International Journal of Scientific Engineering and Technology (ISSN : 2277-1581)
Volume 2 Issue 5, pp : 383-387 1 May 2013

Role of Zernike Moments in Hyperspectral Image Classification

Kavitha K.[1], Arivazhagan S.[2], Dhivya Priya R.[3], Kanaga Sangeetha I.[4]
[1,3,4]Deptt. of E.C.E., Mepco Schlenk Engineering College, Sivakasi, Tamilnadu, India.
[2] Principal, Mepco Schlenk Engineering College, Sivakasi,Tamilnadu, India.
.
[1]vihadakavi@yahoo.com, [2]s_arivu@yahoo.com, [3]dheepa.dhivya@gmail.com, [4]sangeethaik90@gmail.com

Abstract—*Classification of heterogeneous classes present in the Hyperspectral image is one of the recent research issues in the field of remote sensing. This work presents a novel technique that classifies Hyperspectral images that contain number of classes by making use of the image moments. Recently, researchers have introduced a number of neural network models and structured output based methods for classification of these Hyperspectral images they however suffers with the problem of confusion between the classes that are having similar characteristics and hence provides imbalanced solution for the classes with less number of pixels. The polynomial features such as Zernike moments are extracted from the Hyperspectral image and is used for classification. Support Vector Machines with Binary Hierarchical Tree is used for classification of the Hyperspectral data by One Against All methodology. Then, the performance of Zernike moments in Hyperspectral image classification is evaluated.*
Keywords— Zernike moments, Hyperspectral Image Classification, Multi-class Classifier, Support Vector Machine, AVIRIS.

I. INTRODUCTION

Classifying the heterogeneous classes present in the Hyperspectral Image is one of the recent research issues in the field of remote sensing. It is demonstrated that using Hyperspectral image to make quantitative classification and identification of crop or vegetation characteristics is successful [1]. However, the high dimensionality of the Hyperspectral data makes it more difficult to use it for classification. Moreover, there is a lot of redundancy in the data which needs to be removed [2]. Complexity lies in the nature of high dimensionality data and the consequent ground truth demand for supervised classification [3]. Hughes phenomenon implies that the number of labeled training samples required for supervised classification increases as a function of dimensionality. In Remote Sensing, the availability of training samples is often limited and this limitation becomes relevant in case of high number of features. This problem is addressed by identifying a model that is less sensitive to Hughes phenomenon provided it should reduce the redundancy of the dataset available [4].

Classifying the pixels in the Hyperspectral Images and identifying their relevant class belongings depends on the feature extraction and classifier selection processes. Feature extraction is an akin process while classifying images. In Remote Sensing, the number of training samples available is often limited. Hence to avoid the problems caused by the limited training samples, several feature extraction methods based on the wavelet transform have been proposed for Hyperspectral images [5]. Chebyshev Statistics based on Chebyshev Polynomials gives accurate representation of image features. The Chebyshev Polynomial analysis is a form of transformation, where the necessary features are extracted to preserve the second order image statistics [6]. Zernike moments in general are rotation invariant and robust to noise. The higher order Zernike moments represent the global shape of a pattern whereas the lower order Zernike moments grasps the finer details of the pattern [7]. They are proven to have very good image feature representation capabilities, are based on the orthogonal Zernike Radial Polynomials [8]. They are able to achieve a near-zero value of redundancy measure in a set of moment functions where the moments correspond to independent characteristics of the image [9].

In general, Hyperspectral datasets are non-linear in nature. The Kernel methods that are used in machine learning models converts these non-linear datasets into a linear one, thereby making it a useful tool for applications like regression, classification and clustering [10-11].These Kernel methods are suitable for the classification of high dimensional data while the availability of the training samples is limited. Many types of kernels such as Linear, Polynomial, Sigmoid, Radial Basis Function (RBF) etc., are available. Selection of proper kernel gives proper results.

Support Vector Machine (SVM) with Radial Basis Function (RBF) is a preferred combination which balances the complexity and accuracy [12]. The usage of SVM classifier for Hyperspectral Images is shown by J.Gualtieri [13]. Multiclass classifier for Hyperspectral images is explained by B.Scholkopf [14]. Huang and Dixon applied the SVM classification to the Landsat Thematic Mapper (TM) image classification and compared the results with the maximum likelihood classifier (MLC), the neural network classifier, and the decision tree classifier [15-16]. The results show that the SVM achieved higher classification accuracy than those of the other classifiers. In addition, the SVM was also applied to multispectral remotely sensed image classification and achieved higher classification accuracy than those of other traditional classification methods [17]. Foody applied the SVM algorithm to classify the airborne image, and the results indicated that the SVM often achieved a higher accuracy than those of other classification methods [18].

From the literature, it is evident that the image moments are also capable of efficient image representation and hence they are used as features for classification. First the Zernike Moments are extracted as features from the Hyperspectral image. Then the extracted features are used for efficient classification of Hyperspectral image using a SVM classifier.

Rest of the paper is organized as follows. Section-2 deals with the Proposed Work followed by the Experiment Design as Section-3. Section-4 is dedicated to the Results and Discussions. Section-5 gives the conclusion about the work.

II. PROPOSED METHODOLOGY

A. *Zernike Moments*

Zernike features are the coefficients of the Zernike polynomial approximation of the image. While Zernike coefficients are complex-valued, absolute values are used as image descriptors. Zernike first introduced a set of complex polynomials $\{V_{nm}(x,y)\}$ which form a complete orthogonal set over the unit disk of $x^2 + y^2 \leq 1$ in polar coordinates. Since the Zernike moments are defined over the unit circle, two steps were required to convert a rectangular region of each image to a unit circle for calculation of Zernike moments. In the first step, the range of the image should be mapped to the unit circle with its origin at the image's centre. Secondly, the pixels falling outside the unit circle are discarded in the computation process. For example, if we want to compute Zernike moments of a binary image with spatial resolution of 64×64. This binary image is normalized into a unit circle with fixed radius of 32 pixels. The form of the polynomials is defined as:

$$V_{nm}(x,y) = R_{nm}(\rho,\theta) = R_{nm}(\rho)e^{im\theta} \quad (1)$$

where n is positive integer or zero; m is integers subject to constraints n-$|m|$ is even, and $|m| \leq n$; ρ is the length of the vector from the origin to the pixel (x, y); θ is the angle between the vector ρ and x axis in counter-clockwise direction; $R_{nm}(\rho)$ is Radial polynomial defined as:

$$R_{nm}(\rho) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s!\left(\frac{n+|m|}{2}-s\right)!\left(\frac{n-|m|}{2}-s\right)!} \rho^{n-2s} \quad (2)$$

Here, Zernike moments can be calculated using analytical forms of Zernike polynomials instead of complex Zernike polynomials. The Zernike moment of order n with repetition m for function $f(x, y)$ is defined as:

$$A_{nm} = \frac{n+1}{\pi} \iint f(x,y)\, V_{nm}*(x,y)dxdy,\ x^2 + y^2 \leq 1 \quad (3)$$

Where, $V_{nm}*(x,y) = V_{n,-m}(x,y)$.

To compute the Zernike moment of a digital image, we just need to change the integrals with summations:

$$A_{nm} = \frac{n+1}{\pi} \sum_x \sum_y f(x,y) V_{nm}*(x,y)dxdy,\ x^2 + y^2 \leq 1 \quad (4)$$

Suppose we know all Zernike moments A_{nm} of $f(x, y)$ up to order N, we can reconstruct the image by:

$$f'(x,y) = \sum_{n=0}^{N} \sum_{m=0}^{N} A_{nm} V_{nm}(x,y) \quad (5)$$

Zernike polynomials take the unit disk $x^2 + y^2 \leq 1$ as their computation domain. To compute the Zernike moments of a digital image, the range of the image should be mapped to the unit circle with its origin at the image's centre. The lower order Zernike moments capture the overall shape information, and the higher order moments capture the fine details of the image. In addition to these, Zernike moments possesses many attractive properties, such as Rotation Invariance, Invariance to Illumination condition, Robust to Noise, Expression Efficiency, Fast Computation and Excellent multi-level image feature representation capabilities. Besides these, they may also achieve scale and translation invariance properties, if the input image is normalized before the computation of moments.

B. *Feature extraction*

When the input data to an algorithm is too large to be processed and it is suspected to be redundant, then the input data will be transformed into a reduced representation of set of features (also named feature vectors). Transforming the input data into the set of features is called feature extraction. From all the bands of Hyperspectral dataset, the Zernike moments of order $n=3$ are computed and the results are used as features for Hyperspectral image classification with the help of a classifier named SVM.

C. *Support Vector Machine (SVM)*

SVM is a good candidate for remote sensing data classification for a number of reasons. Firstly, an SVM can work well with a small training data set as the selection of a sufficient number of pure training pixels has always been a

International Journal of Scientific Engineering and Technology (ISSN : 2277-1581)
Volume 2 Issue 5, pp : 383-387 1 May 2013

problem with remote sensing data. Secondly, SVM has been found to perform well with high accuracy for problems involving hundreds of features. SVM performs nonlinear classification using kernel trick. Kernel-based methods are based on mapping data from the original input feature space to a kernel feature space of higher dimensionality and solving a linear problem in that space.

The aim is to find a linear separating hyper plane that separates classes of interest. The hyper plane is a plane in a multidimensional space and is also called a decision surface or an optimal separating hyper plane. Radial Basis Function Kernel is used here.

This kernel nonlinearly maps samples into a higher dimensional space, unlike the linear kernel, and so it can handle the case when the relation between class labels and attributes is nonlinear as shown in Figure. 1.

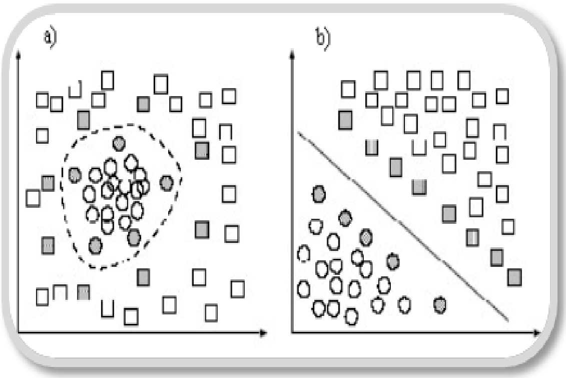

Fig 1. Separable classification using RBF kernel in a) Original Space b) Feature Space

Here, One-against-all (OAA) Binary Hierarchical Tree strategy is used by SVMs while classifying images. It separates the classes hierarchically by considering how one class is separated from others.

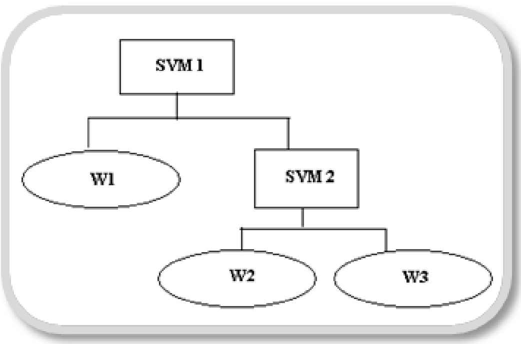

Fig 2. OAA Binary Hierarchical Tree (W_i represents $class_i$)

Likewise so many SVMs can be run to find out the interested classes. While training, care is shown towards the over fitting of the samples.

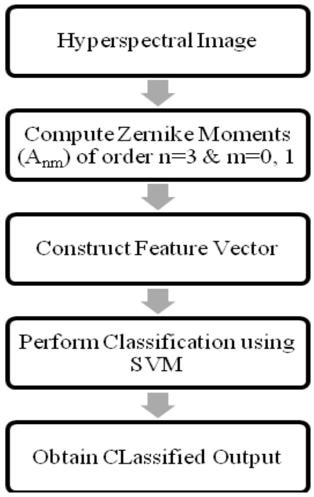

Fig. 3 Proposed Methodology

III. EXPERIMENTAL DESIGN

The experiment is conducted on the AVIRIS Hyperspectral dataset taken over the North-western 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.

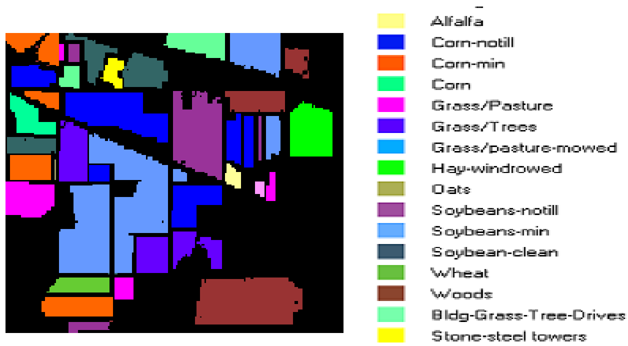

Fig 4. Ground Truth Classes

IV. RESULTS AND DISCUSSIONS

The Feature Vectors derived from the Feature Extraction methods are 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 other classes also trained and by this method, the classes are separated hierarchically.

After that the pixels of interested class are assigned white gray level while others are assigned black. Then the output is displayed and the performance is evaluated. By comparing each class of pixels with the ground truth, the amount of

International Journal of Scientific Engineering and Technology (ISSN : 2277-1581)
Volume 2 Issue 5, pp : 383-387 1 May 2013

misclassified pixels can be found. From that the accuracy can be found.

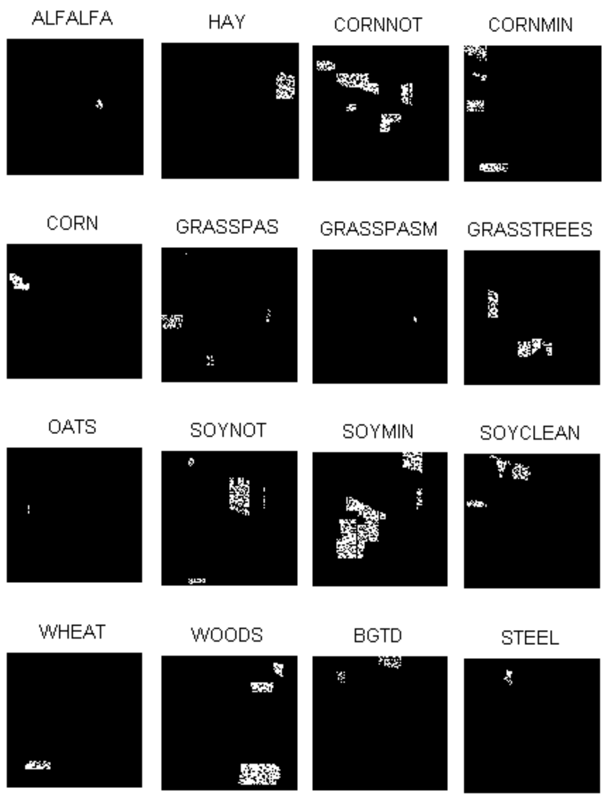

Fig 5. Classification Output

TABLE I : AVERAGE ACCURACY (AA) FOR 16 CLASSES

Class label	class	Average accuracy (%)
C1	Alfalfa	66.67
C2	Corn-notill	57.26
C3	Corn-Min	66.25
C4	Corn	59.95
C5	Grass-Pasture	89.74
C6	Grass-Trees	40.24
C7	Grass-Pasture Mowed	61.54
C8	Hay-Windrowed	68.27
C9	Oats	60
C10	Soybean-notill	61.98
C11	Soybean-Min	64.06
C12	Soybean-Clean	56.51
C13	Wheat	56.60
C14	Woods	61.82
C15	Building Grass Tree Drives	65.00
C16	Stone Steel Tower	28.42
Overall Accuracy		**72.7%**

From Table 1, it is inferred that, for nearly three-fourth of the classes, the average accuracy obtained by the proposed feature set reaches more than 60%. Classes like Grass-pasture and Woods produces accuracy more than 75% and classes like Alfalfa, Corn-min, Hay-windrowed produces accuracy more than 65% and classes like Grass pasture- mowed, Soy- notill, Soy-min, Wheat and Stone-steel towers produces average accuracy more than 60% whereas classes like Corn-notill, Corn and Building Grass Tree Drives attains average accuracy of 40%. By comparing the classification result with the ground available, the overall accuracy of classification is obtained and is found to be 72.7%.

Fig 6. Pseudo Coloured Output of SVM

V. CONCLUSION

The classification of Hyperspectral remote sensing data using Support Vector Machines was investigated. SVM provides very accurate classification, even in the case of a very limited number of training sample and high dimensional data. Feature like Zernike Moments from Raw Image is used for classification using SVM. It is predicted that this feature alone gave overall classification accuracy of 72.7%. It is possible to develop a soft classification algorithm for this type of sensitive classifications. For this case of analysis, knowledge about '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.

ACKNOWLEDGMENT

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

REFERENCES

i. *Shwetank, Jain Kamal & Bhatia K. J, Review of Rice Crop Identification and Classification using Hyper-Spectral Image*

Processing System, International Journal of Computer Science & Communication, vol. 1, no. 1, pp. 253-258,2010.

ii. *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*

iii. *D. Landgrebe, Information Extraction Principles and Methods for Multispectral and Hyperspectral Image Data, Information Processing for Remote Sensing, pp.3-37, World Scientific Singapore, 1999.*

iv. *Gallo A, E. Binaghi A, M. Boschetti B, P.A. Brivio B, A Neural Adaptive Model for feature selection and Hyperspectral data classification, Mirco Boschetti in Topology,2004.*

v. *Hsu, P.H., Spectral Feature Extraction of Hyperspectral Images Using Wavelet Transform, Ph.D. Thesis, Department of Surveying Engineering, National Cheng Kung University, Tainan, Taiwan, R.O.C, 2003.*

vi. *Zaid Omar, Nikolaos Mitianoudis and Tania Stathaki, Region-Based Image Fusion using a combinatory Chebyshev-ICA Method, ICASSP, pp.1213-1216, 2011.*

vii. *W. Kim, Y. Kim, A region-based shape descriptor using Zernike moments, Signal Processing: Image Communication, vol.16, no.1, pp. 95-10, 2000.*

viii. *Thawar Arif, Zyad Shaaban, Lala Krekor, Sami Baba, Object Classification via Geometrical, Zernike and Legendre Moments, Journal of Theoretical and Applied Information Technology, vol.7, no.1, pp.31-37, 2009.*

ix. *R. Mukundan, K.R. Ramakrishnan, Moment Functions in Image Analysis—Theory and Applications, World Scientific, Singapore, 1998.*

x. *B. Scholkopf, A. Smola, Learning with Kernels—Support Vector Machines, Regularization, Optimization and Beyond, MIT Press Series, 2002.*

xi. *G. Camps-Valls, J. L. Rojo-Alvarez, and M. Martinez-Ramon,Kernel Methods in Bioengineering, Signal and Image Processing, Idea Group Publishing, Hershey, PA, 2007.*

xii. *G. Camps-Valls, L. Bruzzone , Kernel-based methods for Hyperspectral image classification,* IEEE Trans. Geosci. Remote Sens. *vol. 43, no. 6, pp. 1351-1362, 2005.*

xiii. *J.Gualtieri and R. CrompSupport vector machines for Hyperspectral remote sensing classification, Proc. 27th AIPR Workshop Advances in Computer Assisted Recognition, Washington DC, pp. 121-132, 1998,.*

xiv. *B. Scholkopf, A. Smola, Learning with Kernels—Support Vector Machines, Regularization, Optimization and Beyond, MIT Press Series.*

xv. *C. Huang, L. S. Davis, and J. R. G. Townshend, An assessment of support vector machines for land cover classification, Int. J. Remote Sens., vol. 23, no. 4, pp. 725–749, Feb. 2002.*

xvi. *B. Dixon and N. Candade, Multispectral landuse classification using neural networks and support vector machines: One or the other, or both, Int. J. Remote Sens., vol. 29, no. 4, pp. 1185–1206, Feb. 2008.*

xvii. *G. Zhu and D. G. Blumberg, Classification using ASTER data and SVM algorithms: The case study of Beer Sheva, Israel, Remote Sens. Environ., vol. 80, no. 2, pp. 233–240, May 2002.*

xviii. *G. M. Foody and A. Mathur, A Relative evaluation of multiclass image classification by support vector machines, IEEE Trans. Geosci. Remote Sens., vol. 42, no. 6, pp. 1335–1343, Jun. 2004.*