A Novel Feature Derivation Technique for SVM based Hyper Spectral Image Classification

Mrs.K.Kavitha

Assistant Professor, Department of ECE
Mepco Schlenk Engineering College
Sivakasi, India

Dr.S.Arivazhagan

Professor & Head, Department of ECE
Mepco Schlenk Engineering College
Sivakasi, India

ABSTRACT

A spatial classification technique incorporating a novel feature derivation method is proposed for classifying the heterogeneous classes present in the hyper spectral images. The classification accuracy can be improved if and only if both the feature extraction and classifier selection are proper. As the classes present in the hyper spectral image are having different textures, textural classification is entertained. Wavelet based textural features extraction is entailed. Hyper spectral images are having dozen numbers of bands. Few mutually distinct bands are selected and wavelet transform is applied. For all the sub bands Gray Level Co-occurrence Matrix (GLCM) are calculated. From GLCMs co-occurrence features are derived for individual pixels. Apart from Co-occurrence features, statistical features are also calculated. Addition of statistical and co-occurrence features of individual pixels at individual bands form New Features for that pixel. By the process of adding these New Features of approximation band and individual sub-bands at the pixel level, Combined Features are derived. These Combined Features are used for classification. Support Vector Machines with Binary Hierarchical Tree (BHT) is developed to classify the data by One Against All (OAA) methodology. Airborne Visible Infra Red Imaging Sensor (AVIRIS) image of Cuprite –Nevada field is inducted for the experiment.

Categories and Subject Descriptors

I.4.6 [**Image Processing and Computer Vision**]: Segmentation-*Pixel Classification.*

General Terms

Algorithms, Experimentation.

Keywords

Multi-class, Wavelet, Co-occurrence features, Feature Extraction, Feature derivation, and Support Vector Machines.

1. INTRODUCTION

Designing the state of the art classifier for the Hyper Spectral image classification is the fastest growing technology in the fields of remote sensing, medicine, change detection, disaster management etc., Hyper spectral images produce spectra of several hundred wavelengths with fine spatial resolution. A Hyper spectral 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].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 hyper spectral images is not the trivial task, as it requires so many factors to be concerned such as :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. This phenomenon is known as Hughes Phenomenon [3] or 'curse of dimensionality'. As the consequence, 'overfitting' results. ie. the classifier performs well on the training samples and poor on the testing samples. So the Hughes phenomenon must be alleviated. 3) Non-linear spread of the data classes. To overcome these problems proper classifier should be selected. The classifier has to support the voluminous data, multi class data and the non-linear 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.

The hyper spectral classification falls into two major categories such as spectral classification and the spatial classification. In spectral classification, the reflectance values of the pixels at different wavelengths are considered. The pixels of different classes exhibit different reflectance at different wavelengths. From these wavelengths and reflectance, required spectral features such as mean reflectance, maximum reflectance, minimum reflectance, standard deviation, and variance can be calculated and can be used for classification. As this type of classification uses, spectral information it is known as *spectral classification.* In *spatial classification*, the spatial arrangement of pixels and their contextual values, textural properties are identified and features are extracted and classification can be done as in this experiment.

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. Support Vector Machines have shown a good performance for classifying high-dimensional data when a limited number of available training samples [12-15].

To improve the classification accuracy, apart from the statistical information, the contextual information should be considered. This neighborhood information improves the accuracy even more than the conventional spectral classification techniques such as Spectral Angle Mapping method. Camps-Valls proposes to estimate the local mean and variance for each pixel using fixed neighborhood [16]. Mercier proposes the same kernel formulation but the spatial information is estimated by wavelet decomposition of the image [17]. Mathieu Fauvel et.al. proposes the adaptive neighborhood of each pixel [18].Yuliya et.al. define the adaptive neighborhood along with the clustering information for the segmentation based classification [4].

The performance of the classifier inclines on the akin process known as feature extraction. The conventional feature extraction techniques are Decision Boundary Feature Extraction (DBFE) and Discriminate Analysis Feature Extraction (DAFE) [1], [19], [20]. DAFE is done based on the ratio of between class variance and within class variance [1]. The shortcoming of this method is its unreliability while the difference between the class mean vectors is small. DBFE determines the linear decision boundary between class variables and directly assigns samples. But it suffers when the training set is small [1]. Non parametric Feature extraction, by using scatter matrices is also possible for hyper spectral images [21]. To find out the inter relationship between the pixels, Gray Level Co-occurrence Matrix is useful [22].Co-occurrence features can be extracted from the GLCM to classify the textural information. Deriving the new features according to the application is an innovative process. Arivazhagan et.al. derived the new features for the decomposed wavelet sub-bands using both statistical and Co-occurrence features and used the derived features for target detection [23-24]. Hiremath et.al. generated the new features using the wavelet packet decomposition and are used to classify the color texture images [25-26]. Among the large number of extracted features, the recursive features can be eliminated by the method of feature selection [27] and the reduced set of significant features can be used for classification.

As classifiers are concerned, Spectral Angle Classifiers are classical one, used for hyper spectral image analysis in early works [28], 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. This is a time consuming work. For very complex boundaries of data this shows poor accuracy.

So, a classifier which is able to perform classification even with very complex boundaries and to overcome all the above cited problems was needed. As the Neural Network based classifiers and Support Vector Machine (SVM) are having that ability, they came into picture around 1990s. As the internal process of the Neural Network classifier is a hidden one and complex, a simple learning machine based on statistical learning theory became

popular [29]. More over SVM requires no pre processing on the data. Support Vector Machine advocates good results in the linear domain classifications and classification of the multi classes present in the image [30].

But, the hyper spectral domain is a non-linear one. Non-linear domain can be converted into the linear domain by using kernel trick. SVMs with kernel functions are used [31]-[34]. Many types of kernels like linear, polynomial, radial Basis Function (RBF), Sigmoid etc., are available. Selection of proper kernel gives proper results. SVM with Radial Basis Function (RBF) is a preferred combination which balances the complexity and accuracy [35]. Selection of optimal kernel can be done by Genetic Algorithm based manner [36]. Gaussian Radial Basis Function Kernel are also used according to the data type [37].

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 the features and deriving the novel features from the extracted features are the ever green processes. So, in this proposed methodology the correlation between neighborhood and Co-occurrence properties of individual pixels are calculated for all the sub-band of the wavelet decomposed image. New features are derived for the individual sub-bands obtained using both Co-occurrence and statistical features. By adding the New Features of the sub-bands, the Combined Features can be derived. Such Combined Features are expected to yield the better spatially classified image. Rest of the paper is organized in the following manner. Chapter-2 deals with the Proposed Work followed by the background of Support Vector Machine as Chapter-3. Chapter-4 describes the Experiment Design and Chapter-5 is dedicated to the Results and Discussions. Chapter –6 gives the conclusion about the work.

2. PROPOSED WORK
Classification is the process of assigning the pixels into the class based on the extracted features.

2.1. Feature Extraction

A feature is nothing but the significant representative of an image which can be used for classification, since it has a property which distinguishes one class from other. The extracted features provide the characteristics of input pixel to the classifier. The spatial features can be extracted by statistical and co-occurrence methods.

The wavelet transform is able to decorrelate the data and improves the classification rate. It provides orientation-sensitive information which is essential in texture analysis. The computational complexity is significantly reduced by considering the wavelet decomposition. So wavelet transforms are frequently used in classification process. All bands of the hyper spectral image are decomposed into sub bands by using wavelet transforms. For all the pixels in the individual sub-bands both statistical features and co-occurrence features are calculated. The mean, variance and standard deviation are the important statistical features. Statistical features provide the gray level information of pixels, while the co-occurrence features give the inter-pixel relationship. Spatial Gray Level Co-occurrence

Matrix (GLCM) estimates the image properties related to the second-order statistics. The features are calculated using the formulas given in Table 1.

In Texture analysis procedure, relative positions of pixels in image should be considered. In the co-occurrence matrix the distance can be chosen from 1 to 8 and the directions in the conventional way given in[22] which can be any one of $0^0, 45^0, 90^0, 135^0, 180^0, 225^0, 270^0, 315^0$. From the Co-occurrence matrix, eight different are extracted using the formulae in Table 1. While calculating the features, '1' distance and '0^0' direction is considered for this experiment.

Table 1. Co-occurrence Features

Feature	Formula
Entropy	Entropy $= -\sum_i \sum_j P[i,j] \times \log P[i,j]$
Energy	Energy $= \sum_i \sum_j P^2[i,j]$
Contrast	Contrast $= \sum_i \sum_j (i-j)^2 P[i,j]$
Homogeneity	Homogeneity $= \sum_i \sum_j (P[i,j]/(1+|i-j|))$
SumMean	SumMean $= (1/2)[\sum_i \sum_j iP[i,j] + \sum_i \sum_j jP[i,j]]$
Variance	Variance $= (1/2)[\sum_i \sum_j (i-\mu)^2 P[i,j] + \sum_i \sum_j (j-\mu)^2 P[i,j]]$
Maximum Probability	Max $\{P[i,j]\}$
Cluster Tendency	Cluster Tendency $= \sum_{i,j} (i+j-2\mu) k \, P[i,j]$

2.2. Feature Selection

All the extracted features are used for classification and accuracy is calculated. By dropping a feature, accuracy is recalculated. This process is repeated until the selection the significant features. This process of selecting the feature is known as Sequential Backward Selection (SBS) method. In the other feature selection process only one feature alone is considered and accuracy is calculated. By adding other features one by one, the accuracies are recalculated. This process of selecting the feature is known as Sequential Forward Selection (SFS) method. In this experiment SBS is incorporated and three most significant co-occurrence features (energy, entropy and contrast) are identified.

2.3. Derivation of Combined Features

The statistical mean and the significant Co-occurrence features such as energy, entropy and contrast are added up for individual pixels to form a New Feature for all the sub-bands (approximation, horizontal, vertical and diagonal sub-bands) individually. Combined Feature-I, II, III are calculated from the New feature.

2.3.1. List of Derived Features:

1. New Feature-I (NF-I)	:	Statistical mean+Co-occurence (energy+entropy+contrast) for Approximation Band
2. New Feature-II (NF-II)	:	Statistical mean+Co-occurence (energy+entropy+contrast) for Horizontal Band
3. New Feature-III (NF-III)	:	Statistical mean+Co-occurence (energy+entropy+contrast) for Vertical Band
4. New Feature-IV (NF-IV)	:	Statistical mean+Co-occurence (energy+entropy+contrast) for Diagonal Band

From these New Features, Combined Features can be derived.

|		
a) Combined Feature-I (CF-I)	:	NF-I+NF-II (Combination of feature in the Approximation and Horizontal Bands)
b) Combined Feature-II (CF-II)	:	NF-I+NF-III (Combination of feature in the Approximation and Vertical Bands)
c) Combined Feature-III (CF-III)	:	NF-I+NF-IV (Combination of feature in the Approximation and Diagonal Bands)

The derived Combined Features are used for spatial classifications.
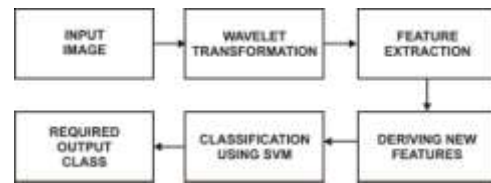
Fig. 1. Proposed Methodology

3. SUPPORT VECTOR MACHINES (SVM)

Support vector Machines performs the robust non-linear classification with kernel trick. It outperforms other classifiers even with small numbers of the available training samples.
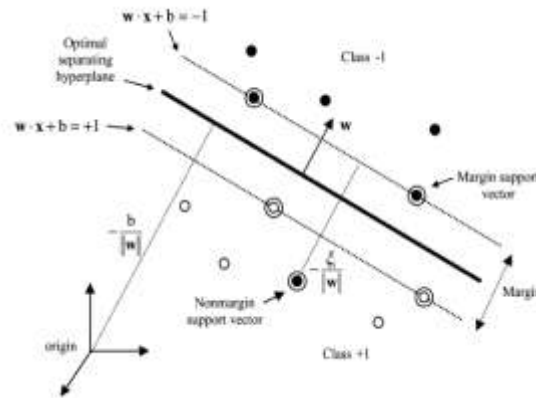
Fig. 2. SVM Hyperplane

SVM finds the separating hyper plane in some feature space inducted by the kernel function while all the computations are done in the original space itself [29]. For the given training set, the decision function is found by solving the convex optimization problem.

$$\max_{\alpha} g(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j k(x^i, x^j) \tag{1}$$

$$subject \ to \quad 0 \le \alpha_i \le C \, and \, \sum_{i=1}^{l} \alpha_i y_i = 0$$

Where 'α' are the Lagrange co-efficient. 'C' is a positive constant that is used to penalize the training errors and 'k' is the kernel.

When optimal solution is found, i.e. the α_i, the classification of a sample X is achieved by looking to which side of the hyper plane it belongs:

$$y = \text{sgn}\left(\sum_{i=1}^{l} \alpha_i y_i k(x^i, x) + b\right) \qquad (2)$$

Radial Basis Function (RBF) kernel is used in this experiment. The derived Combined Features are used for the classification.

One Against All (OAA) strategy is used while classifying the images. By which one class is separated from others. Thus the classes are separated hierarchically.
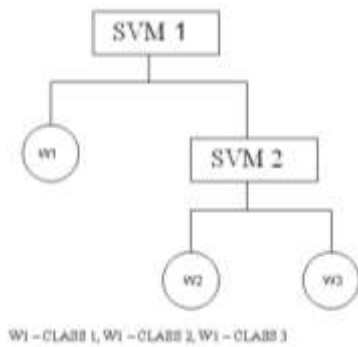

Fig. 3 OAA Binary Hierarchical Tree (BHT)

Like wise so many SVMs can be run to find out the interested classes. While training, care is shown towards the reduction of over fitting effect.

4. EXPERIMENT DESIGN

4.1. Hyper Spectral Data Preparation

To execute the proposed methodology, the Cuprite-Nevada: a site of geologic interest is chosen. The image has been captured by Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor. The real data set contains 8 different minerals. Iron rich minerals are abundant in this field. Among the original data set, this experiment aims towards only three minerals of interest such as Goethite, Hematite and Fe-Chlorite. Both goethite and Hematite are Iron bearing oxide minerals found in soil and exhibits slightly different reflectance in some wavelengths which is difficult to classify. But they are having different texture characteristics. Fe-Chlorite is a distinct iron bearing mineral. So, these three are selected as interested classes to be identified. AVIRIS sensor acquires data in the wavelength region from 0.38 μm to 2.50 μm and produces 220 spectral bands. The bands near 1.4 μm and 1.9 μm are eliminated, since they evidence the strong atmospheric absorption due to water vapor [22].

Around 1μm both Goethite and Hematite shows single dip, as indicated in the Fig.2 which can be useful to easily classify them. With the remaining 63 bands the near by bands which are having

correlation can be discarded by calculating the mutual information [23], [24] between bands.
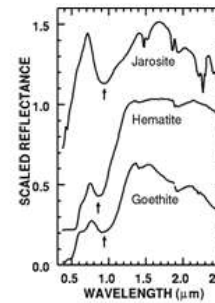

Fig. 3. Spectral Response of Minerals

4.2 Feature Vector Formation

As explained in the chapter -2, the combined features are derived for each pixel for the sub-bands obtained by the Discrete Wavelet Decomposition process. The same process is repeated for all the bands. The feature vector is formed by taking its combined feature at all the bands. Now, all the pixels are having eight different combined features and they are assigned as a vector. These vectors are used for classification.

$$CF^T = [CF_1, CF_2, CF_3, CF_4, CF_5, CF_6, CF_7, CF_8]$$

5. RESULTS AND DISCUSSIONS

Combined features (Statistical mean and Co-occurrence energy, entropy and contrast), by the feature extraction methods are used for classification. Neglecting other features, accuracy is minimally affected. Among the band combinations, CF-III (Approximation band combined with the Diagonal band) outperforms all other combinations. The result obtained using CF-III feature is shown in the Fig.3.

Pixels are randomly chosen from each class and their combined features are trained. Instead of testing the pixels of the respective class, all the pixels are tested against the training samples. The classifier produces the output, whether the pixel under test belongs to the interested class or not.

Thus the pixels of interested classes are identified among the whole data set. Similarly other classes are also trained. For these new trained values, total numbers of pixels are tested. By this way, classes are separated hierarchically. The training and testing information and v-fold cross validation accuracies are shown in the Table 2.

After identifying the pixels of interested class it is labeled and indicated by white gray level. All other pixels are assigned black gray level. Then the pixels are displayed. Like wise all the Bands are classified spatially and the Fig.5 gives the classified images.

Chlorite is classified better with the accuracy of 91.54% in the Band3, followed by Band4. Band7 and Band3 exhibits better classification rate for Hematite. Band8 alone gives good classification for Goethite.Band5 illustrates poor accuracy in case of Hematite, Goethite but it performs good at Chlorite, since the number of training samples are less compared to that of Chlorite. This is the same case for the Goethite in Band4 and Band3.

The number of pixels belong to Chlorite is less compared to that of Goethite and Hematite. So, Chlorite is classified well even with smaller number of training samples.

Moreover the inter-pixel relationship dictates the classification always. As the contrast is one of the factors, which is included in the classification, high contrast neighborhood shows good classification.

Accuracies are calculated by the v-fold cross validation method in which the training set is sub-divided into v-number of subsets and accuracies are cross validated and results are shown in table 2.
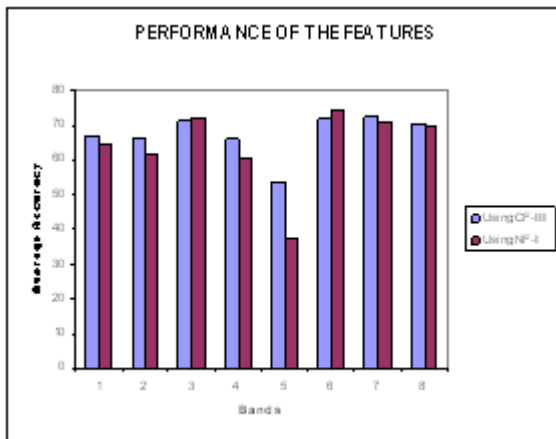
Table 2.Classification Details

Bands		Class1 (Chlorite)	Class2 (Hematite)	Class3 (Goethite)
Band1	Training	320	362	229
	Testing	37179	37179	37179
	Accuracy	71.65	55.92	73.58
Band2	Training	229	352	403
	Testing	37179	37179	37179
	Accuracy	75.98	61.55	60.40
Band3	Training	362	193	344
	Testing	16059	16059	16059
	Accuracy	91.54	75.47	47.28
Band4	Training	397	427	334
	Testing	16059	16059	16059
	Accuracy	75.266	73.89	47.82
Band5	Training	450	346	369
	Testing	16059	16059	16059
	Accuracy	73.47	40.46	47.65
Band6	Training	320	345	456
	Testing	15900	15900	15900
	Accuracy	83.65	67.76	63.64
Band7	Training	347	288	318
	Testing	15900	15900	15900
	Accuracy	64.54	89.76	62.45
Band8	Training	389	180	311
	Testing	16059	16059	16059
	Accuracy	77.87	55.84	76.61

Fig. 4. Average Accuracy for different Features

For the quantitative comparison, the average accuracies of the bands for features CF-III and NF-I is tabulated and graphically represented in Fig.4.
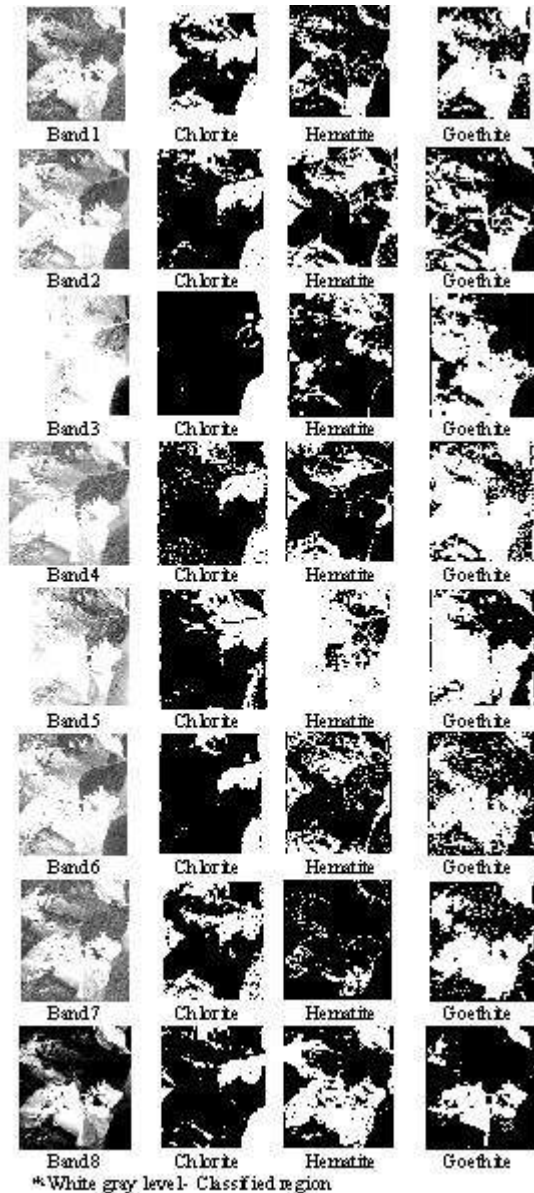
Table 3. Average Accuracies for the Features

Band/ Feature used	Using CF-III	Using NF-I
Band1	67.05	64.17
Band2	65.97	61.56
Band3	71.43	71.8
Band4	65.66	60.43
Band5	53.86	37.12
Band6	71.68	74.15
Band7	72.72	70.67
Band8	70.11	69.72

From the graph it is evident that the proposed derived feature acts as a potential member for most of the bands. It gives the improved accuracy which will be useful for the textural classifications.

Fig. 5. Classified Images

Even though the accuracies are lesser for the experiment, this type of textural classification can be used as a metric to cross validate the performance other classifiers. Moreover, this textural classification can be used to classify the objects having the identical reflectance or other similar spectral properties at the same wavelength but having different textural properties.

6. CONCLUSION

The Co-occurrence features can be calculated for different distance and for different directions from the Co-occurrence matrix and can be used for classification. Different types of composite kernels and different types of wavelets can be selected and the accuracies can be tested. It is possible to develop a soft classification algorithm for this type of sensitive

classifications. For this case of analysis, only knowing 'which' class a pixel belongs to is not sufficient. If, the information about, 'how much' the pixel belongs to a particular class is known, it will be more useful for classification. As the Soft Classification extends its application to the sub-pixel levels, it can able to reduce the misclassifications and can be used for this kind of applications..

7. ACKNOWLEDGEMENT

8. 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, Jón 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 using 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 Back-Propagation Neural Network for mineralogical mapping from AVIRIS data," Int. J.Remote Sens., vol. 20, no. 1, pp. 97–110, 1999.

[10] Hernández-Espinosa, M. Fernández-Redondo, and J. Torres-Sospedra, 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] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Vila-Franc´es, and J. Calpe-Maravilla, "Composite Kernels for Hyperspectral Image Classification," IEEE Geosci. Remote Sens. Lett., vol. 3, no. 1, pp. 93–97, Jan. 2006.

[17] G. Mercier and F. Girard-Ardhuin, "Partially supervised oil-slick detection by SAR imagery using kernel expansion," IEEE Trans. Geosci. Remote Sens., vol. 44, no. 10, pp. 2839–2846, Oct. 2006.

[18] Mathieu Fauvel , Jocelyn Chanussot , Jon Atli Benediktsson, "Adaptive Pixel Neighborhood Definition for the Classification of Hyperspectral images with Support Vector Machines and Composite Kernel, in Proc. ICASSP, May 2006, pp. II-813–II-816.

[19] C. Lee and D. A. Landgrebe, "Feature Extraction based on Decision Boundaries, "IEEE Trans. Pattern Anal. Mach. Intell., vol. 15, no. 4, pp. 388–400, Apr. 1993.

[20] L. O. Jimenez and D. A. Landgrebe, "Hyperspectral data analysis and Supervised Feature Reduction via Projection Pursuit," IEEE Trans. Geosci. Remote Sens., vol. 37, no. 6, pp. 2653–2667, Nov. 1999.

[21] Bor-Chen Kuo, David A. Landgrebe, "Nonparametric Weighted Feature Extraction for Classification", IEEE Trans. Geosci. Remote Sens., Vol. 42, No. 5, pp.1096-1105, May 2004

[22] R. M. Haralick, K. Shanmugam, I. Dinstein, "Texture Features for image classification." IEEE Trans. System Man Cybernetics, Vol.8, No.6, pp. 610-621,1973

[23] S. Arivazhagan, L. Ganesan,S.DeivaLakshmi, "Texture Classification using Wavelet Statistical Features", Journal of the Institution of Engineers(India),Vol.85, pp.51-55, 2003

[24] S. Arivazhagan, L. Ganesan, "Texture Classification using Wavelet Transform", Pattern Recognition Letters, Vol.24, pp. 1513-1521, 2003.

[25] P.S.Hiremath, S.Shivashankar, "Wavelet Based Features For Texture Classification", GVIP Journal, Vol. 6, No. 3,pp.55-58, December, 2006.

[26] 16. P.S.Hiremath, S. Shivashankar, and Jagadeesh Pujari," Wavelet Based Features For Color Texture Classification With Application To CBIR", International Journal of Computer Science and Network Security, Vol.6 No.9A,pp.124-132, September 2006.

[27] 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.

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

[29] V. N. Vapnik, Statistical Learning Theory. New York: Wiley, 1998.

[30] Farid Melgani, Lorenzo Bruzzone, "Classification of Hyperspectral Remote Sensing Images With Support Vector Machines", IEEE Trans. Geosci. Remote Sensing, vol. 42, no. 8, pp. 1778–1790, August 2004.

[31] Gustavo Camps-Valls, Joan Vila, "Composite Kernels for Hyper Spectral Image Classification", IEEE Trans. Geosci. Remote Sens. Letters, vol. 3, no. 1, pp. 93–97 January 2006

[32] M. Murat Dundar and A. Landgrebe, "A Cost-effective Semisupervised Classifier Approach with Kernels," IEEE Trans. Geosci. Remote Sens., vol. 42, no. 1, pp. 264–270, Jan. 2004.

[33] Mingmin Chi, "Semi Supervised Classification of Hyperspectral Images by SVMs Optimized in the Primal", IEEE Trans. Geosci. Remote Sensing, Vol. 45, no. 6, pp. 1870–1880, June 2007.

[34] Gustavo Camps-Valls, Tatyana V. Bandos Marsheva, Dengyong Zhou, "Semi-Supervised Graph-Based Hyperspectral Image Classification", IEEE Trans. Geosci. Remote Sensing, Vol. 45, no. 10, pp. 3044–3054, October 2007.

[35] 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

[36] 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

[37] 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.

[38] Brian J. Ross and Anthony G. Gualtieri, Frank Fueten, Paul Budkewitsch, " Hyper spectral Image Analysis Using Genetic Programming", Genetic and Evolutionary Computation Conference, pp: 1196 - 1203, 2002.

[39] Baofeng Guo, Steve R. Gunn, R. I. Damper, James D.B. Nelson ,"Customizing Kernel Functions for SVM-Based Hyper spectral Image Classification", IEEE Image Processing, Vol. 47, no. 4, pp. 622–629, Apr-2008.

[40] David Landgrebe, "Information Extraction Principles and Methods for Multispectral and Hyper spectral Image Data", pp.1-30,November1998.DOI: www:cobweb.ecn.purdue.edu/~landgreb/GeoComp99.pdf