Hyperspectral Dimensionality Reduction via Localized Discriminant Bases

Shilpa Venkataraman, Lori Mann Bruce, *Senior Member IEEE*,
Anil Cheriyadat, Abhinav Mathur, *Student Member IEEE*

Electrical and Computer Engineering Department, Mississippi State University, Mississippi State, MS 39762

***Abstract* —To overcome the dimensionality curse of hyperspectral data, the authors of the paper have investigated the use of grouping the spectral bands along with localized discriminant bases, followed by decision fusion to develop an ATR system for data reduction and enhanced classification of hyperspectral data. The proposed system is robust to the availability of limited training data. Initially, the entire span of spectral bands in the hyperspectral data is subdivided into subspaces or groups based on a performance metric. The groups are not allowed to grow beyond what is supported by the amount of available training data. Feature extraction is done using supervised methods as well as unsupervised methods. Further, decision level fusion is applied to the features extracted from each group. To reduce the effect of conflicting decisions by individual groups, a voting scheme called Qualified Majority Voting is adopted to combine decisions. The effectiveness of the proposed system is tested using a data set consisting of hyperspectral signatures of a target class, Cogongrass (*Imperata Cylindrica)*, and a non-target class, Johnsongrass (*Sorghum halepense).* Cogongrass is an invasive species of plant whose monitoring has become important due to the extensive ecosystem damage that it causes. A comparison of target detection accuracies by the proposed system before and after decision fusion is done to illustrate the effect of the influence of each group of spectral bands on the final decision and to illustrate the benefit of using decision fusion with multiclassifiers.**

***Keywords*—hyperspectral, dimensionality reduction, feature extraction, multiclassifiers, data fusion, decision fusion, qualified majority voting**

I. INTRODUCTION

Modern hyperspectral sensors acquire data in hundreds of spectral bands for each pixel in a scene, thereby allowing subtly different classes to be discriminated. Since the number of spectral bands is very large in hyperspectral data, the number of training samples needed for accurate classification is immense, and often unrealistic to achieve. This problem is known as Hughes phenomenon [1] or the "curse of dimensionality." In traditional hyperspectral feature extraction methods, features are extracted from the data in the high dimensional space. However, the study in [2] on the characteristics of high dimensional spaces has shown that high dimensional data can be projected to a lower dimensional space without the loss of significant discrimination information. Features are extracted from lower dimensional subspaces formed by linear projections of the original high dimensional space. The extracted features become inputs to classifiers for labeling the data into target and non-target classes.

In the proposed method, the entire hyperspectral signature of a very high spectral resolution dataset is divided into adjacent disjoint groups based on a performance metric. This paper investigates several performance metrics that are functions of band correlation and/or class separation. The resulting groups are linearly independent, contiguous, and cover the entire spectrum of the data. Features are extracted from each of these groups using supervised and unsupervised statistical methods. Next, decision level fusion is applied to the features extracted from each group. To reduce the effect of conflicting decisions by individual groups, a voting scheme called Qualified Majority Voting (QMV) [3, 4] is adopted to combine decisions.

The proposed algorithm is tested on hyperspectral signatures obtained from the Analytical Spectral Device (ASD) [5] handheld spectroradiometer, consisting of 1625 spectral bands in the range of 350-1975nm. A comparison is done between traditional hyperspectral feature extraction methods, such as greedy search best band selection [6-7], and the proposed algorithm. The results show that this technique out-performs traditional methods in terms of classification accuracy as well as computation time.

II. METHODOLOGIES

Figure 1 shows a overall block diagram of the proposed methodologies.

A. Manual vs. Automated Grouping

Grouping algorithms can be used to sub-divide hyperspectral data into groups of spectral bands, from which features can be extracted to perform classification. Several methods of grouping can be applied to divide the entire hyperspectral subspace. The number of groups formed determines the reduced dimensionality. The authors have proposed and implemented a method by which the entire hyperspectral subspace can be projected into lower dimensional subspace in two different ways – manually grouping spectral bands into groups of fixed length, or grouping spectral bands based on a performance metric. Manual grouping is done by forcing a fixed number of bands into a single group. The entire hyperspectral signature consists

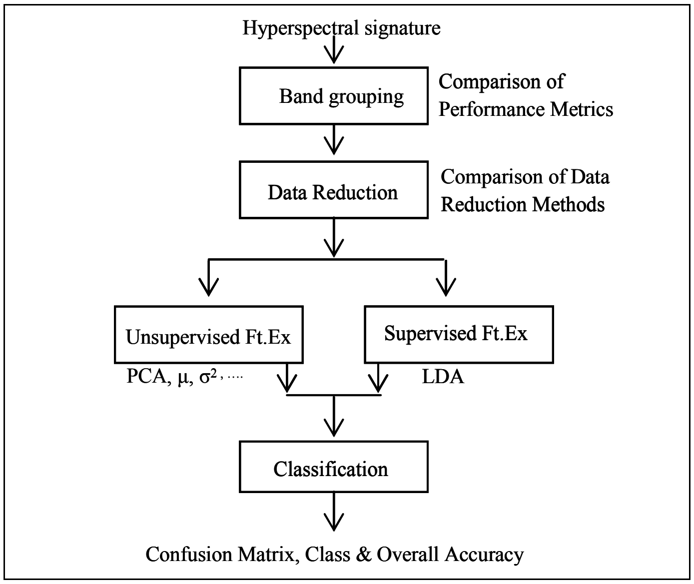

Figure 1. Overall block diagram of proposed methods.

of hundreds of spectral bands, which are manually divided into groups of size 10, 20, and so on. Grouping spectral bands based on a performance metric is an automated process where bands groupings are based on the maximization of a predefined performance metric. Here, the groups are not necessarily of equal lengths. Initially, adjacent pairs of spectral bands are grouped; the next step is to combine groups with adjacent bands or adjacent groups. This process is iterative until a stopping criterion is satisfied. In this study, two stopping criteria were used: *(i)* further grouping with adjacent groups or spectral bands does not significantly increment the current value of the performance metric, or *(ii)* the size of the groups is not allowed to grow beyond what is supported by the amount of available training data.

B. Performance Metrics

It is desirable that the adjacent bands of hyperspectral data exhibit high correlation and maximum separation between the classes [4]. Hence, the lower dimensional subspaces are formed based on a performance criterion which is a function of these two properties. In this study, Bhattaacharyya distance (BD), Jeffries Matusita distance (JM), and area under the receiver operating characteristics curve (ROC) are used as measures of class separation. Ten performance metrics were investigated: Correlation (Corr), BD, JM, ROC, product of Corr and BD (Corr*BD), product of Corr and JM (Corr*JM), product of Corr and ROC (Corr*ROC), weighted sum of Corr and BD (Corr+BD), weighted sum of Corr and JM (Corr+JM), and weighted sum of Corr and ROC (Corr+ROC). For each of the weighted sum metrics, weights ranged from 0 to 1 in step sizes of 0.1, and the sum of the two weights were forced to one.

C. Feature Extraction

Once the hyperspectral signature's bands have been appropriately grouped, features are extracted from each group. The authors analyzed both unsupervised and supervised methods of feature extraction. Unsupervised feature extraction methods included statistical measures, such and mean and variance, as well as eigen-based methods such as principal component analysis (PCA). Supervised feature extraction methods included Fisher's linear discriminant analysis (LDA).

D. Classification and Decision Fusion

Both parametric, such as maximum likelihood (ML), and nonparametric, such as k-nearest neighbor (k-NN), classifiers were investigated. Two types of classification analyses were conducted. In Analysis I, the features extracted from the band groups were forced into a single k-NN or ML classifier. That is, only one classifier utilized. In Analysis II, the features extracted from each group are applied to that group's classifier. That is, multiple classifiers are used. If there were *G* groups of spectral bands, *G* classifiers (k-NN or ML) were used. Each classifier resulted in its own classification decision, and the *G* decisions were fused using QMV. The QMV utilized a weighted vote from each of the *G* classifiers, where the weights were based on the accuracy of that particular group's classifier. If the group's classification accuracy (based on training data) was ≤75%, then the weight was zero (i.e. that group was not actually allowed to vote). If the group's classification accuracy was >75%, then the weight was proportional to the group's accuracy. These voting groups were referred to as the "discriminant bases". The final result, for both Analysis I and Analysis II, was a confusion matrix detailing the system's classifications of the testing data.

III. DATA COLLECTION

The signatures were obtained with an ASD Fieldspec Pro handheld spectroradiometer [5], which has a spectral range of 350 – 2500 nm, spectral resolution of 3 nm @ 700 nm and 10 nm @ 1400/2100 nm, and uses a single 512 element silicon photodiode array for sampling 350 - 1000 nm and two separate, graded index Indium-Gallium-Arsenide photodiodes for the 1000 - 2500 nm range [26]. The signatures were taken in good weather conditions in Mississippi, U.S.A., in 2000-2004 with the fiber optic sensor held NADIR at approximately shoulder height (4 feet) above ground. A 25° IFOV foreoptic was used, and the ASD unit was set to average ten signatures to produce each sample signature.

The data set consists of hyperspectral signatures of a target class, Cogongrass (*Imperata Cylindrica)*, and a non-target class, Johnsongrass (*Sorghum halepense).* Cogongrass is among the world's worst weeds, and it spreads at a rapid rate, displacing desirable native vegetation [ref]. Johnsongrass was chosen as the class to differentiate from Cogongrass in order to simulate realistic conditions, since many of the areas that are invaded by Cogongrass have substantial amounts of Johnsongrass present. In total, 286 and 130 measurements were collected for the Cogongrass and Johnsongrass classes, respectively. These signatures were jack-knifed into training and testing data sets for this study.

0-7803-9050-4/05/$20.00 ©2005 IEEE.

IV. Experimental Results and Discussion

Figure 2 shows the number of groups formed for each of the performance metrics. Table I shows the number of discriminant bases chosen when each of the classifiers is used for each automated grouping performance metric, feature extraction, and classifier investigated. Figure 3 shows the classifier performance for Analysis I and Analysis II when various feature extraction methods were investigated. It can be concluded that using supervised methods, such as LDA, is very beneficial. Figures 4 and 5 show the overall accuracies when using LDA as the feature extraction technique for the k-NN and ML classifiers, respectively. In Figures 4 and 5, results for both grouping methods (manual and automated) and all ten performance metrics are shown. From Figures 4 and 5, it is clear that the multiclassifiers (Analysis II) provide significant accuracy improvements. Interestingly, when using manual grouping, group sizes of 10, 20, or 30 result in accuracies that are on par with the automated grouping. For the automated grouping, the weighted sum performance metrics result in the highest overall classification accuracies. However, the accuracies are only slightly higher than those for the product performance metrics, particularly Corr*BD and Corr*JM. And since the weighted sum metrics are more likely to become over trained, as compared to the product metrics, the product metrics may be more desirable.

The proposed methods gave improved results, compared to traditional methods, such as the greedy search best band selection [6-7]. For this study's dataset, the computation time of the greedy search algorithm was 3.02 hours. The grouping algorithm followed by decision fusion seems more complex; however, the computational time is drastically less. Grouping of the hyperspectral data took an average of 2 sec (for manual grouping) and 28.8 sec (for grouping using Corr+ROC performance metric). A comparison between analyses I and II clearly shows that decision fusion using QMV improves classification accuracies significantly. However, the multiclassifier approach did not require significantly more computational time. After the grouping stage, the time taken for performing analyses I and II were 20.75 sec and 20.79 sec, respectively.

Table I. Number of Discriminant Bases (Best Groups) for Automated Band Grouping

	Mean kNN	Mean ML	Var kNN	Var ML	PCA kNN	PCA ML	LDA kNN	LDA ML
Corr	44	109	44	109	44	109	86	48
BD	14	36	13	36	14	36	91	84
JM	14	36	13	36	14	36	91	84
ROC	15	34	14	34	14	34	91	83
Corr*BD	14	33	13	33	13	33	91	83
Corr*JM	14	31	14	38	14	37	108	85
Corr*ROC	16	50	19	50	25	55	125	88
Corr+BD	16	35	13	35	15	35	87	84
Corr+JM	15	33	14	35	14	34	89	81
Corr+ROC	17	48	25	57	25	51	120	82

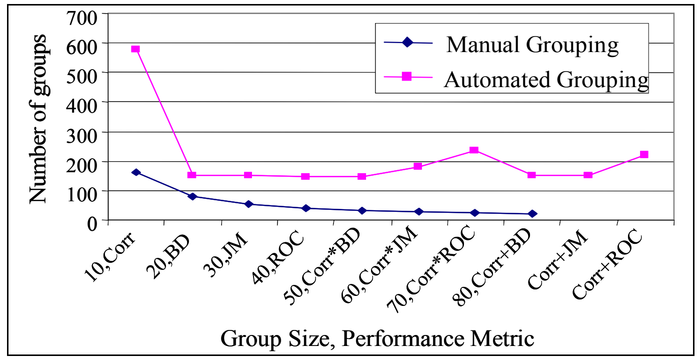

Figure 2. Total number of spectral band groups formed versus automated grouping performance metric (e.g. Corr, BD, etc) and versus manual group size (10, 20, etc).

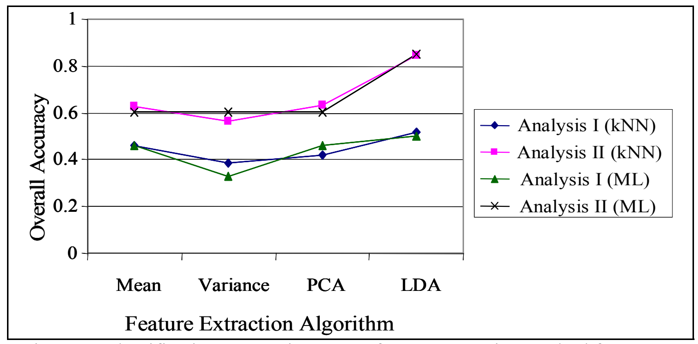

Figure 3. Classification accuracies versus feature extraction method for automated band grouping when using Corr*BD performance metric.

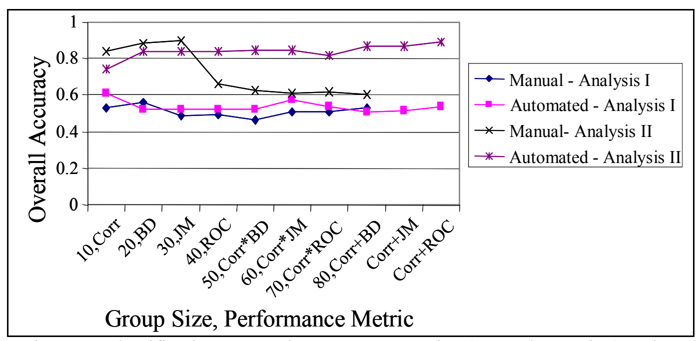

Figure 4. Classification accuracies versus group size (manual grouping) and versus performance metric (automated grouping) for Analysis I and II when using LDA feature extraction and k-NN classifier.

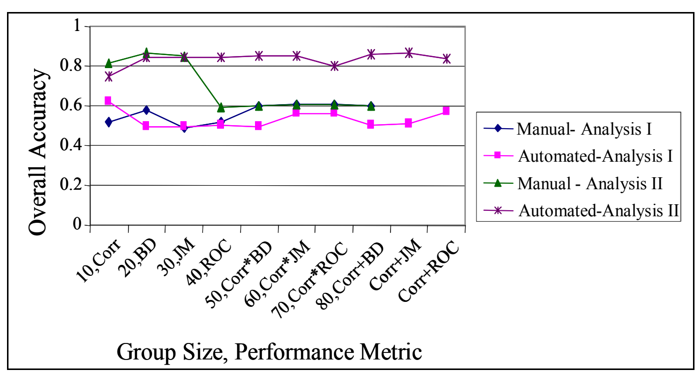

Figure 5. Classification accuracies versus group size (manual grouping) and versus performance metric (automated grouping) for Analysis I and II when using LDA feature extraction and ML classifier.

V. Conclusions

A new approach for analyzing and classifying high dimensional hyperspectral datasets has been proposed by the authors. The proposed system utilizes four approaches, (i) grouping of adjacent spectral bands, (ii) feature extraction from resultant groups, (iii) classification of feature vectors, and (iv) optionally, fusion of the groups' classification decision if a multiclassifier approach was used in (iii). The approaches were tested on experimental hyperspectral datasets, where the goal was to discriminate two similar vegetations, Cogongrass and Johnsongrass. Ten performance metrics were investigated for the band grouping, where weighted sums of correlation and class separation (BD and JM) resulted in the highest target detection accuracies. Unsupervised and supervised feature extraction methods were investigated, and it was shown that the supervised methods, specifically LDA, significantly outperformed the other methods. Two types of classifiers were investigated, k-NN and ML, resulting in similar classification accuracies. Finally, a single classifier versus a multiclassifier approach was investigated. The multiclassifier approach, using QMV for decision fusion, resulted in significantly higher classification accuracies than the single classifier method. In general, the automated band grouping using LDA feature extraction and multiclassifiers with QMV resulted in classification accuracies between 80% and 90%. These results were quite impressive and promising, considering the difficulty level of the dataset, where two similar grasses were the target and non-target classes. Furthermore, the computational time required for the proposed methods was significantly less than that required for existing methods, such as the greedy search best band selection.

References

[1] G.F. Hughes, "On the Mean Accuracy of Statistical Pattern Recognizers," *IEEE Trans. Infor. Theory,* vol. IT-14, no. 1, pp. 55-63, 1968.

[2] D. A Landgrebe, "Hyperspectral Image Data Analysis as a High Dimensional Signal Processing Problem," *IEEE Signal Processing Magazine*, vol. 19, pp. 17–28, Jan 2002.

[3] J. Kittler, M. Hatef, P. W. Duin and J. Matas, "On Combining Classifiers," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol.20, pp.226-239, Mar.1998

[4] A. Cheriyadat, L.M. Bruce, and A.Mathur, "Decision level fusion using best-bases for hyperspectral classification," *IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data,* Washington, D.C, Oct. 2003.

[5] ASD FieldSpec Pro FR portable spectroradiometer, Analytical Spectral Device Inc. 5335 Sterling Drive, Boulder, CO 80301-2344 USA.

[6] A. Mathur, L. M. Bruce, J. Byrd, "Discrimination of Subtly Different Vegetative Species via Hyperspectral Data," *Proc. IEEE International Geoscience and Remote Sensing Symposium*, vol. 2, pp. 805 –807, 2002.

[7] A. Mathur, L. M. Bruce, A. M. Cheriyadat and H. H. Lin, "Hyperspec – Analysis of Handheld Spectroradiometer Data," *Proc. IEEE International Geoscience and Remote Sensing Symposium*, Toulouse, France, July 2003.