE Photonics, pp.1-10, Jan-1999

[2] C.-I. Chang, Hyperspectral Imaging: Techniques for Spectral Detection and Classification. Dordrecht, the Netherlands: Kluwer, 2003.

[3] G. F. Hughes, "On the mean accuracy of statistical pattern recognizers, "IEEE Trans. Inf. Theory, vol. IT-14, no. 1, pp. 55–63, Jan. 1968.

[4] Yuliya Tarabalka, Jon Atli Benediktsson, Jocelyn Chanussot, "Spectral–Spatial Classification of Hyperspectral Imagery Based on Partitional Clustering Techniques", IEEE Trans. Geosci. Remote Sensing, vol. 47, no. 8, pp. 2973–2987, Aug. 2009.

[5] D. A. Landgrebe, Signal Theory Methods in Multispectral Remote Sensing. New York: Wiley, 2003.

[6] P. K. Goel, S. O. Prasher, R. M. Patel, J. A. Landry, R. B. Bonnell, and A. A. Viau, "Classification of Hyperspectral Data by Decision trees and Artificial Neural Networks to identify weed stress and nitrogen status of corn," Comput. Electron. Agric., vol. 39, no. 2, pp. 67–93, 2003.

[7] H. Zhou, Z. Mao, and D. Wang, "Classification of coastal areas by airborne Hyperspectral image," in Proc. SPIE Opt. Technol. Atmos., Ocean, Environ. Stud., May 2005, vol. 5832, pp. 471–476.

[8] S. Subramanian, N. Gat, M. Sheffield, J. Barhen, and N. Toomarian, "Methodology for Hyperspectral Image Classification uses Novel Neural Network," in Proc. SPIE Algorithms Multispectral Hyperspectral Imagery III, A. E. Iverson, S. S. Shen, Eds., Aug. 1997, vol. 3071, pp. 128– 137.

[9] H. Yang, F. V. D. Meer, W. Bakker, and Z. J. Tan, "A BackPropagation Neural Network for mineralogical mapping from AVIRIS data," Int. J.Remote Sens., vol. 20, no. 1, pp. 97–110, 1999.

[10]Hernandez-Espinosa, M. Fernández-Redondo, and J. TorresSospedra, Some experiments with ensembles of neural networks for classification of hyperspectral images," in Proc. ISNN, vol. 1, 2004, pp. 912–917.

[11]Vaiphasa, "Innovative genetic algorithm for hyperspectral image classification, "in Proc. Int. Conf. Map Asia, 2003, p. 20.

[12]G. Camps-Valls and L. Bruzzone, "Kernel-based methods for Hyperspectral image classification," IEEE Trans. Geosci. Remote Sens.,

vol. 43,no. 6, pp. 1351–1362, Jun. 2005.

[13]M. Fauvel, "Spectral and spatial methods for the classification of urban remote sensing data," Ph.D. dissertation, Grenoble Inst. Technol., Grenoble, France, 2007.

[14]A. Gualtieri and R. F. Cromp, "Support Vector Machines for Hyperspectral remote sensing classification," Proc. SPIE, vol. 3584, pp. 221–232, Jan. 1998.

[15]M. Fauvel, J. Chanussot, and J. A. Benediktsson, "Evaluation of Kernels for Multiclass Classification of Hyperspectral remote sensing data," in Proc. ICASSP, May 2006, pp. II-813–II-816.

[16] Rick Archibald and George Fann," Feature Selection and Classification of Hyper spectral Images with Support Vector Machine", IEEE Trans. Geosci. Remote Sensing Letters, Vol.4, No.4, pp.674-678, October-2007.

[17]David Landgrebe, "Information Extraction Principles and Methods for Multispectral and Hyperspectral Image Data", pp.1-30, November 1998

[18]Gustavo Camps-Valls, Lorenzo Bruzzone, "Kernel-Based Methods for Hyperspectral Image Classification", IEEE Trans. Geosci. Remote Sensing, Vol.43, No.6; pp: 1351- 1363, June 2005

[19] Yakoub Bazi, Farid Melgani, "Toward an Optimal SVM Classification System for Hyper-spectral Remote Sensing Images", IEEE Trans. Geoscience. Remote Sensing, Vol.44, No.11; pp: 3374-3376, November 2006

[20] Asbjørn Berge, Anne H. Schistad Solberg, "Structured Gaussian Components for Hyper spectral Image Classification", IEEE Trans. Geosci. Remote Sensing,Vol.44, No.11; pp: 3386-3396, November 2006.

[21]Feng-Chia li "Comparison of the primitive classifiers with Extreme Learning Machine in credit scoring" Dept. of Inf. Manage., Jen-Teh Junior Coll., Miaoli, Taiwan IEEE international conference dated on 8-11 dec 2009 in pp 685 - 688

[22] Suresh.S of Electr. & Electron. Eng., Nanyang Technol. Univ., Singapore "Image Quality Measurement Using Sparse Extreme Learning Machine Classifier"9 th international conference dated on 5-8 dec 2006 on pp-1-6