World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:4, No:10, 2010

Combined Feature Based Hyperspectral Image Classification Technique Using Support Vector Machines

Mrs.K.Kavitha and S.Arivazhagan

***Abstract*—**A spatial classification technique incorporating a State of Art Feature Extraction algorithm is proposed in this paper for classifying a heterogeneous classes present in 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 in the hyper spectral images are assumed to have different textures, textural classification is entertained. Run Length feature extraction is entailed along with the Principal Components and Independent Components. A Hyperspectral Image of Indiana Site taken by AVIRIS is inducted for the experiment. Among the original 220 bands, a subset of 120 bands is selected. Gray Level Run Length Matrix (GLRLM) is calculated for the selected forty bands. From GLRLMs the Run Length features for individual pixels are calculated. The Principle Components are calculated for other forty bands. Independent Components are calculated for next forty bands. As Principal & Independent Components have the ability to represent the textural content of pixels, they are treated as features. The summation of Run Length features, Principal Components, and Independent Components forms the Combined Features which are used for classification. SVM with Binary Hierarchical Tree is used to classify the hyper spectral image. Results are validated with ground truth and accuracies are calculated.

***Keywords*—**Multi-class, Run Length features, PCA, ICA, classification and Support Vector Machines.

I. INTRODUCTION

HYPER Spectral image classification is the fastest growing technology in the fields of remote sensing and medicine.

Hyper spectral images produce spectra of several hundred wavelengths. A Hyper spectral image has hundreds of bands, where as a multi spectral has 4 to 7 bands only [1]. So, the hyper spectral images provide more information for within class discrimination, i.e., they can discriminate between different types of rocks and different types of vegetation while multi spectral images can discriminate between rock and vegetation. Such hyper spectral images can be used effectively to classify the heterogeneous classes which are present in the image. For classification of multi-class information in an image, it is necessary to derive proper set of features and to choose proper classifier.

K.Kavitha is the Assistant Professor with the Department of Electronics and Communication Engineering, Mepco Schlenk Engineering College, Sivakasi-626 005, TamilNadu ,India (Corresponding Author to provide phone: 09486488787; e-mail: vihadakavi@yahoo.com).

Dr. S.Arivazhagan is Professor and Head with the Department of Electronics and Communication Engineering, Mepco Schlenk Engineering College, Sivakasi-626 005, TamilNadu ,India (sarivu@mepcoeng.ac.in).

Classification of hyper spectral images is not the trivial task, as it requires so many factors to be concerned, such as

(i) the large number of land-cover class to be handled; (ii) High number of spectral bands but low number of the availability of training samples. This phenomenon is known as Hughes Phenomenon [2] or 'curse of dimensionality'. As the consequence, 'over fitting' results. i.e., the classifier performs well on the training samples and poor on the testing samples. This Hughes phenomenon must be alleviated; (iii) Non-linear spread of the data classes. To overcome these problems proper classifier should be selected in such a way that it has to support voluminous data, multi class data and the non-linear dataset. Moreover proper features must be fed to the classifiers to obtain the high accuracy rate.

In spatial classification, the spatial arrangement of pixels and their contextual values, textural properties are identified by means of the extracted features and classification can be done. To analyse the textural properties, the repetitiveness of the gray levels and the primitives of the gray levels which can distinguish the different classes must be retrieved. By this way, the basic attribute known as feature can be used to identify the region of interest.

An extensive literature is available on the pixel-level processing technique, i.e., techniques that assign each pixel to one of the classes based on its extracted features [3]. Galloway (1975) et.al explained the usage of conventional Run Length features such as Short Run Emphasis (SRE), Long Run Emphasis (LRE), Gray Level Non uniformity (GLN), Run Length Non uniformity (RLN) and Run Percentage for texture analysis [4]. Haralick et.al. (1978) used the Co-Occurrence features for classifying the textures [5]. Run Length features are used to analyze the natural textures in [6]. Xiaoou Tang (1998) introduces the Dominant Run Length feature such as Short-run Low Gray Level Emphasis, Short-run High Gray Level Emphasis, Long-run Low Gray Level Emphasis and Long-run High Gray Level Emphasis to extract the discriminant information for successful classification [7].

Maneesha Singh and Sameer Singh (2002) used the Run Length features to analyze the textures and they compared the performance of RL features with that of other texture features such as GLCM features [8] and the recognition rate was calculated. Poonguzhali & Ravindran (2008) extracted the RL features to classify the focal lesion in ultrasound live images [9]. Wiselin Jiji and Ganesan (2009) used the Run Length features along with the unsupervised method for classification of textures [10]. Renuca Methre & M.

Ravindranath extracted the wavelet based RL feature for texture image retrieval application [11].

Canonical Analysis such as PCA [12] and ICA can also be used for feature extraction. These features are expected to yield the better results for hyper spectral image classification. Andreas weingessel (2000) developed a fast algorithm to calculate the Principal Components [13].

Suyken et. Al (2003) introduced the Principal Component Analysis along with SVM formulation [14]. Wavelet based and lifting based PCA has been introduced by Yuka Higashijima et. al. [15]. Zhan – Li Sun et. al.(2005) uses PCA along with Fuzzy C – means clustering for extracting the features of multispectral images [16].

Apart from Principal Component Analysis, Independent Component Analysis also plays a vital role in extracting features. Independent Components provides the statistically independent random variables behind the object of interest which will be useful for classification [17]-[20] Shah et. al (2004) obtained ICA features for hyperspectral image classification [21]. Palmason (2005) extracted the independent component along with morphological preprocessing to classify the urban area from the hyperspectral data [22]. Jian Yang et. al (2007) constructed a ICA baseline for face recognition along with PCA [23]. Xiao benlin et. al (2008) used the ICA for change detection & classification of multispectral images [24]. Weibo Zou et.al (2002) used the wavelet domain for extracting the Independent Components [25].

All the above mentioned works extracted features such as Run Length Features, Principal Components and Independent Components and individually used the same to classify the images. In this proposed work, the analysis is made when the features are summed up and are used for classification.

As classifiers are concerned, Spectral Angle Classifiers are classical one, used for hyper spectral image analysis in early works [24] in which the input image spectra are compared with reference spectra for classification. The classes are 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 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 theory became popular [27]. Unlike neural networks, SVM requires no pre processing on the data. Support Vector Machines advocate good results in the linear domain classification [28].

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 [28]-[29]. 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 computational complexity and accuracy [31].

From the literature, it is evident that the care must be shown towards the feature extraction and classifier selection. The runs of gray levels are useful in identifying the objects or classes. Principal and Independent Components are also the representatives of the classes. So Run Length features, Principal Components and Independent Components are combined and are used for classification. More over innovative feature extraction and feature selection are the ever wanted processes. In this proposed methodology the Gray Level Run Length properties of individual pixel are calculated and Run Length features are derived and are used for classification. Combined Features are generated from RL features and Principal and Independent Components. Such Combined Features are expected to yield the better spatially classified image. The rest of the paper is organized in the following manner. Section-II deals with the Proposed Work followed by the Experiment Design as Section-III. Section-IV is dedicated to the Results and Discussions. Section–V gives the Conclusion about the work.

II. Procedure for Paper Submission

Classification is the process of assigning the pixels into the class, based on the derived features.

A. Feature Extraction

As a feature is the significant representative of an image, it can be used to distinguish one class from other. Feature extraction is the akin process while classifying the images. The extracted features exhibits characteristics of input pixel which the basic requirement of the classifier to make decisions about the class belonging of the pixels. The spatial features can be extracted by statistical and Run Length methods.

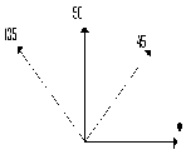
B. Run Length Matrix

Run Length Matrix (RLM) can be used to extract the features from the image components. RLM gives the repetitive runs of each grey levels, there by gives the inter-relation and the texture primitives. A gray level run is a set of consecutive points having the same gray level value. The number of pixels in the run is known the length of the run which reflects the size of the texture element.

With the observation that, in a coarse texture, relatively long gray level runs would occur more often and that a fine texture should contain primarily short runs, The matrix element (i,j) specifies the number of times the image contains Run Length j in the given direction having the gray level i. which gives the distinguishing primitive of the image. For a given image, an element p(i,j) in the Run Length matrix ‘p’ is defined as the number of runs with gray level ‘i’ and Run Length ‘j’. The run-length matrices are calculated in all the four directions (0^0, 45^0, 90^0, 135^0) as shown. For all the directions, Run-Length Matrices are found and features are extracted.

These features can be used for classification and are shown in the Fig.1.

Run Length Orientations



(a)

Sample Image

1	1	1	2	2
3	4	2	2	3
4	4	4	4	4
5	5	3	3	3
1	1	3	4	5

(b)

RLM for 0^0

GRAY LEVEL	RUN LENGTH: 0	1	2	3	4	5
	1	0	1	1	0	0
	2	0	2	0	0	0
	3	3	0	1	0	0
	4	2	0	0	0	1
	5	1	1	0	0	0

(c)

RLM for 45^0

GRAY LEVEL	RUN LENGTH: 45	1	2	3	4	6
	1	5	0	0	0	0
	2	0	2	0	0	0
	3	4	0	1	0	0
	4	6	1	0	0	0
	5	3	0	0	0	0

(d)

RLM for 90^0

GRAY LEVEL	RUN LENGTH: 90	1	2	3	4	5
	1	5	0	0	0	0
	2	2	1	0	0	0
	3	4	1	0	0	0
	4	6	1	0	0	0
	5	3	0	0	0	0

(e)

RLM for 135^0

GRAY LEVEL	RUN LENGTH: 135	1	2	3	4	5
	1	5	0	0	0	0
	2	4	0	0	0	0
	3	6	0	0	0	0
	4	6	1	0	0	0
	5	3	3	0	0	0

(f)

Fig.1 Run Length Matrix Calculation

The traditional run-length features are,

1. Short Run Emphasis (SRE)

$$SRE = \frac{1}{nr}\sum_{i=1}^{M}\sum_{j=1}^{N} p(i,j)/j^2 = \frac{1}{nr}\sum_{j=1}^{N} pr(j)/j^2$$

2. Long Run Emphasis (LRE)

$$LRE = \frac{1}{nr}\sum_{i=1}^{M}\sum_{j=1}^{N} p(i,j).j^2 = \frac{1}{nr}\sum_{j=1}^{N} pr(j).j^2$$

3 .Gray Level Non uniformity (GLN)

$$GLN = \frac{1}{nr}\sum_{i=1}^{M}(\sum_{j=1}^{M} p(i,j))^2 = \frac{1}{nr}\sum_{i=1}^{M} pg(i)^2$$

4. Run Length Non uniformity (RLN)

$$RLN = \frac{1}{nr}\sum_{j=1}^{N}(\sum_{i=1}^{M} p(i,j))^2 = \frac{1}{nr}\sum_{j=1}^{N} pr(j)^2$$

5. Run Percentage

$$RP = \frac{nr}{np}$$

where P is the run-length matrix, P(i,j) is an element of the run-length matrix at the position (i,j) and 'nr' is the number of runs in the image. 'np' is the number of pixels.

The attainable accuracy is comparatively lesser while using traditional run-length features,. In order to improve the accuracy of classification four new run-length features are extracted and are known as Dominant Run Length features. Dominant Run Length features use the gray values of pixels along with their runs.

C. Dominant Run Length features

The four dominant Run Length features are,

1. Short-run low gray level emphasis

A.

$$SRLGE = \frac{1}{nr}\sum_{i=1}^{M}\sum_{j=1}^{N} P(i,j)/(i^2.j^2)$$

2. Short-run high gray level emphasis

B.

$$SRHGE = \frac{1}{nr}\sum_{i=1}^{M}\sum_{j=1}^{N} P(i,j).i^2/j^2$$

3. Long-run low gray level emphasis

C.

$$LRLGE = \frac{1}{nr}\sum_{i=1}^{M}\sum_{j=1}^{N} P(i,j).j^2/i^2$$

4. Long-run high gray level emphasis

$$LRHGE = \frac{1}{nr}\sum_{i=1}^{M}\sum_{j=1}^{N} P(i,j).i^2.j^2$$

where P is the run-length matrix, P(i,j) is an element of the run-length matrix at the position (i,j) and 'nr' is the number of runs in the image. 'np' is the number of pixels. For all the four principle directions, the dominant features are calculated.

D. PCA

PCA is optimal in the mean square sense for data representation. So, the hyper spectral data are reduced from several hundreds of data channel into few data channels. Hence the dimensionality can also be reduced without losing the required information.

The steps involved in Principal Component Extraction are as follows:

1. Get the image
2. Calculate mean and covariance matrix
3. Calculate Eigen vectors from the co-variance matrix
4. Find Eigen values
5. Order the covariance matrix by Eigen value highest and lowest.
6. Eigen vectors with highest Eigen value is the Principal Component, which contains significant information
7. Lowest Eigen value components can be discarded.

The significant components known as Principal Components can be used as the features for classification as they depict more information about the image and are good representatives of image.

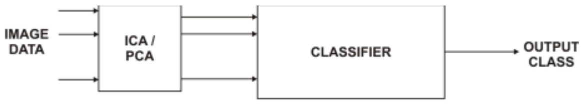


Fig. 2 ICA Process

E. ICA

It is a statistical technique which reveals the hidden factors. ICA separates the data into underlying information components. As ICA separates the underlying information components of the image data, it can be used as a feature extraction technique. ICA generates the variables which are not only decorrelated, but also statistically independent from each other. PCA makes the data uncorrelated while ICA makes the data as independent as possible. This property is useful in discriminating the different classes in the image.

ICA features can be used for classification, as they exhibits high order relationship among image pixels and ICA seeks to find the directions of projections of the data .

ICA features are higher order uncorrelated statistics. While PCA features are second order uncorrelated statistics. Second order statistics may inadequate to represent all the objects in the image. To exploit the information from the multivariate data such as hyper spectral data, higher order statistics are required. ICA features are capable of finding the underlying sources in such applications in which PCA features fail.

ICA concept can better be analysed by considering a Cocktail Party Problem [22]. Imagine a room where two people are speaking simultaneously. Two microphones held at different locations. The microphones give two recorded time signals, which can be denoted by $x_1(t)$ and $x_2(t)$, with x_1 and x_2 the amplitudes, and t the time index. Each of these recorded signal is a weighted sum of the speech signals emitted by the two speakers, which can be denoted by $s_1(t)$ and $s_2(t)$. This could be expressed as a linear equation:

$$x_1(t) = a_{11}s_1 + a_{12}s_2 \quad (1)$$

$$x_2(t) = a_{21}s_1 + a_{22}s_2 \quad (2)$$

where a_{11}, a_{12}, a_{21}, and a_{22} are some parameters that depend on the distances of the microphones from the speakers. It would be very useful to estimate the two original speech signals s_1 (t) and s_2 (t), using only the recorded signals $x_1(t)$ and $x_2(t)$. This is called the cocktail-party [22] problem.

To illustrate this analysis, two signal are shown in Fig.1. The original speech signals could look something like those in Fig. 1 and the mixed signals could look like those in Fig. 2. The problem is to recover the data in Fig. 1 using only the data in Fig. 2.

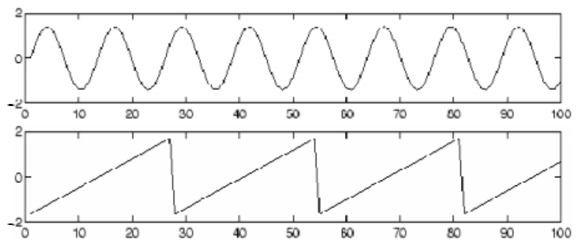


Fig.3 The Original Signals

Actually, if the parameters a_{ij} are known, then the linear equation in (1) is solved by classical methods. If a_{ij} is unknown the problem is considerably more difficult. For most of the practical applications aij is unknown.

To solve this problem and to estimate a_{ij} the statistical properties of the signals $s_i(t)$ are used. It is observed that the signals $s_1(t)$ and $s_2(t)$ are statistically independent at each time instant 't'. This property can be used to separate the original signals $s_1(t)$ and $s_2(t)$ from the observed mixture of signals $x_1(t)$ and $x_2(t)$.

This cock-tail party problem can also be extended to the images. Independent Components can be obtained from the images. The extracted components can have the ability to highlight the desired class.

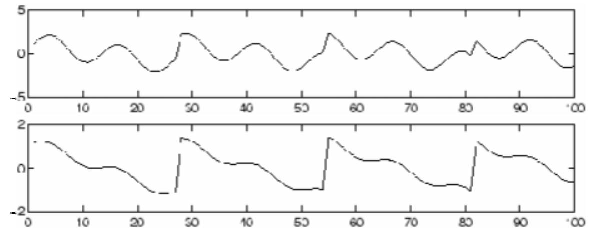


Fig.4 The observed mixtures of the source signals

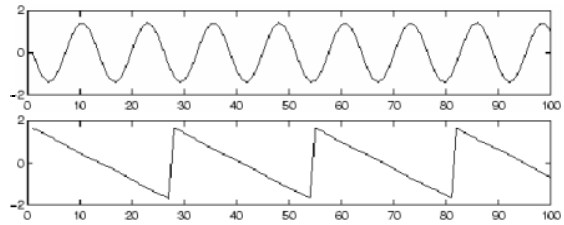


Fig.5 The estimates of the original source signals

Formation of New Features:

Gray Level Emphasis is extracted and is added up for individual pixels to form a new feature for all the directions separately. As first five Principal Components depict the information of 97.8%, they are alone selected and summed up. It is known as New Feature –V. As all the Independent Components are having the meaningful and distinct information, all the Independent Components are added together to form the New feature-VI.

New Feature-I	: Summation of RL features in the Direction of 0^0
New Feature-II	: Summation of RL features in the Direction of 45^0
New Feature-III	: Summation of RL features in the Direction of 90^0
New Feature-IV	: Summation of RL features in the Direction of 135^0
New Feature-V	: Summation of First five Principal Components
New Feature-VI	: Summation of All the Independent Components

Combined Features are formed separately for all the Directions.

Combined Feature for *0⁰*

Combined Feature-I (CF-I) : NF-I
Combined Feature-II (CF-II) : NF-I+NF-V
Combined Feature-I (CF-III) : NF-I+NF-VI
Combined Feature-I (CF-I) : NF-I+NF-V+NF-VI

Combined Feature for 45^0

Combined Feature-I (CF-I) : NF-II
Combined Feature-II (CF-II) : NF-II+NF-V
Combined Feature-I (CF-III) : NF-II+NF-VI
Combined Feature-I (CF-I) : NF-II+NF-V+NF-VI

Combined Feature for 90^0

Combined Feature-I (CF-I) : NF-III
Combined Feature-II (CF-II) : NF-III+NF-V
Combined Feature-I (CF-III) : NF-III+NF-VI
Combined Feature-I (CF-I) : NF-III+NF-V+NF-VI

Combined Feature for 135^0

Combined Feature-I (CF-I) : NF-IV
Combined Feature-II (CF-II) : NF-IV+NF-V
Combined Feature-I (CF-III) : NF-IV+NF-VI
Combined Feature-I (CF-I) : NF-IV+NF-V+NF-VI

The generated Combined Features are used for spatial classifications. The overall proposed methodology is shown in the pictorial representation form as shown in the Fig. 4

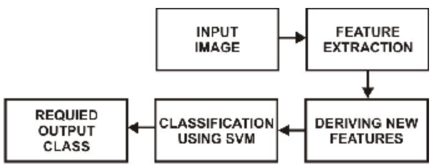


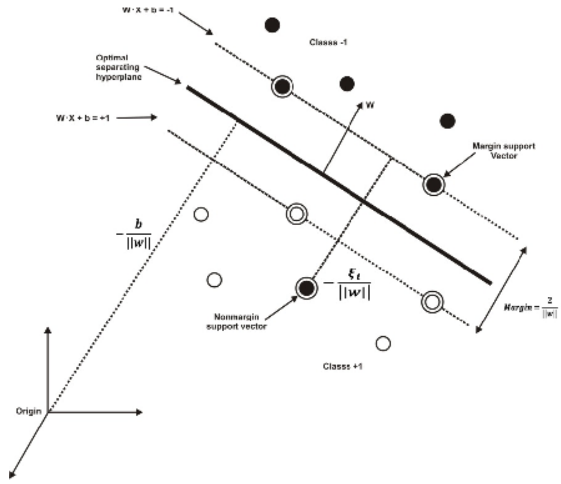
Fig. 6 Proposed Methodology

F. Classification

Support Vector Machines use the statistical learning theory which maximizes the distance between training samples of two classes. This approach gives SVMs a high generalization capacity which requires only a few training samples without compromising accuracy requirement and outperforms other classifiers.

Support vector Machines performs the robust non-linear classification with kernel trick. 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. For the given training set, the decision function is found by solving the convex optimization problem.

SVM is used for both spatial and spectral classifications. SVM is having the capability of separating the non-linear classes by introducing kernel functions. Radial Basis Function (RBF) kernel is used in this experiment. 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.

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

$$subject\ \ to\quad 0 \le \alpha_i \le C\, and \sum_{i=1}^{l} \alpha_i y_i = 0 \tag{3}$$

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:

$$y = \operatorname{sgn}\left(\sum_{i=1}^{l} \alpha_i y_i k(x^i, x) + b \right) \tag{4}$$

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.

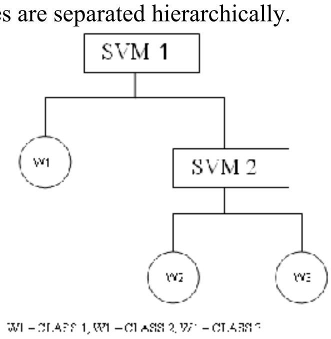


Fig. 7 OAA Binary Hierarchical Tree (BHT)

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

III. Experiment Design

For evaluating the performance of the proposed method, a sample hyper spectral image which is taken over northwest Indiana's Indian pine test site is selected. The site is sensed by AVIRIS. The data consists of 220 bands and each band consists of 145 x 145 pixels. This data set is chosen because its ground truth is available. The original data set contains 16 different land cover classes as shown in Table. 1. Among 16 classes, three classes are discarded due to the insufficient number of training samples. The remaining 13 classes are selected to be classified. During the learning phase of classification, 2596 training samples are selected. 2900 testing samples are used for validating the performance. While using SVM, RBF kernel is used.

IV. Results and Discussions

Combined Features derived from the feature extraction methods are used for classification. Pixels are randomly chosen from each class and their Combined Features are used for training. All other pixels are tested against the training samples. The classifier produces the output, whether the pixel under test belongs to the interested trained class or not. Thus the pixels of interested classes are identified among the whole data set. Similarly other classes are also trained. Randomly selected pixels are tested against the training samples. By this way, classes are separated hierarchically. The training and testing information are shown in the Table 1.

Fig. 8 Ground Truth Classes

After identifying the pixels of interested class it is labelled and indicated by white gray level. All other pixels are assigned black gray level. Then the pixels are displayed. The output is subtracted from the labelled ground truth and number of misclassified pixels is calculated. The accuracy of each class is calculated and is shown Table 2.

In coarse texture like Alfafa, the gray level runs are longer. But fine texture like buildings takes the short runs. Likewise, the runs of gray levels are varying for each class. So, it is possible for classifying the classes by using Run Length features. But Run Length features fail to distinguish the classes which contain more or less same textural properties. For identifying those classes, which are outliers while using Run Length features, Principal Components and Independent Components are used.

For the class Soy-min, the Run Length features predominates the Principal Components since Soy-min class has the good primitive information for classification.

Table 1 Classification Details

Sl.No.	Class	Class No.	No. of Training samples used	No. of Testing Samples
1	Wheat	C1	126	140
2	Corn-min	C2	144	280
3	Soy-min	C3	408	400
4	Soy notilll	C4	330	360
5	Woods	C5	527	550
6	Grass Pasture	C6	198	150
7	Alfalfa	C7	18	60
8	Building	C8	80	100
9	Oats	C9	18	40
10	Grass-Trees	C10	184	200
11	Soy-Clean	C11	140	200
12	Hay-windrowed	C12	315	320
13	Corn-notill	C13	108	100
Total			**2596**	**2900**

By observing the accuracies, it is evident that the Combined Features yield the good accuracies for all the classes. For classes like Soybean-min, Alfalfa, the Run Length features have the potential for high accurate classification. The classes like oats and Grass-Trees, the Run Length features with 45^0 give good results. Corn-notill emits good accuracy for Run Length features with 135^0.

After adding the Principal Components to the Run-Length features, the accuracy is increased by the good amount. From the scatter plot which shows the distribution of Run Length features, Principal Components and Independent Components, it is evident that they are having the mixed values or non-linear values which cannot be linearly separable by the low level classifiers like minimum distance classifier or maximum likelihood classifiers etc.

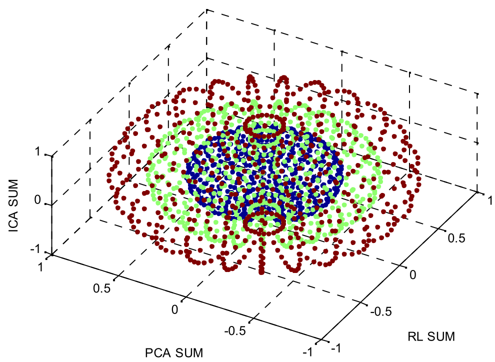


Fig. 9 Feature Distribution Plot

But these can be classified by SVMs by incorporating the trick. The RBF kernel is used and it produces sufficient accuracies.

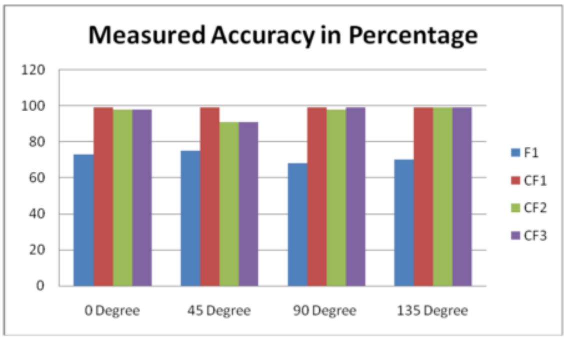


Fig.10 Accuracy Comparison

V. Conclusion

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.

Acknowledgement

The authors are grateful to Prof. David. A. Landgrebe for providing the AVIRIS data set along with the ground truth and for providing the MultiSpec package.

References

[1] David Landgrebe, "Some Fundamentals and Methods for Hyper Spectral Image Data Analysis", *SPIE Photonics,* pp.1-10, Jan-1999.
[2] 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.
[3] Yulia Tarabalka, Jon Atli Beneditksson, Joycelyn Chanussot,"Spectral-Spatial Classification of Hyperspectral Imagery based on Partitional Clustering Techniques", IEEE Trans. On Geoscience and Remote sensing,Vol-47,No-8,Aug 2009, pp 2973-2987.
[4] Galloway,"Texture information in Run Length Matrices",IEEE Transactions on Image Processing,Nov-1998,Vol.7,No.11, pp:1602-1609.
[5] 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.
[6] Horng-Hai Loh, Jia Guu Leu," The Analysis of natural Textures using Run Length Features", IEEE Transactions on Industrial Electronics, May 1988, Vol.35,No.2, pp:323-328
[7] Xiaoou Tang,"Dominant Run- Length Method for Image Classification" Department of Applied Ocean Physics and Engineering, Woods Hole Oceanographic Institution, Woods Hole, Report-MA 02543.
[8] Maneesha Singh and Sameer Singh,"Spatial Texture Analysis: A Comparative Study", IEEE conference on Image Processing, 2002, pp:676-679.
[9] Poonguzhali, and G.Ravindran," Automatic classification of focal Lesion in Ultrasound Liver Images using Combined Texture features, Information Technology Journal, 2008, pp:205-209.
[10] Wiseline Jiji and L.Ganesan," Unsupervised Texure Classification, Journal of Theoretical and Applied Information Technology,2009, pp:373-381
[11] Renuka Methra, and M.Ravindranath, " Exploring Spatial Information in Spectral features for Texture Image retrival", International Journal of Computer and network Security, Dec.2009,Vol 1,No.3, pp:46-49
[12] Lindsay I Smith, "A tutorial On Principal Components Analysis", ICIC International Journal, Febuary2002, Vol. 4, No.5
[13] Andreas Weingessel and Kurt Hornik, "Local PCA Algorithms", IEEE Transactions on Neural Networks, November 2000, Vol.11, No.6.
[14] J. A. K. Suykens, T. Van Gestel, J.Vandewalle, and B. De Moor, "A Support Vector Machine Formulation to PCA Analysis and Its Kernel Version", IEEE Transactions on Neural Networks, March 2003, Vol. 14, No. 2, pp:447-450
[15] Yuka Higashijima, Shigeru Takano, and Koichi Niijima, "Image Classification by Lifting Wavelet PCA", Journal of Kyushu University, Japan.
[16] Zhan-Li Sun,De-Shuang Huang,Yiu-Ming Cheung,"Using FCMC, FVS, and PCA techniques for feature Extraction of Multispectral Images, IEEE Transactions on Geosciences and Remotesensing Letters,Apr-2005, Vol.2,No.2, pp: 108-112
[17] M.Asuncion Vicente, Patrik O.Hoyer and Aapo Hyvarinen, "Equivalence of Some Common Linear Feature Extraction Techniques for Appearance-Based Object Recognition Tasks", IEEE Transactions on Pattern Analysis and Machine Intelligence, May 2007, Vol.29, No.5
[18] A. Hyvärinen, J. Karhunen, and E. Oja, "Independent Component Analysis"
[19] Hyvarinen, A. and E. Oja, "A fast fixed-point algorithm for independent component analysis", Journal of Neural Computation, vol.9, pp.1483-1492, 1997
[20] Mark D. Plumbley, and Erkki Oja, "A Nonnegative PCA Algorithm for Independent Component Analysis", IEEE Transactions on Neural Networks, January 2004, Vol.15, No. 1.
[21] C.A.Shah, P.Watanachaturaporn,p.K.varshney,M.K.Arora," Some recent results on Hyperspectral Image Classification", IEEE conference on Geosciences and remotesensing,2004, pp:346-353.
[22] Jon Aevar Palmason, Jon Atli Benedikson,Johannes R.Sveinsson,Jocelyn Chanussot," Classification of Hyperspectral data from Urban Areas using Morphological Preprocessing and Independent Component Analysis, IEEE conference on Geosciences and Remotesensing,2005,Vol.1, pp:176-180
[23] Jian Yang, David Zhang, and JingYu Yang, "Constructing PCA Baseline Algorithms to Reevaluate ICA-Based Face-Recognition Performance", IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics, August 2007, Vol.37, No. 4, pp:1015-1021
[24] Xiao Benlin Li Fangfang,Mao Xingliang, Jin Huazhong, " The International Archives of the photogrammetry, remote Sensing and Spatial Information sciences, Beijing 2008, Vol. XXXVII, Part B7.
[25] Weibo Zou, Yan Li King Chuen Lo, Zheru Chi," Improvement of Image Classification with Wavelet and Independent Component Analysis(ICA) based on a structrured Neural Networks, International Joint Conference on Neural Networks, July 2006, pp:3949-3954.
[26] David Landgrebe, "Information Extraction Principles and Methods for Multispectral and Hyper spectral Image Data", pp.1-30, November 1998.
[27] V. N. Vapnik, Statistical Learning Theory. New York: Wiley, 1998.
[28] Farid Melgani, Lorenzo Bruzzone, "Classification of Hyperspectral Remote Sensing Images With Support Vector Machines", IEEE Trans. Geosci. Remote Sensing, August 2004,vol. 42, no. 8, pp. 1778–1790
[29] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for Hyperspectral image classification," IEEE Trans. Geosci. Remote Sens., Jun. 2005, vol. 43,no. 6, pp. 1351–1362
[30] M. Fauvel, "Spectral and spatial methods for the classification of urban remote sensing data," Ph.D.dissertation, Grenoble Inst. Technol., Grenoble, France, 2007.
Gualtieri and R. F. Cromp, "Support Vector Machines for Hyperspectral remote sensing classification," Proc. SPIE, vol. 3584, pp. 221–232, Jan. 1998.
[31] M. Fauvel, J. Chanussot, and J. A. Benediktsson,"Evaluation of Kernels for Multiclass Classification of Hyperspectral remote sensing data," in Proc. ICASSP, May2006, pp. II-813–II-816.
[32] Gustavo Camps-Valls, Lorenzo Bruzzone, "Kernel-Based Methods for Hyperspectral Image Classification", IEEE Trans. Geosci. Remote Sensing, June 2005, Vol.43, No.6; pp:1351-1363

TABLE II
CLASSIFICATION ACCURACY DETAILS

Direction	Feature/Class	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	Average Accuracy
0^0	RL Sum	68.00	47.00	98.64	60.41	74.00	67.00	93.17	63.00	95.00	90.00	68.00	51.00	74.00	73.02
	RL Sum+PCA	99.70	99.74	94.54	99.82	99.81	99.70	100	99.89	97.24	98.43	99.58	99.53	99.00	98.99
	RL SUM+ICA	99.78	100.00	93.32	99.81	99.76	100	100	99.37	98.73	98.44	97.66	99.28	99.93	98.92
	RLS+PCA+ICA	99.73	99.62	93.62	99.81	99.68	99.46	100	98.74	97.24	98.63	97.66	99.02	99.88	98.69
Direction	Feature/Class	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	Average Accuracy
45^0	RL Sum	53.00	61.00	97.00	68.00	71.00	90.00	93.00	63.00	95.00	90.00	68.00	51.00	74.00	74.92
	RL SUM+PCA	99.58	99.54	93.11	99.58	99.8	99.22	99.98	99.63	99.09	99.33	97.65	99.75	96.96	98.70
	RL SUM+ICA	99.57	99.99	93.49	97.31	99.15	99.37	99.9	99.87	97.24	99.12	97.67	99.73	98.25	90.99
	RLS+PCA+ICA	99.87	99.57	93.11	99.27	97.65	99.9	99.9	98.92	97.27	98.09	97.65	99.73	97.21	90.80
Direction	Feature/Class	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	Average Accuracy
90^0	RL Sum	67.00	46.00	98.00	60.00	74.00	67.00	94.00	55.62	66.88	81.00	44.00	50.00	77.33	67.75
	RL SUM+PCA	99.71	99.74	94.98	99.84	99.8	99.69	100	99.84	98.34	99.32	97.65	99.72	97.26	98.91
	RL SUM+ICA	99.73	99.88	94.32	99.84	99.94	99.87	99.99	99.84	98.11	97.73	99.25	99.33	94.96	98.67
	RLS+PCA+ICA	99.79	100	94.19	99.69	99.84	99.77	99.99	99.68	98.43	99.38	98.94	99.69	98.95	99.10
Direction	Feature/Class	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	Average Accuracy
135^0	RL Sum	69.00	48.00	98.00	60.00	76.20	67.00	97.00	53.00	79.00	82.00	44.19	50.24	83.00	69.74
	RL SUM+PCA	99.59	99.25	94.98	99.80	99.20	99.73	99.90	98.85	97.24	99.48	98.77	99.73	98.77	98.86
	RL SUM+ICA	99.68	99.66	94.54	99.82	99.78	99.69	99.96	99.56	98.36	99.31	99.53	99.73	96.90	98.96
	RLS+PCA+ICA	99.68	99.88	94.52	99.30	99.76	99.36	99.96	98.78	97.24	97.05	97.65	99.68	98.79	98.59