IDENTIFICATION OF PERTINENT REGIONS IN SPECTRO-TEMPORAL MAPS FOR VEGETATIVE TARGET DETECTION

Abhinav Mathur, Lori Mann Bruce, Wilfredo Robles*, John Madsen*

Electrical and Computer Engineering Department, *Plant and Soil Sciences Department, Mississippi State University

ABSTRACT

Two classes of methods are designed for extracting features from spectro-temporal reflectance maps. Methods designed for these two approaches include various stepwise selection methods, windowing, and clustering techniques. The first class of methods is based on the consideration that all the elements of the spectro-temporal map are independent of each other (Mathur *et al.*, 2006a). The second class of methods is based on the consideration that the elements of the spectro-temporal map have some vicinal dependency among them (Mathur *et al.*, 2006b). Various data analyses are performed to evaluate the accuracies of the proposed methods. These include sub-sampling the original data at different rates in both spectral and temporal dimensions and then extracting features. Another set of analysis is done on data simulated according to various satellite and airborne sensor profiles. The efficacies of the new methods are demonstrated within an aquatic invasive species detection application, namely discriminating Waterhyacinth from other aquatic vegetation such as American Lotus.

INTRODUCTION

Various dimensionality reduction and feature extraction methods have been developed for both hyperspectral and multitemporal data (Kuo *et al.*, 2002 and Jimenez *et al.*, 1994). Temporal changes in hyperspectral signatures of vegetation, that occur due to varying growth patterns, could be utilized to discriminate between different plant species. As a plant goes through various growth stages, these differences could possibly be observed in the hyperspectral signature characteristics that change over time.

In recent years, a substantial amount of research has been done towards understanding and exploiting the underlying temporal features in multispectral and hyperspectral data. There are two basic forms of temporal approaches that researchers have studied. The first one is the bi-temporal approach, where two different time samples are considered and compared to extract useful information. The second one is the multi-temporal approach, where more than two time samples are considered and used to extract pertinent features to solve the classification problem at hand.

Kumar *et al.* (2001) stated that often best spectral features algorithms ignore the fact that adjacent spectral bands are highly correlated. They also pointed out that using just a 'single' global set of features is inefficient utilization of the available data. The authors proposed two best features selection techniques. The top-down approach breaks up the bands in to two sets (not necessarily equal) and replaces the final bands with the mean values. The bottom-up approach linearly merges the adjacent bands that are highly correlated to form bigger sets such that the band combination is projected in the Fisher's direction in order to maximize class discrimination.

Zabih *et al.* (2004) pointed out that in feature space clustering, spatial correlation between adjacent pixels in the image data is neglected. Therefore the feature space clusters lack spatial coherence. The authors presented a clustering technique that addresses this issue. They tested this method on two types of images. The first one was a synthetically generated image that composed of a bright region in a dark background, both of the same area. The second was a color image. This image had 16 labels to be segmented. Their technique is based on finding good clusters based on a quality measure and the spatial coherence. The output was a set of spatially coherent clusters, which could then be used in any segmentation algorithm.

For this study, an aquatic invasive species detection application is chosen, namely discriminating Waterhyacinth (*Eichornia crassipes*) from other aquatic vegetation such as American Lotus (*Nelumbo lutea*). Invasive aquatic plants like Waterhyacinth affect drainage for agriculture and forestry, aesthetics, drinking water quality, commercial and sport fishing, fish and wildlife habitat, habitats for other plants,

flood control, human and animal health, hydropower generation, irrigation, navigation, recreational boating, swimming, water conservation and transport, and, ultimately, land values. Most invasive aquatic plant species do not have natural control agents or competitors, as they have been introduced to this country from abroad, and they tend to dominate the aquatic systems to which they are exposed (Rockwell, 2003).

Waterhyacinth grows in ponds, canals, freshwater and coastal marshes, lakes, and back water sloughs and oxbows along rivers. It reproduces primarily from runners or stolons via vegetative means. This method of vegetative reproduction allows the plant to quickly colonize large areas in relatively short periods of time. American lotus grows in pools, marshes and swamps. It is also found in ponds, lakes, backwater areas and shallow embayments of reservoirs. It reproduces by seed which may remain viable for decades. These two aquatic plant species are chosen for testing the proposed methods as they have very similar physical characteristics and thus they simulate a real life target detection problem.

METHODOLOGIES

This work is based on the analysis of hyperspectral, hypertemporal data. This data was acquired over a period if 16 weeks from 24th June to 26th October 2005 using an Analytical Spectral Devices (ASD) handheld spectroradiometer (Hatchell, 1999). The efficacies of the new methods are demonstrated within an aquatic invasive species detection application, namely discriminating Waterhyacinth from other aquatic vegetation such as American Lotus. These two aquatic plant species are chosen for testing the proposed methods as they have very similar physical characteristics and thus they simulate a real life target detection problem.

Hyperspectral, multitemporal data is three-dimensional data that can be organized into a map where the x-axis is time, y-axis is wavelength, and z-axis is reflectance. This form of data organization is termed the spectro-temporal map (Mathur *et al.*, 2006a and Mathur *et al.*, 2006b) (Figure 1). These spectro-temporal maps are then analyzed using the two methods proposed by Mathur *et al.* (2006a and 2006b). The proposed methods optimize the selection of pertinent features from the spectro-temporal space data. This work designs two classes of methods for extracting features from spectro-temporal reflectance maps. Methods designed for these two approaches include various stepwise selection methods, windowing approaching, and clustering techniques.
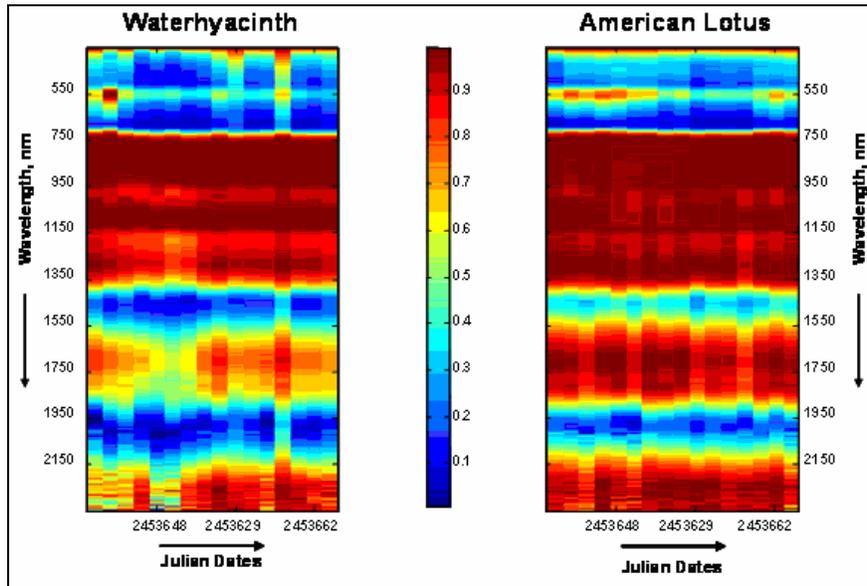
Figure 1: Spectro-temporal maps

The first class of methods is based on the consideration that all the elements of the spectro-temporal map are independent of each other. This method addresses the need to identify and extract the most useful information from a spectro-temporal map. Usefulness is measured in terms of signal classification on target detection. For this reason, instead of using the entire spectro-temporal map, pertinent features are extracted to reduce dimensionality. The classification accuracy increases if the distributions of the classes are statistically more separate in the feature space (Mathur *et al.*, 2006a).

The second class of methods is based on the consideration that the elements of the spectro-temporal map have some vicinal dependency among them. The hypothesis is that as time progresses, the spectral response of different plants change differently. Thus, there should be areas in the spectro-temporal maps of the two plant species that are different. These areas can then be used as robust features to accurately distinguish the two species (Mathur *et al.*, 2006b).

The efficacies of the new methods are demonstrated within an aquatic invasive species detection application, namely discriminating Waterhyacinth from other aquatic vegetation such as American Lotus.

DATA DEGRADATION

The spectro-temporal maps were used for estimating the efficacies of the proposed algorithms. Also these maps were degraded in the spectral as well as temporal dimensions before applying the algorithms. This degradation was done by sub-sampling the original spectro-temporal data at different rates in both spectral and temporal dimensions and then extracting features. For example in one case the spectro-temporal map is sub-sampled by a sub-sample rate of 500 in the spectral dimension and by a sub-sample rate of 4 in the temporal dimension. The resulting degraded data will have 4 bands (2150 original bands sub-sampled by a sub-sample rate of 500) and 4 dates (16 dates sub-sampled by a sub-sample rate of 4).

As the original data is a weekly data, sub-sample rates in the temporal dimension simulate the conditions of bi-monthly data (sub-sample rate 8), monthly data (sub-sample rate 4), bi-weekly (sub-sample rate 2) and weekly data (sub-sample rate 1). This analysis provided information about the tradeoff between the spectral and temporal resolutions required and desired overall accuracies. These results are shown in Table 2 to Table 7.

SIMULATED SENSORS

The available ASD data was used to simulate different sensors' data taking the available specifications into account. This simulated data was analyzed with the proposed feature extraction and classification methods to estimate their performance on satellite data that will be available in the future. This process will act as a pilot study to estimate the efficacy of the proposed methods when applied to data acquired from different air-borne and space-borne sensors. The original ASD data used to model sensor data is shown in Figure 2.

It is important note here that the simulation of the sensor data has been performed under ideal conditions. The system may not perform as well on actual sensor data for the following reasons. Firstly, the ASD data is taken of plant species planted in tanks under controlled environmental conditions. These conditions are not similar to those of ponds and lakes. The data is collected using handheld spectroradiometers. Therefore, the error introduced due to the atmosphere is minimum. Compared to ASD data, the actual sensor data will contain much more atmospheric noise as it is a satellite based or airborne. Secondly, as the ASD sensor is a handheld spectroradiometer, it collects reflectance spectra of true pixels as it is held close to the material under observation. Much more mixing of pixels will be introduced in the case of the actual sensor data depending upon the spatial resolution of the sensor. Thirdly, the satellite and airborne image data typically have geo-rectification performed on them. These effects will further decrease the quality of the data. These three factors will hinder the overall classification accuracies of the actual Hyperion data. Thus the accuracies reported here do not reflect the real life situation.

This analysis can however serve as a feasibility study for the end user. These results can provide the end user with information about which type of data will best solve the classification problem at hand and will assist in making a decision about which sensor is best suited for the task.
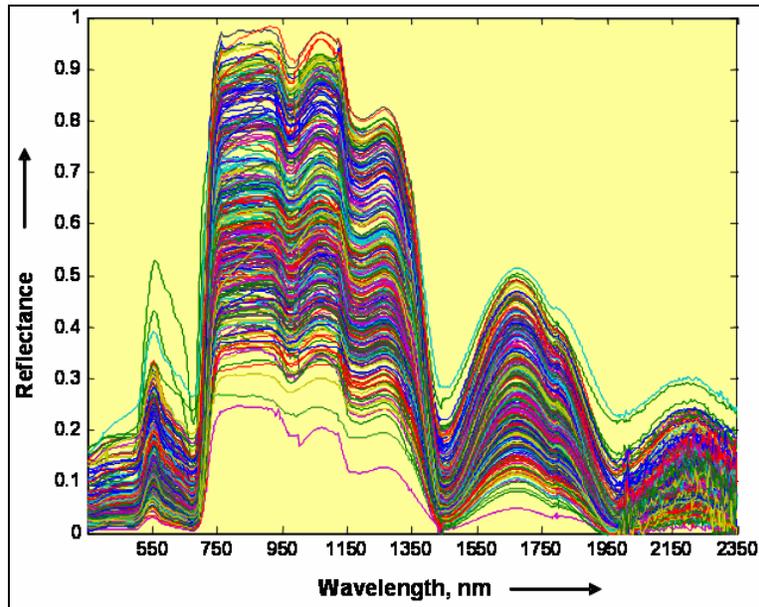
Figure 2: Original ASD waterhyacinth data.

The sensor profiles used are:

Advanced Land Imagery (ALI) (Bryant *et al.*)
Compact Airborne Spectrographic imager (CASI) (CASI 550 Airborne Hyperspectral Solutions)
GeoVantage (GeoVantage Precision Navigated Imagery: APPLICATIONS)
Hyperion (Carman)
Ikonos (Current and Future Sensor Systems)
Rdacs (Lobitz *et al.*, 1997)
TRWIS-2 (Pearlman *et al.*)
TRWIS-b (Pearlman *et al.*)
TRWIS-d (Pearlman *et al.*)

As an example, the sensor profile and simulated data is shown in Figure 3 and Table 1 (CASI 550 Airborne Hyperspectral Solutions) respectively.

Table 1: Sensor profile specifications for CASI.

Property name	Property value
Spectral range	400 – 1000 nm
Spectral Channels	288
Spatial Pixels	550
Total Field of View	40.4 Deg.
IFOV	1.34 mRad
Spectral Width Sampling/Row	1.9 nm
Spectral Resolution (FWHM)	2.2 nm
Pixel Size	22.5 x 15 microns
Dynamic range	14-bits (16384:1)
Burst Data Rate – Mega-pixels/second	1.25 MPix/sec
Spectral Smile	±0.8 pixels
Keystone Distortion	±0.8 pixels
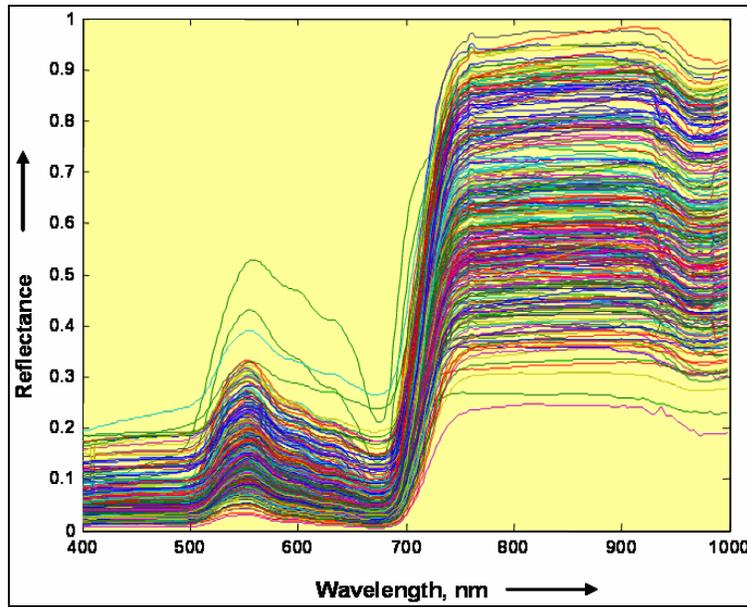
Figure 3: Data modeled according to CASI sensor profile

RESULTS

Table 2: Stepwise feature selection algorithm accuracies (%) of data with water band and thermal band removed

Spectral sub-sampling rate (no of bands)	Temporal sub-sampling rate (data collection frequency)			
	8 (bi-monthly)	4 (monthly)	2 (bi-weekly)	1 (weekly)
1000 (2)	90	91	91	98
500 (4)	99	100	100	100
200 (10)	96	98	99	99
100 (21)	95	100	100	100
50 (43)	92	100	100	100
20 (107)	99	100	100	100
10 (215)	99	100	100	100
5 (430)	97	100	100	100

Table 3: Clustering algorithm accuracies (%) of data with water band and thermal band removed

Spectral sub-sampling rate (no of bands)	Temporal sub-sampling rate (data collection frequency)			
	8 (bi-monthly)	4 (monthly)	2 (bi-weekly)	1 (weekly)
1000 (2)	97	98	98	100
500 (4)	98	100	100	100
200 (10)	100	100	100	100
100 (21)	100	100	100	100
50 (43)	100	100	100	100
20 (107)	100	100	100	100
10 (215)	100	100	100	100
5 (430)	100	100	100	100

Table 4: Stepwise feature selection algorithm accuracies (%) of data with water band, thermal band and continuum removed

Spectral sub-sampling rate (no of bands)	Temporal sub-sampling rate (data collection frequency)			
	8 (bi-monthly)	4 (monthly)	2 (bi-weekly)	1 (weekly)
1000 (2)	73	66	92	92
500 (4)	87	76	85	85
200 (10)	93	91	88	91
100 (21)	98	93	93	100
50 (43)	100	94	98	100
20 (107)	100	98	99	100
10 (215)	100	98	99	100
5 (430)	99	99	99	100

Table 5: Clustering algorithm accuracies (%) of data with water band, thermal band and continuum removed

Spectral sub-sampling rate (no of bands)	Temporal sub-sampling rate (data collection frequency)			
	8 (bi-monthly)	4 (monthly)	2 (bi-weekly)	1 (weekly)
1000 (2)	79	94	92	92
500 (4)	86	87	97	98
200 (10)	100	100	100	100
100 (21)	100	100	100	100
50 (43)	100	100	100	100
20 (107)	100	100	100	100
10 (215)	100	100	100	100
5 (430)	100	100	100	100

Table 6: Stepwise feature selection algorithm accuracies (%) of data with water band and continuum removed

Spectral sub-sampling rate (no of bands)	Temporal sub-sampling rate (data collection frequency)			
	8 (bi-monthly)	4 (monthly)	2 (bi-weekly)	1 (weekly)
1000 (2)	86	70	98	78
500 (4)	89	79	98	74
200 (10)	95	90	100	99
100 (21)	97	93	98	100
50 (43)	97	99	100	100
20 (107)	100	99	98	100
10 (215)	100	99	100	100
5 (430)	100	100	100	100

Table 7: Clustering algorithm accuracies (%) of data with water band and continuum removed

Spectral sub-sampling rate (no of bands)	Temporal sub-sampling rate (data collection frequency)			
	8 (bi-monthly)	4 (monthly)	2 (bi-weekly)	1 (weekly)
1000 (2)	87	92	98	98
500 (4)	100	89	99	97
200 (10)	100	100	99	100
100 (21)	100	100	100	100
50 (43)	99	100	100	100
20 (107)	100	100	100	100
10 (215)	100	100	100	100
5 (430)	100	100	100	0

Table 8: Stepwise feature selection overall accuracies (%) for data modeled according to different sensor profiles

Sensor	Temporal sub-sampling rate (data collection frequency)			
	8 (bi-monthly)	4 (monthly)	2 (bi-weekly)	1 (weekly)
ALI	99	99	100	99
CASI	99	100	100	99
GeoVantage	96	94	98	99
Hyperion	98	99	99	99
Ikonos	92	90	99	100
Rdacs	94	93	98	98
Trwis 2	100	100	100	100
Trwis b	94	98	99	97
Trwis d	98	100	100	100

Table 9: Clustering overall accuracies (%) for data modeled according to different sensor profiles

Sensor	Temporal degradation step			
	8 (bi-monthly)	4 (monthly)	2 (bi-weekly)	1 (weekly)
ALI	100	100	100	100
CASI	100	100	100	100
GeoVantage	100	100	100	100
Hyperion	100	100	100	100
Ikonos	100	100	100	100
Rdacs	100	100	100	100
Trwis 2	100	100	100	100
Trwis b	100	100	100	100
Trwis d	100	100	100	100

DISCUSSION AND CONCLUSIONS

As seen from Tables 2 to 7, the over all accuracies generally increase as the degradation decreases in both spectral and temporal directions. This observation can be attributed to the fact that as the degradation decreases, the features available initially are more. This provides more information about the difference between the two plant species to the two proposed feature extraction techniques. Thus the algorithms extract optimum features that are based on more details available about the spectral and temporal change differences in the two plant species under consideration.

It is however observed that the increase in overall accuracies for the clustering algorithms is quicker than that for the stepwise feature selection. Higher accuracies are also observed for the clustering algorithm than the stepwise feature selection in the case of data modeled according to the different sensor profiles. This general observation indicates that the clustering algorithm performs better as compared to the stepwise feature extraction algorithm. This inference could be explained as follows.

The stepwise feature selection algorithm considers all the features in the spectro-temporal map as individual independent features. Thus the algorithm finds isolated features in the map that are supposed to work well together towards solving the overall classification problem. It forms a set of robust features that perform this task. It however, does not consider that the spectro-temporal features are actually highly correlated at many locations. This high correlation can attributed to the hypothesis that the changes in the reflectance of closely spaced wavelengths do not always change dramatically over time. This consideration is exploited in the clustering algorithm. The final clusters formed using this method, provide more localized group of features that work very well towards classifying the two plant species' data. Specifically, Table 8 and Table 9 provide a pilot study as to how effective the two algorithms will be if the data from the different considered sensors is available. This analysis can provide valuable information to the end user about what type of data is best suited to solve the classification problem at hand.

Table 2 to Table 9 provide expected accuracies under scenarios of different time intervals between data collection. In these tables, analysis has been done for this interval to be two months, one month, two weeks and one week. For most cases, the accuracies increase as this interval decreases.

REFERENCES

Bryant, R., M.S. Moran, S. McElroy, C. Holifield, K. Thome, T. Miura, and S.F. Biggar. Data Continuity of Earth Observation (EO-1) Advanced Land Imager (ALI) and Landsat TM and ETM+. http://eo1.gsfc.nasa.gov/new/validationReport/Technology/Documents/Tech.Val.Report/Science_Summary_Moran_Bryant_1.pdf

Carman, S. Hyperion Grating Imaging Spectrometer. http://eo1.gsfc.nasa.gov/overview/Workshop/06.pdf

Hatchell, D.C. (1999). Analytical Spectral Devices, Inc. (ASD) Technical Guide, third edition.

Jimenez, L., and D. A. Landgrebe (1994). High dimensional feature reduction via projection pursuit. *Geoscience and Remote Sensing Symposium*, 2:1145 – 1147.

Kumar, S., J. Ghosh, and M. M. Crawford (2001). Best-Bases Feature Extraction Algorithms for lassification of Hyperspectral Data. *IEEE Transactions on Geoscience And Remote Sensing*, 39(7).

Kuo, B., and D. A. Landgrebe (2002). Hyperspectral data classification using nonparametric weighted feature extraction. *Geoscience and Remote Sensing Symposium*, 3:1428 – 1430.

Lobitz, B., L. Johnson, C. Hlavka, R. Armstrong and C. Bell (1997). Grapevine Remote sensing Analysis of Phylloxera Early Stress (GRAPES): Remote Sensing Analysis Summary. *NASA Technical Memorandum 112218*.

Mathur, A., L. M. Bruce, D. W. Johnson, W. Robles, and J. Madsen (2006a). Automated Step-Wise Selection of Hyperspectral-Hypertemporal Features for Target Detection. *IEEE International Geoscience and Remote Sensing Symposium* - pending.

Mathur, A., L. M. Bruce, D. W. Johnson, W. Robles, and J. Madsen (2006b). Exploiting Hyperspectral-Hypertemporal Imagery via a Novel Clustering Approach. *IEEE International Geoscience and Remote Sensing Symposium* – pending.

Pearlman, J. S., J. Marmo, M. Folkman, and S. Sandor, "Lewis Satellite Hyperspectral Payloads And Applications Demonstrations. http://ltpwww.gsfc.nasa.gov/ISSSR-95/lewissat.htm

Rockwell, Jr. H. W (2003). Summary of a Survey of the Literature on the Economic Impact of Aquatic Weeds. *The Economic Impact of Aquatic Weeds*.

Zabih, R., V. Kolmogorov (2004). Spatially Coherent Clustering Using Graph Cuts. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2(II): 437444.

CASI 550 Airborne Hyperspectral Solutions. http://www.itres.com/files/attachments/files/Itres_CASI550.pdf

Current and Future Sensor Systems. http://geo.arc.nasa.gov/sge/health/sensor/cfsensor.html

GeoVantage Precision Navigated Imagery: APPLICATIONS. http://www.geovantage.com/products/index.html