# SOME ISSUES RELATED WITH SUB-PIXEL CLASSIFICATION USING HYPERION DATA

A. Kumar[a*], H. A. Min[b],

[a]Indian Institute of Remote Sensing, Dehradun, India. anil@iirs.gov.in
[b]Satellite Data Processing and Digital Mapping Station, Union of Myanmar, hlaingaungmin@gmail.com

**Commission VI, WG VI/4** *ISPRS-2008-WgS – PS: WG VII/3* – **Information Extraction from Hyperspectral Data**

**KEY WORDS:** Hyperspectral Data, Sub-pixel, Mixed Pixels, Support Vector Machine (SVM), Fuzzy Error Matrix (FERM).

**ABSTRACT:**
The most significant breakthrough in remote sensing has been the development of hyperspectral sensors. Due to the spectral resolution limitations of conventional multispectral remote sensing, hyperspectral remote sensing techniques have been introduced in 1990s. Natural earth surface is composed of heterogeneous surface features, so the features in the satellite data also have mixed pixel effect. The mixed pixels are treated as noise or uncertainty in class allocation of a pixel. Conventional hard classification algorithms may thus produce inaccurate classification outputs. The application of soft or sub-pixel classification methods may be adopted for classification of images acquired in complex and uncertain environment. The main objective of this research work have been to study the effect of feature dimensionality as well as the effect of training sample size on three different types of sub-pixel classification algorithms, like; statistical classifier-linear mixture model (LMM), learning classifier- artificial neural network (ANN) and statistical learning classifier- support vector machine (SVM), using Hyperion EO1 data. In this work mixed pixels have been used at allocation and testing stages and sub-pixel classification outputs have been evaluated using fuzzy error matrix (FERM). FERM method for evaluation of sub-pixel evaluation was not available in commercial software, so, in-house SMIC-Sub-pixel Multispectral Image Classifier package has been used. Earth Observation satellite, Hyperion EO1 data, acquisition date, April, 2005 and IRS-P6, LISS-III data acquisition date March, 2005 have been used for classification output and testing the classification output respectively. Thus the finding of this research illustrate that upcoming statistical learning classifier, SVM, can produce thematic information with higher sub-pixel classification accuracies in comparison to conventional classifiers, from higher spectral dimensional, hyperspectral remotely sensed data fraught with mixed pixels.

## 1. INTRODUCTION

The most widely used method for extracting information from remotely sensed data is image classification. Research into the problem of land cover classification using multispectral remotely sensed data has been ongoing since the early 1970s, when ERTS (later Landsat - 1) multi-spectral scanner (MSS) data became available. Classification techniques such as the parallelopiped, minimum distance to mean and maximum likelihood methods were developed. Perhaps surprisingly, the accuracy of land cover classification was not necessarily improved (and, in some cases, was reduced) by the use of higher spatial resolution data and by the availability of additional bands in the near and mid- infrared wavebands (Cushnie 1987). Woodcock and Strahler (1987) suggested that this phenomenon is the consequence of an increase in within class spectral variability as spatial resolution increases (i.e. pixel size decreases). Hughes (1968), have reported that classification accuracy decreased, as additional features were included. This effect has been termed "the curse of dimensionality".

In 1990's onwards hyperspectral data was available to remote sensing application users. So, till 1990's the available classification techniques were not sufficiently powerful to identify patterns in hyperspectral data. In this study one statistical based classifier; Linear Mixture Model (LMM), learning based classifier; Artificial Neural networks (ANN) and statistical learning based classifier; Support Vector Machine (SVM) have been evaluated.

There are two main types of classification method, namely supervised and unsupervised approach. Both methods may be applied to perform hard and soft classification. In hard classification, pixel is allocated to one and only one class, which may produce erroneous results, particularly in classifying coarse spatial resolution images. It is important that soft classification is used to produce class proportions within a pixel in order to increase the classification accuracy and to produce meaningful and appropriate land cover composition. The main objective is to study the performance of three classification methods (LMM, ANN and SVM) while applying soft approach at allocation as well as at testing stage. So the objectives of this research work are stated as follows;

1. To study the effect of number of features on sub-pixel classification accuracy.
2. To study the effect of training data size on sub-pixel classification accuracy.
3. To compare the sub-pixel classification algorithms (i.e. LMM, ANN, SVM) while using optimum parameters.

## 2. LITERATURE REVIEW

Mahesh and Mather 2006 have studied the sub-pixel classification of hyperspectral DAIS and Landsat-7 ETM+ data applying four algorithms (such as, maximum likelihood, decision tree, artificial neural network and support vector machine) and made the comparison of accuracy assessment, for the area of La Mancha Alta, South of Madrid, Spain. They conclude that no reduction in classification accuracy was observed as the number of bands was increased, even with the small training data set of 100 pixels per class. However, classification accuracy starts to stabilize once a threshold number of bands are reached. The SVM produce higher classification accuracies than others with small training data sets. The effect of using different sampling plans was investigated and it was found that ML classifier produces higher classification accuracies when the training data were sampled randomly than those achieved using a systematic sampling plan. Both sampling plans produced similar results with SVM, DT and ANN. The level of classification accuracy when 13 MNF components were used were lower than those obtained by classifying the raw data, indicating that the MNF technique may not be effective for dimensionality reduction in the context of classification with this type of data. The use of DT based feature selection techniques and the accuracy achieved was close to the level reached using raw data, suggesting that the DT approach can be effectively used for feature selection with hyperspectral data. The ML classifier shows a greater dependence on the characteristics of the training data than do the other methods. This result indicates that the ML method does not generalize well to unknown cases. The SVM algorithm is least affected by the nature of the training data.

Aziz, M.A. 2004 has evaluated the soft classifiers for multi-spectral remote sensing data, and this study has focused on two statistical classifiers; maximum likelihood classifier (MLC) and linear mixture model (LMM), two fuzzy set theory based classifiers; fuzzy c mean (FCM) and probability c mean (PCM) and two neural network classifiers; back propagation neural network (BPNN) and competitive learning neural network followed by learning vector quantizers (CLNN-LVQ). IRS 1B LISS-2 data has been used for classified and IRS 1C PAN image derived reference map registered to LISS-2 has been used for testing image. The hypothesis of fuzzy error matrix (FERM) has been promoted to assess the accuracy of soft classification. As the formulation of majority of these classifiers and accuracy measures in the existing commercial image processing software are not available, so Soft Classification Methods and Accuracy assessment Package (SCMAP) has also been developed. The results showed that the distribution free classifiers based on fuzzy set and neural network produced more accurate classification than the statistical classifiers. An improvement in accuracy of 8% to 12% was observed. It was shown that how PCM classifier was robust to the existence of noise in the data. CLNN-LVQ produced the highest classification accuracy of 53.89% and showed an improvement of more than 5% over the FCM. The accuracy of hard classification was further increased by including a priori probabilities in BPNN and MLC classifiers. A new approach to include a priori probabilities by way of replicating the training data of a class in accordance with the proportional area covered by that class on ground was suggested. The accuracy of BPNN classifier increased by 20% whereas the accuracy of MLC increased by 7% on the inclusion of a priori probabilities. Evaluation of soft classification through FERM based measures led to an improvement of the order of 20% in the accuracy of the classification over the accuracy determined from traditional error matrix based measures for the same classification. thus, it is recommended that soft classification outputs from any classifier should not be hardened for evaluation purposes, as this may results into loss of information. LMM as soft classifier produced the lowest accuracy whereas BPNN and PCM as soft classifiers produced the highest map accuracy of about 73%, which was an improvement of 20% over the highest accuracy achieved by the unsupervised classifiers. When the images are dominated by mixed pixels, their incorporation not just in allocation stage through generation of soft outputs, but also in training and test stages were also assessed. The results showed that by properly accounting for mixed pixels in all stages, same level of accuracy could be achieved as would have been obtained by using pure pixels in all stages.

Mahesh and Mather 2003 have done support vector classifiers for land cover classification. They have studied for two project areas; the first area used in the report is near town of Littleport in eastern England. The second is a wetland area of La Mancha region of Spain. For the Littleport area, $ETM^+$ data acquired on $19^{th}$ June 2000 is used. The classification problem involves the identification of seven land cover types (wheat, potato, sugar beet, onion, peas, lettuce and beans) for the $ETM^+$ data set. For the La Mancha study area, hyperspectral data acquired on $29^{th}$ June 2000 by the DAIS 7915 airborne imaging spectrometer were available. Eight different land cover types (wheat, water body, dry salt lake, hydrophytic vegetation, vineyards, bare soil, pasture lands and buildup area) were specified. Random sampling was used to collect the training and test for both data sets. Total selected pixels were divided into two part, one for training and one for testing the classifiers, so as to remove any possible bias resulting from the use of same set of pixels for both testing and training phases. A standard back propagation neural network classifier was used. All user defined parameters are set as recommended by Kavzoglu (2003). Like ANN classifier the performance of SVM depends on a number of user defined parameters, which may influence the final classification accuracy. For the study, a radial basic kernel with penalty value C=5000 is used for both data sets. The values parameters were chosen after a number of trials and the same parameters are used with the DAIS data. Results obtained using $ETM^+$ data suggests that the SVM classifier perform well in comparision with ANN and MLC. Further the training time taken by SVM is 0.3 minutes in comparison of 58 minutes by the ANN on a dual processor machine. Results suggest that SVM performance is statistically significant in comparison with ANN and MLC classifiers. To study the behavior of SVM classifier with hyperspectral data a total of 65 bands are used as the combination of first 5 bands, first 10 bands, etc. giving a total of 13 experiments. Results obtained from analysis of hyperspectral data suggested that SVM classifier increase almost continuously as a function of number of features, with the size of training data set held constant, whereas the overall classification accuracies produced by the ML, DT and ANN classifiers decline slightly once the number of bands exceeds 50 or so. They concluded that SVM outperforms MLC and ANN in terms of classification accuracy with both data sets. Several user-defined parameters affect the performance of SVM classifier, but it is easy to find appropriate values for these parameters than it is for parameters defining the ANN classifier. The level of classification accuracy achieved by SVM classifier is better than both MLC and ANN classifiers when used with small number of training data.

The review of literature suggests that there is a range of soft classification methods proposed and implemented by different researchers. From among a number of soft classification methods, this paper has focused on statistical method (Linear mixture Model), learning method (Artificial Neural network), and statistical learning method (Support Vector Machine). The details of these algorithms have been given in following sections. Fraction images generated from LMM, ANN and SVMs methods have been evaluated using FERM. This is a new approach that has been developed to assess the accuracy of soft classifiers (Binaghi *et al*., 1999). The elements of the fuzzy error matrix represent class proportions, corresponding to soft reference data ($R_n$) and soft classified data ($C_m$), to class n and m, respectively.

## 3. TEST SITE AND DATA USED

The test site for this research work was near Nepali farm, south west of Rishikesh administratively belonging to Dehradun district of the Uttarakhand State. The district lies between the parallels of 30°04'14. 04" and 30°00'52.92" latitude and 78°13'12. 44" and 78°17'01.58" longitude. In this project, the hyperspectral data of Earth Observation Satellite EO1, Hyperion data has been used for classification and Indian Remote Sensing Satellite (IRS-P6), LiSS-3 data has been used for testing. Another dataset IRS-P6, AWiFS and LISS-3 acquired at same time has been used to test and prove the earlier results (results from Hyperion EO1 and LISS-3 data) for the project work. Survey of India, toposheet, 53 J/4, 53 J/8, 53 K/1 and 53 K/5 at 1:50,000 scale and GPS field data were used for image registration, training and reference data collection.

## 4. OPTIMUM PARAMETERS OF CLASSIFICATION ALGORITHMS AND OUTPUTS

While using different classification algorithms like; one statistical algorithm (Linear Mixture Model), one learning algorithm (Artificial Neural Network), and one statistical learning algorithm (Support Vector Machines) in this research work, the optimum learning parameters (Aziz, M. A. 2004, unpublished Ph. D thesis, IIT Roorkee) for ANN, are shown in Table 1 and logistic function as activation function has been used. For SVMs, four kernels were studied with, the optimum penalty value *C* has been used from Varshney and Arora, 2004. The optimum penalty values adopted for different kernel types are shown in Table 2.

| | | |
|---|---|---|
| 1 | Training Rate | 0.2 |
| 2 | Training momentum | 0.5 |
| 3 | RMS | 0.01 |
| 4 | Number of Hidden Layers | 1 |
| 5 | Number of Training Iterations | 100 |

Table 1. Optimum parameters for ANN

| Sr. No | Kernel Type | Penalty value (*C*) | |
|---|---|---|---|
| | | Hyperspectral | Multispectral |
| 1 | Linear | 50 | 100 |
| 2 | Polynomial Function with | 500 | 500 |
| 3 | Radial Basic Function | 5000 | 7500 |
| 4 | Sigmoid | 7500 | 0.75 |

Table 2. Optimum penalty values for different kernel types

For the accuracy assessment of each classified images generated from different algorithms and with different training sample sizes have been computed by using fuzzy error matrix (FERM) approach in Sub-pixel Multispectral Image Classifier (SMIC) software. The classified images and reference images had been taken as Hyperion EO1 rule images and LISS-3 rule images respectively. For accuracy assessment of the rule images, 500 (100 samples per class; total five classes) samples, outside the training area, have been collected from fraction images generated from LISS-3 image of IRS-P6 satellite, dated March 2005, as reference image as well as from fraction images generated from Hyperion EO1 image, dated April 2005. These testing sample sizes have been equivalent to the sample size of 75 to 100 pixels per class as recommended by Congalton (1991) for accuracy assessment purpose.

## 5. VERIFICATION OF CLASSIFICATION RESULTS

It was difficult to get the Hyperion EO1 data and reference data of LISS-3 data of same time. So to check how was the performance of these classifiers with classified (AWiFS data) and reference (LISS-3 data) data set acquired at same time had been used from IRS P6 satellite. This test site was located west of Dehradun city, of the Uttarakhand State. The area of interest (AOI), number of land cover classes, number of training samples, classification algorithms with their optimum parameters and all other processing procedures were set with the same as in previous dataset.

## 6. RESULTS AND DISCUSSIONS

In this research work influence of feature dimensionality on classification accuracy is assessed using the Hyperion EO1 data. Also it has been accessed the relative performance (in terms of classification accuracy) of the three classification algorithms and to check how classification accuracy varies with a fixed number of training data as the number of bands is progressively increased from 30 to 150. Five subsets of Hyperion EO1 bands had been extracted, comprising bands, 1-30, 1-60, 1-90, 1-120 and 1-150, respectively. The classifications were performed using the LMM, ANN, SVMs (used four kernel types) with five training data set sizes of 100, 200, 300, 400 and 500 sample pixels per class randomly. The effect of dimensionality on classification accuracy achieved from three classifiers (LMM, ANN, SVM) at fixed training data size of 300 is shown in figure 1.

Classification accuracy using the SVM algorithm (with sigmoid kernel) has been 71.184% when first thirty bands (1-30 bands of Hyperion EO1 data) were used and 76.294% when all bands (1-150 bands of Hyperion EO1 data) were used. The maximum classification accuracy of LMM, 85.318% occurred when the bands combination of 1-150 bands dataset of Hyperion EO1 data with 400 samples per class was used. For ANN classifier, the highest classification accuracy, 68.033% was in the band combination of 1-90 bands data with 500 pixels per class. The maximum accuracy of SVM classifier with sigmoid kernel, 90.846% had been produced by 1-120 bands data set of 500 training samples per classes. Of the three classifiers used, the SVM with sigmoid kernel gave the highest classification accuracy when 1-120 bands data had been used while ANN method generally gave the lowest classification accuracy value over all training sample sizes. But the difference between the lowest and highest accuracy was not too much different. Generally, the classification accuracy of SVM with sigmoid kernel and LMM had been increased when the numbers of bands were increased. For ANN, the maximum classification accuracy had been occurred when the number of bands reached at 1-90 bands and after that it decreases when the number of bands where further increased. But classification accuracy of SVM with other kernel types, such as linear kernel, polynomial function with degree 2 and radial basic function has variations. Based on the results following objectives where achieved.

1. The SVM with sigmoid kernel produced the higher classification accuracies compared to ANN and LMM even with small number of training sample data set, suggesting its appropriateness in situations where the training sample data are difficult to collect.
2. Classification accuracy had been increased in case of SVMs with increased in number of training samples as well as with increased in number of bands. Out of four kernels used in SVM classifier, sigmoid kernel gave the higher classification accuracy. Other classification algorithms were not consistent.
3. ANN gave the lower classification accuracies through out all the bands combination and different number of training samples data sets.
4. The highest accuracy was not over 90.846% in this analysis by using different acquired date of satellite data. In this project the acquire date was nearly one month different, (i.e. test image LISS-3 was one month before the classified image of Hyperion EO1 data. The time different was not too much, but the crop season was totally different, LISS-3 data is in nearly harvested season but the Hyperion EO1 data was after harvested season.
5. The further investigation for classification accuracy of three classifiers used in this work was done using AWIFS data as input and LISS-3 data as test image with same data acquisition dates, to verify the performance of three classification algorithms used. It had proven that the classification algorithms used in this study generally have the same trends as in hyperspectral dataset (Hyperion EO1 and LISS-3 data).

## 7. CONCLUSION AND FUTURE SCOPE

The objective of this project was to understand the behaviour of classification algorithms on effect of dimensionality of feature space and the effect of training sample size. In this research work, while using three classification algorithms, allocation of land cover classes, testing of classified data had been done with fuzzy error matrix (FERM) approach. The conclusions of this research work are as follows;

1. When band was set to minimum, 30 bands and it was increased up to 150 bands, SVM (Sigmoid) gave the maximum classification accuracy in all cases. It does not follow Hughes (1968) law.
2. When training data was set to 100 samples and changed up to 500 samples/class, SVM (Sigmoid) gave the maximum classification accuracy in all cases.

## 8. FUTURE SCOPE

Hyperspectral remote sensing holds the potential to provide a high spectral data obtained about the earth surface features. By using hyperspectral data, a sub-pixel classification method can give the proper high classification accuracy and more accurate results to identify the minor features, (within classes). Classification accuracy depends on a number of factors, of which the nature of training samples, the number of bands used, the number of classes to be identified relative to the spatial resolution of the image and the properties of the classification algorithms are the most important. This research evaluates the effect of these factors on classification accuracy using test data of Hyperion E01 data and IRS P6, LISS-3 data. There is no evidence to support the view that classification accuracy inevitably declines as the data dimensionality and the number of training samples increase. The use of another verification data set, proved that these effects were also appropriate to encourage this research.

Classification accuracy also depends on the optimum parameters of the classification algorithms; different types of data and different algorithms have different parameters. So the learning of these optimum parameters should be carried out to train the algorithms and hence by using these optimum parameters, the exact accurate results can be calculated. The acquisition date of allocation and test images should be the same. Detail study of mineral detection, vegetation and forest species level identification, etc. which could be carried out with much higher spectral and spatial resolution.

## REFERENCES

Aziz, M. A. 2004. Evaluation of soft classifiers for remote sensing data, unpublished Ph. D thesis, Indian Institute of Technology Roorkee, Roorkee, India.

Congalton R. G., A review of assessing the accuracy of classification of remotely sensed data. Remote Sensing Environment, 37, 35-47 (1991). [doi:10.1016/0034-4257(91)90048-B

Hughes G. F, 1968. On the Mean Accuracy of Statistical Pattern Recognizers., IEEE transactions on Information Theory, 14, pp.55-63

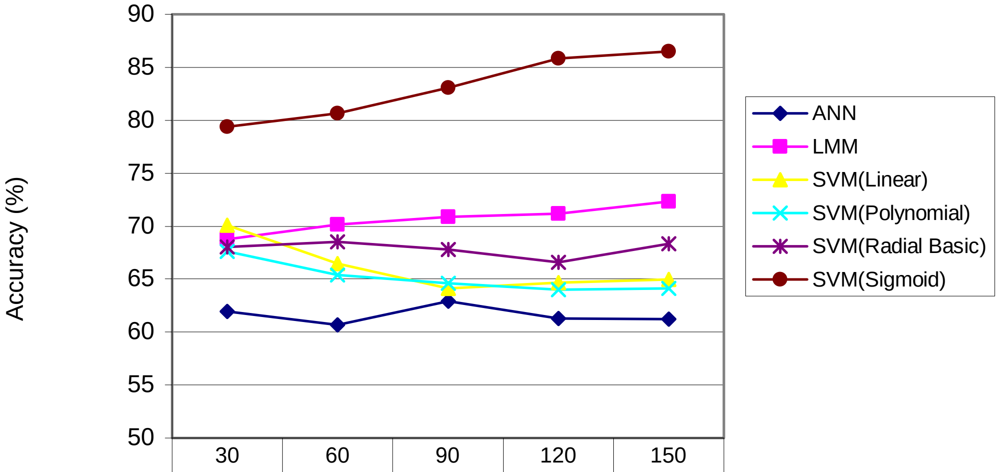
Kavzoglu, T. and Mather, P. M. 2003: The use of backpropagating artificial neural networks in land cover classification., IJRS, 24, pp. 4907-4939.

Mahesh and Mather, Support vector classification for land cover classification, Processing's of Map India 2003 Image processing and Interpretation theme.

Mahesh and Mather, 2006; Some issues in the classification of DAIS hyperspectral data, IJRS, 27 (14). pp. 2895-2916.

Varshney P. K. and Arora M. K., Advance Image Processing Techniques for Remotely Sensed Hyperspectral Data, Springer-Verlag, Berlin (2004).

Woodcock, C. E. and Strahler, A. H. 1987: The factor of scale in remote sensing., Remote Sensing of Environment, 21, pp. 311-332.




| | 30 | 60 | 90 | 120 | 150 |
|---|---|---|---|---|---|
| ANN | 61.921 | 60.695 | 62.928 | 61.275 | 61.226 |
| LMM | 68.742 | 70.167 | 70.866 | 71.184 | 72.318 |
| SVM(Linear) | 70.105 | 66.483 | 64.096 | 64.685 | 64.989 |
| SVM(Polynomial) | 67.645 | 65.412 | 64.579 | 64.005 | 64.115 |
| SVM(Radial Basic) | 68.053 | 68.500 | 67.800 | 66.600 | 68.356 |
| SVM(Sigmoid) | 79.400 | 80.639 | 83.069 | 85.825 | 86.530 |

Number of Bands

Figure 1. Variation in classification accuracy with increasing numbers of bands for training sample sizes of 300 pixels per class