

OPEN ACCESS

Hyperspectral image classification using Support Vector Machine

To cite this article: T A Moughal 2013 *J. Phys.: Conf. Ser.* **439** 012042

View the [article online](#) for updates and enhancements.

You may also like

- [Theoretical investigation on the excited state intramolecular proton transfer in Me₂N substituted flavonoid by the time-dependent density functional theory method](#)
Hang Yin, , Ying Shi et al.
- [All-fiber transparent piezoelectric harvester with a cooperatively enhanced structure](#)
Yiin-Kuen Fuh, Hsi-Chun Ho, Bo-Sheng Wang et al.
- [Microstructured hybrid nanocomposite flexible piezoresistive sensor and its sensitivity analysis by mechanical finite-element simulation](#)
Yin He, Liduan Zhao, Xiujuan Wang et al.



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Hyperspectral image classification using Support Vector Machine

T A Mughal

Laboratory of Complex Systems and Intelligent Control, School of Mathematical Sciences, Beijing Normal University, Beijing, China

E-mail: tauqir_mughal@yahoo.com

Abstract. Classification of land cover hyperspectral images is a very challenging task due to the unfavourable ratio between the number of spectral bands and the number of training samples. The focus in many applications is to investigate an effective classifier in terms of accuracy. The conventional multiclass classifiers have the ability to map the class of interest but the considerable efforts and large training sets are required to fully describe the classes spectrally. Support Vector Machine (SVM) is suggested in this paper to deal with the multiclass problem of hyperspectral imagery. The attraction to this method is that it locates the optimal hyper plane between the class of interest and the rest of the classes to separate them in a new high-dimensional feature space by taking into account only the training samples that lie on the edge of the class distributions known as support vectors and the use of the kernel functions made the classifier more flexible by making it robust against the outliers. A comparative study has undertaken to find an effective classifier by comparing Support Vector Machine (SVM) to the other two well known classifiers i.e. Maximum likelihood (ML) and Spectral Angle Mapper (SAM). At first, the Minimum Noise Fraction (MNF) was applied to extract the best possible features from the hyperspectral imagery and then the resulting subset of the features was applied to the classifiers. Experimental results illustrate that the integration of MNF and SVM technique significantly reduced the classification complexity and improves the classification accuracy.

1. Introduction

Land cover is an elementary variable that impacts on and links many parts of the human and physical environments [1]. Thus, information on the spatial distribution of the land cover classes is of vital importance for the investigation of environmental processes. Satellite remote sensing techniques are widely used for the environmental monitoring. Hyperspectral imagery is a valuable source from which one can extract detailed information about earth surface phenomena and objects. In fact, the sensors are characterized by a very high spectral resolution that usually results in hundreds of narrow spectral channels. Remote sensing images acquired by hyperspectral sensors, such as the widely used Airborne Visible Infrared Imaging Spectrometer (AVIRIS) and Hyperspectral Digital Imagery Collection Experiment (HYDICE) sensors, have shown their usefulness in numerous earth observation applications [2],[3]. HYDICE sensor data is used in this study for land cover classification. Hyperspectral Digital Imagery Collection Experiment (HYDICE) sensor is an experimental earth imaging instrument with very high spectral resolution. This sensor provides the sample spectrum in 210 spectral channels, nominally 10 nanometers wide covering the spectral range from 400 to 2500 nanometers.



The analysis of the hyperspectral data is not a trivial task. In particular, many factors made the analysis quite complex, such as the large spatial variability of the hyperspectral signature of each land cover class, atmospheric effects; and the curse of dimensionality [1]. Therefore, the main objective of this study is to search for an effective classifier by extracting the best possible features appropriate for the land cover classification. Three classifiers have been used for the comparison purpose.

Maximum Likelihood (ML) is a very popular parametric classifier, being widely used in pattern recognition and image classification [4]. It usually acquires higher classification accuracy compared to other traditional classification approaches. It assumes that each band is normally distributed and the chosen training samples are comprised of exhaustively defined set of classes. For hyperspectral data with tens of hundreds of spectral bands, the efficient training pixels (exhaustively defined) to locate for the discrimination of land cover classes is not an easy task. Whereas, the classification accuracy of ML classifier is based on the accurate selection of the training samples. Thus, for the hyperspectral imagery with poorly represented labelled training samples, it is preferable to adopt an alternative to the standard multiclass classifier. One approach is to adopt a Spectral Angle Mapper (SAM) technique, which compares each pixel in the image with every endmember for each class instead of comparing it to the training pixels [5]. The constraint to this classifier is, while selecting the endmembers; it does not take into account the sub-pixel values and becomes more problematic for the heterogeneous earth's surface. Further improvement in the classification accuracy can be attained by the development of Support Vector Machines (SVMs), which makes a breakthrough of classification for hyperspectral data. SVM do not require an estimation of the statistical distribution of classes to carry out the classification task, whereas it only defines the classification model by exploiting the concept of margin maximization by taking into account only few training pixels. SVM is an effective method of statistical learning theory, compared with the traditional classification methods; it is suitable for small samples learning, besides, it has better generalization ability and high efficiency for learning [6]. By comparing the three classifiers i.e. Maximum Likelihood (ML), Spectral Angle Mapper (SAM) and Support Vector Machine (SVM), the efficient monitoring of land cover classification can be achieved.

The rest of this paper is organized in four sections. Section II describes the data description and area under study. Section III describes different classifiers that can be used for the analysis purpose. Section IV deals with the experimental phase of the work and summarizes the observations. Finally, Section V presents the concluding remarks and future directions to complete this paper.

2. Data description and study area

The hyperspectral data that will be used for the analysis in this study is taken by the airborne Hyperspectral Digital Imagery Collection Experiment (HYDICE) sensor available in the student CD-ROM [7]. It was collected for the Mall in Washington, DC, with 210 bands covering 0.4–2.4 μm spectral region. This image has very high spatial resolution of about 2.8 m. The HYDICE image size is 307×1280 pixels. The dataset contains 1280 scan lines with 307 pixels in each scan line. There are seven information classes in the Washington, DC data includes road, grass, shadow, trails, trees, roof and water. Training and test samples are available for this scene.

Table 1. HYDICE sensor data information classes and samples.

Classes	Training Samples	Testing Samples
Trails	183	42
Shadow	223	48
Trees	136	192
Grass	970	300
Water	1,172	83
Roof	1,263	147
Roads	651	197



Figure 1. HYDICE sensor data.

3. Methodology

The methodology adopted in this study consists of two steps. First, to extract the best possible features appropriate for the classification of land cover classes and second, to find an effective classifier in terms of accuracy. Three classification methods have been proposed for comparative analysis in this study i.e. Maximum Likelihood, Spectral Angle Mapper and Support Vector Machine.

3.1. Data pre-processing

To get rid of the curse of dimensionality by selecting the appropriate features for the classification of HYDICE sensor data, Minimum Noise Fraction (MNF) serves the solution in this study. At first, the HYDICE sensor data contains 191 hyperspectral bands. MNF was adopted to reduce the dimensionality and computational requirements for the further processing of the HYDICE sensor data for the land cover classification.

The training samples were already available for the HYDICE sensor data as shown in table 1. There were seven classes of interest i.e. Trails, Shadow, Trees, Grass, Water, Roof and Roads that needed to be classified. The training samples were used to train ML and SVM classifiers to categorize the seven land cover classes from HYDICE sensor data for the comparative analysis.

3.2. Classification Techniques

Three classification techniques are adopted in this study to classify the land cover classes from the HYDICE sensor data.

3.2.1. Maximum Likelihood (ML). Maximum Likelihood (ML) is the parametric classifier based on the assumptions of normally distributed data for each class and exhaustively selected set of classes. For classifying the land cover classes, number of studies used the ML classifier as a benchmark to compare its classification accuracy with the other newly developed classifiers [8]. It is considered as a standard approach to thematic mapping from the remotely sensed imagery. In real life applications hardly the nature of the distribution is known. It is preferable to use the non-parametric classifiers that are free from assumptions. For the sake of comparison of the classification accuracies and to validate the suitability of the non-parametric classifier for classifying the land cover classes in this study, ML classification is conducted.

3.2.2. Spectral Angle Mapper (SAM). Spectral Angle Mapper (SAM) is a supervised classification algorithm, which utilizes spectral angular information for classification of hyperspectral image data [9]. It permits the rapid classification by calculating the spectral similarity between the image spectrums to reference reflectance spectra [10]-[12]. The reference spectra can either be attained from the field measurements or taken directly from the image. The reference spectra for this study is taken from image directly. SAM measures the spectral similarity by calculating the angle between the image and reference spectra, treating them as vectors in the n-dimensional feature space. The smaller angles between the two spectrums indicate high similarity and viceversa. The solar illumination factors does not affect this classifier. Moreover, It is very powerful classifier as it contains the influence of the shading effects to highlight the target reflectance characteristics. The major drawback faced by this classifier is that it assumes the endmembers chosen to classify an image by representing the pure spectra of a reference material, whereas the earth's surface is heterogenous in many ways and consists of mixed pixels. This classifier is adopted in this study because in general, the spectral mixture problem decreases with higher spatial resolution images like HYDICE.

3.2.3. Support Vector Machine (SVM). The Support Vector Machine (SVM) is a classification method based on the statistical information of remote sensing images [13]. Recently, particular attention has been devoted to support vector machines (SVMs) for the classification of hyperspectral remote sensing images [14]-[15]. SVM is a non-parametric binary classifier that locates the optimal hyper plane between the two classes to separate them in a new high-dimensional feature space by taking into account only the training samples that lie on the edge of the class distributions known as support vectors. Moreover, it does not require the assumption of normality and is insensitive to the curse of dimensionality. SVMs have often been found to provide higher classification accuracies than other widely used pattern recognition techniques, such as the maximum likelihood. Furthermore, SVMs appear to be especially advantageous in the presence of heterogeneous classes for which only few training samples are available. The SVMs were originally developed to solve binary classification problems. The implementation of SVMs in multiclass classification problem is possible by formulating SVMs directly as a multiclass optimization problem, but as the number of the classes needed to be classified increases, the number of parameters to be estimated increases, it in return affects the SVMs classification performance in terms of accuracy. All the three classifiers were used to classify an image to investigate for an effective classifier.

3.3. Accuracy Assessment

The confusion matrix was used to assess the accuracy measures for all the three classification procedures by using the available ground truth pixels. Also the overall accuracy (OA), Kappa coefficient (K), producer's accuracy (PA) and the user's accuracy (UA) was keenly observed. The overall accuracy is the percentage of all validation pixels correctly classified, whereas the user's and producer's accuracy provide information about the commission and omission errors associated with the individual classes, respectively. Unlike the overall accuracy, Kappa takes into account the possibility of agreements occurring by chance in a random classification [16].

The spectrum after removing the noise and extracting the suitable features acquired from the Minimum Noise fraction (MNF) were classified by the three above mentioned methods and their classification accuracies can be compared by analyzing the confusion matrices.

4. Results and Discussion

The output from three classification techniques has shown in figure 2 illustrating the seven land cover classes. By visual inspection of the classification results shown in figure 2 compared to the original image shown in figure 1, it can be seen that the ML classification significantly overestimated the Roof class and underestimated the Shadow class. Whereas, the SAM classification highly over estimated all the classes except the Water and Grass class. This is due to the reference spectra that had taken directly from the image. As the image was heterogeneous and reference spectra did not take into account the sub pixel information, therefore most of the classes are misclassified. However, the SVM classifier obtained relatively accurate classification results for all the land cover classes.

The results in terms of classification accuracy obtained by the three classifiers are summarized in table 2, table 3 and table 4. The SVM exhibited the best Overall Accuracy (OA) of 78.39%, i.e., the percentage of correctly classified pixels among all the test pixels considered, with a gain of 9.81% and 9.91% over the classical ML and SAM classifiers, respectively. Also the Kappa coefficient (K) is 0.7335 for the SVM, which is higher as compared to the ML (0.6074) and SAM (0.6163) classifiers. Moreover, for the comparative analysis of three classifiers, figure 3 and figure 4 shows a bar chart for the Producer's accuracy (PA) and User's accuracy (UA) in percentage of ML, SAM and SVM.

The following discussion concerns the quantitative and comparative analysis of three different classification techniques for hyperspectral image, including parametric and non-parametric approaches.

The accuracy of ML classifier was calculated as it provides the benchmark for the assessment of SVM classifier. The most critical class for the analysis in table 2 is the shadow class. The classification result of ML for shadow class shows the omission error of 93.75% higher as compared to the SAM (41.67%) and SVM (37.50%); hence ML is omitting the shadow pixels from the shadow class and under estimates this class. This can be better visualized by figure 3. It can be further observed from table 2 that the commission error for the roof class is also higher i.e. 58.62% in comparison to the SAM (42.51%) and SVM (45.03), which means that roof class is largely over estimated. Hence the use of ML classifier for classifying the land cover classes is not recommended in this study.

In order to monitor the environmental changes appropriately, search for an effective classifier for the classification of land cover is of crucial interest. Thus, for the comparative analysis, the confusion matrix for the SAM classifier is calculated. Table 3 shows that most of the classes are over estimated by using SAM classifier except the Grass and Water class, which shows the user's accuracy of 86.79% and 97.18% respectively.

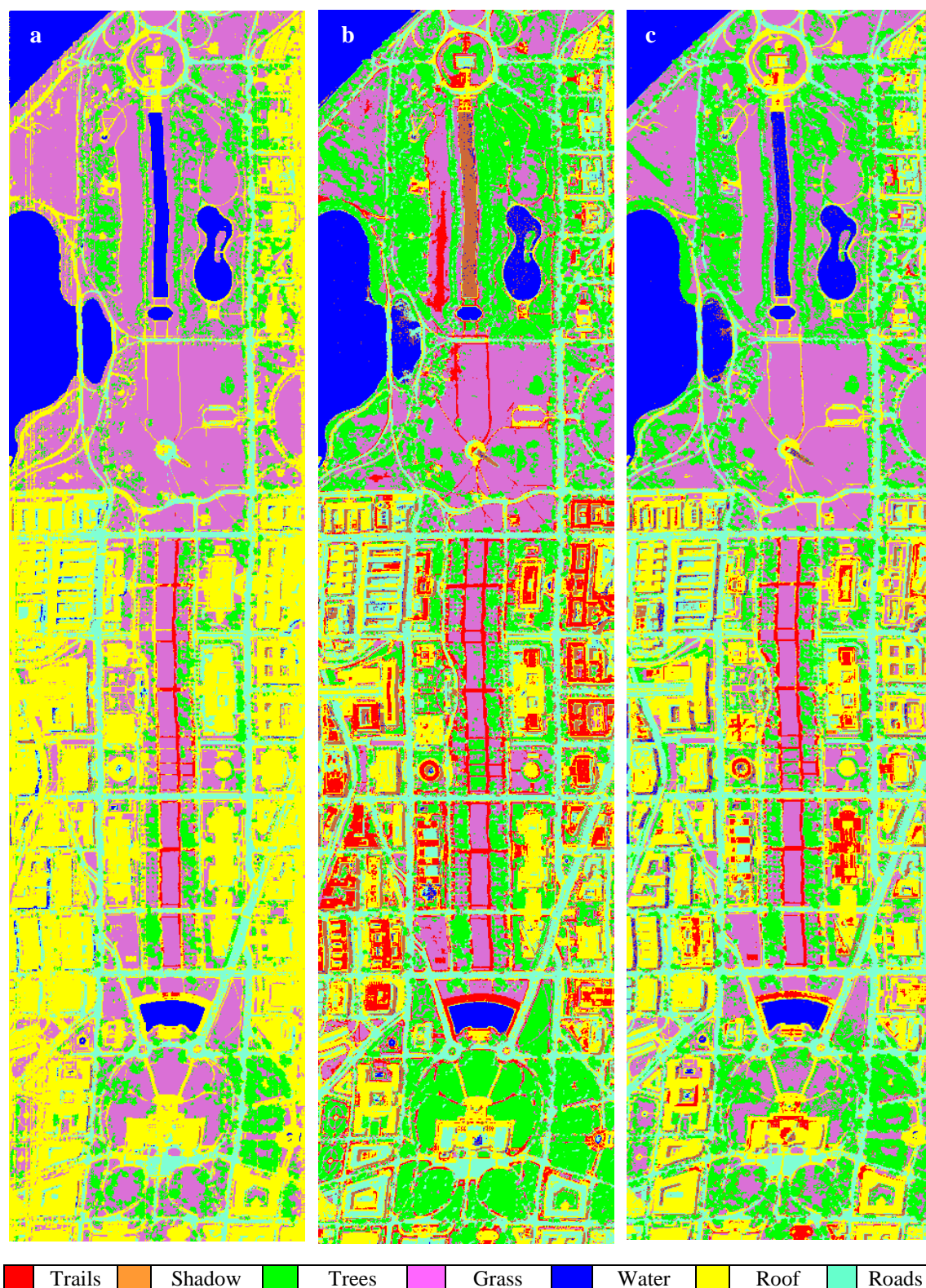


Figure 2. Classification results. (a) ML classification (b) SAM classification and (c) SVM classification.

Table 2. Confusion Matrix for Conventional ML.

	ML	Actual							PA (%)
		Trails	Shadow	Trees	Grass	Water	Roof	Roads	
Predicted	Trails	19	0	0	0	0	0	0	45.24
	Shadow	0	3	0	0	0	0	0	6.25
	Trees	0	1	94	5	0	0	5	48.96
	Grass	1	0	75	266	0	1	1	88.67
	Water	0	1	0	0	83	0	0	100
	Roof	22	22	23	29	0	144	108	97.96
	Roads	0	21	0	0	0	2	83	42.13
	Total	42	48	192	300	83	147	197	K = 0.6074
	UA (%)	100	100	89.52	77.33	98.81	41.38	78.30	OA=68.58%

Table 3. Confusion Matrix for SAM Classification.

	SAM	Actual							PA (%)
		Trails	Shadow	Trees	Grass	Water	Roof	Roads	
Predicted	Trails	39	0	0	23	0	25	7	92.86
	Shadow	0	28	0	0	11	1	3	58.33
	Trees	0	1	154	90	0	0	1	80.21
	Grass	0	0	22	184	0	3	3	61.33
	Water	0	2	0	0	69	0	0	83.13
	Roof	3	0	3	3	0	96	62	65.31
	Roads	0	17	13	0	3	22	121	61.42
	Total	42	48	192	300	83	147	197	K = 0.6163
	UA (%)	41.49	65.12	62.60	86.79	97.18	57.49	68.75	OA=68.48%

Table 4. Confusion Matrix for SVM Classification.

	SVM	Actual							PA (%)
		Trails	Shadow	Trees	Grass	Water	Roof	Roads	
Predicted	Trails	33	0	0	0	0	24	0	78.57
	Shadow	0	30	1	0	4	1	5	62.50
	Trees	0	1	158	25	0	0	1	82.29
	Grass	0	0	21	254	0	5	2	84.67
	Water	0	5	0	0	79	0	0	95.18
	Roof	9	0	0	20	0	105	57	71.43
	Roads	0	12	12	1	0	12	132	67.01
	Total	42	48	192	300	83	147	197	K = 0.7335
	UA (%)	57.89	73.17	85.41	90.07	94.05	54.97	78.11	OA=78.39%

The confusion matrix acquired for the SVM shows the best overall accuracy (78.39%) as well as the Kappa coefficient (0.7335) for the classification of land cover classes as shown in table 4. Moreover, the SVM shows the good tradeoff between the user's and producer's accuracies for all the land cover classes.

The results reported in table 4 confirm the superiority of the multiclass SVM in terms of both overall accuracy and Kappa coefficient. The good classification performance demonstrated by successful machine learning technique i.e. the support vector machine (SVM), using spectral signatures as input features, have improved by the incorporation of intelligent feature extraction technique (MNF), which reduces the dimensionality of the data to the right subspace without losing the original information that allows for the separation of the information classes.

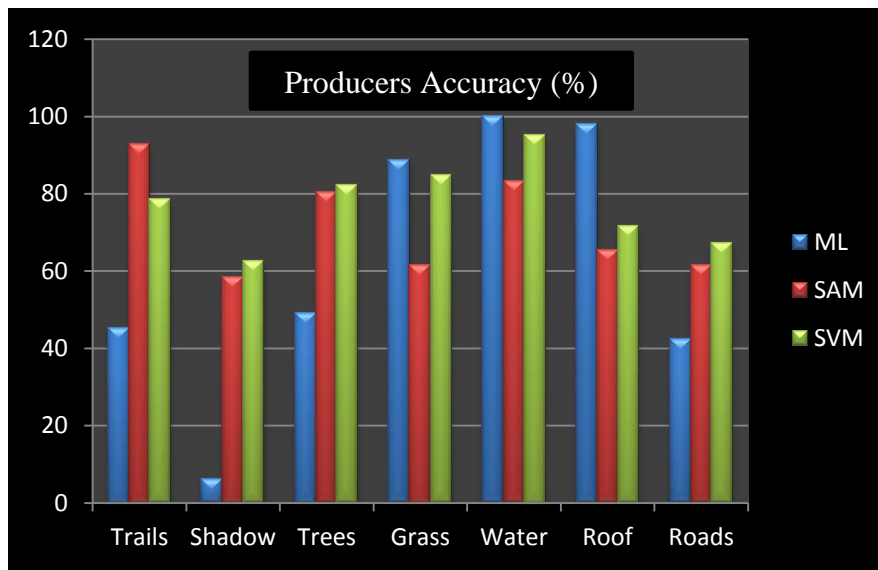


Figure 3. Producer's Accuracy (%) for Maximum likelihood, Spectral Angle Mapper and Support Vector Machine classifiers.

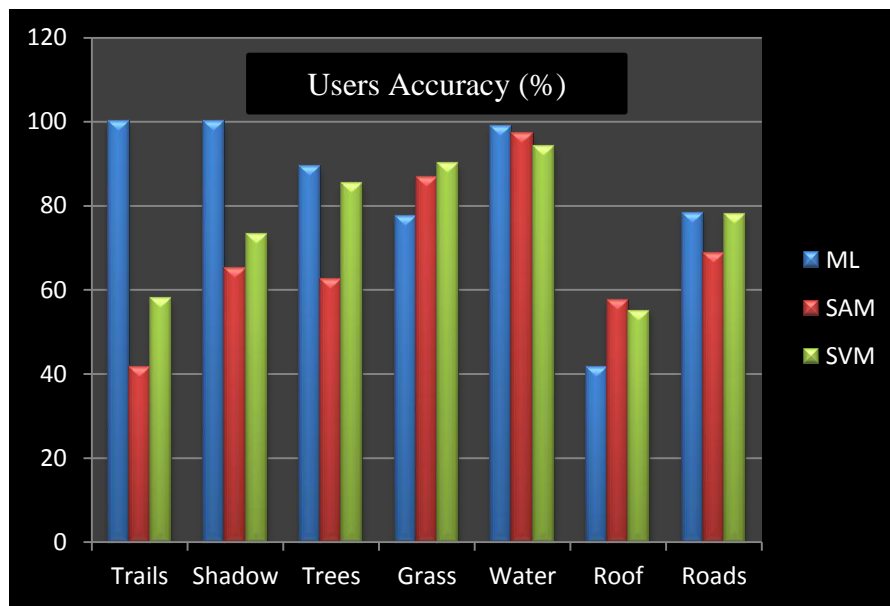


Figure 4. User's Accuracy for Maximum likelihood, Spectral Angle Mapper and Support Vector Machine classifiers.

Figure 3 and figure 4 shows the extreme classification behavior for some classes classified by the ML classifier. For example, the Shadow class shows the low producer's accuracy in comparison to a very high user's accuracy. Such a behavior of a classifier is not desirable as there is a need to trade off between the omission and the commission error. Moreover, SAM also shows the high producer's accuracy with a very low user's accuracy for the Trail class. Overall the SVM provides the better classification accuracy results for all the land cover classes.

4. Conclusion and Future Direction

In this study, the problem of the classification of hyperspectral remote sensing data using Maximum Likelihood (ML), Spectral Angle Mapper (SAM) and Support Vector Machine (SVM) have been addressed. In order to assess the effectiveness of the land cover classification methodologies, the main objective considered was to search for an effective classifier by extracting the best possible features using MNF, appropriate for the land cover classification. The results obtained from the HYDICE sensor dataset showed that SVM is much more effective than other conventional classifiers (i.e., the ML and the SAM classifier) in terms of classification accuracy, computational time, and stability to parameter settings.

Another important aspect to be pointed out is the intrinsic good generalization capability of SVM, which stems from the selection of the hyperplane that maximizes the geometrical margin between classes. In a hyperspectral context, the maximum margin solution allows to fully exploit the discrimination capability of the relatively few training samples available.

The major drawback to this solution is that the large spatial variability of the hyperspectral signature of each information class given the limited information present in the training set adversely affects the classification accuracy. However, it is worth noting that to solve the problem of the spatial variability of the hyperspectral signature of classes effectively, good generalization properties of the classifiers should be coupled with other techniques.

Further research will be done from the view of improving the classification accuracy and reducing the calculation time.

References

- [1] King R B 2002 *Land cover mapping principles: a return to interpretation fundamentals*. Int. J. Remote Sens. **23** 3525-3546
- [2] Zhang J, Zhang Y and Zhou T 2001 *Proc. International conference on* **1** 882-885
- [3] Benediktsson J A, Palmason J A and Sveinsson J R 2005 *IEEE Trans. Geosci. Remote Sens.* **43** 480-491
- [4] Jia X P and Richards J A 1994 *IEEE Trans. Geosci. Remote Sens.* **32** 274-281
- [5] Girouard G, Bannari A, Harti A E and Desrochers A 2004 Istanbul Turkey XXth ISPRS 599 – 604
- [6] Melgani F and Bruzzone L 2004 *IEEE Trans. Geosci. Remote Sens.* **42** 1778-1790
- [7] Landgrebe D A 2003 *Signal Theory Methods in Multispectral Remote Sensing* Hoboken, NJ: John Wiley & Sons
- [8] Huang C, Davis L S and Townshend J R G 2002 *Int. J. Remote Sens.* **23** 725-749
- [9] Kruse F A, Lefkoff A B, Boardman J W, Heidebrecht K B, Shapiro A T, Barloon P J and Goetz A F H 1993 *Remote Sens. Environ.* **44** 145-163
- [10] Crosta A P, Sabine C and Taranik J V 1998 *Remote Sens. Environ.* **65** 309-319
- [11] Schwarz J and Staenz K 2001 *Can. J. Remote Sens.* **27** 216-224
- [12] Hunter E L and Power C H 2002 *Int. J. Remote Sens.* **23** 2989-3008
- [13] Cortes C and Vapnik V 1995 *Support-Vector Networks, Machine Learning* Lorenza Saitta Netherlands Kluwer Academic Publisher, Boston **20** 273-297
- [14] Vapnik V N 1998 *Statistical Learning Theory*. New York: Wiley
- [15] Burges C J C 1998 *A tutorial on support vector machines for pattern recognition, Data Mining Knowledge Discovery* **2** 121-167
- [16] Brown D G, Lusch D P and Duda K A 1998 *Geomorphology* **21** 233 – 250