Characterizing the Gradient of the Spectral Curve for Object-Based Random Forest Image Classification Using Airborne Hyperspectral Datasets

Eric Ariel Salas[1] and Sakthi Kumaran Subburayalu[1]

[1]Central State University

November 21, 2022

Abstract

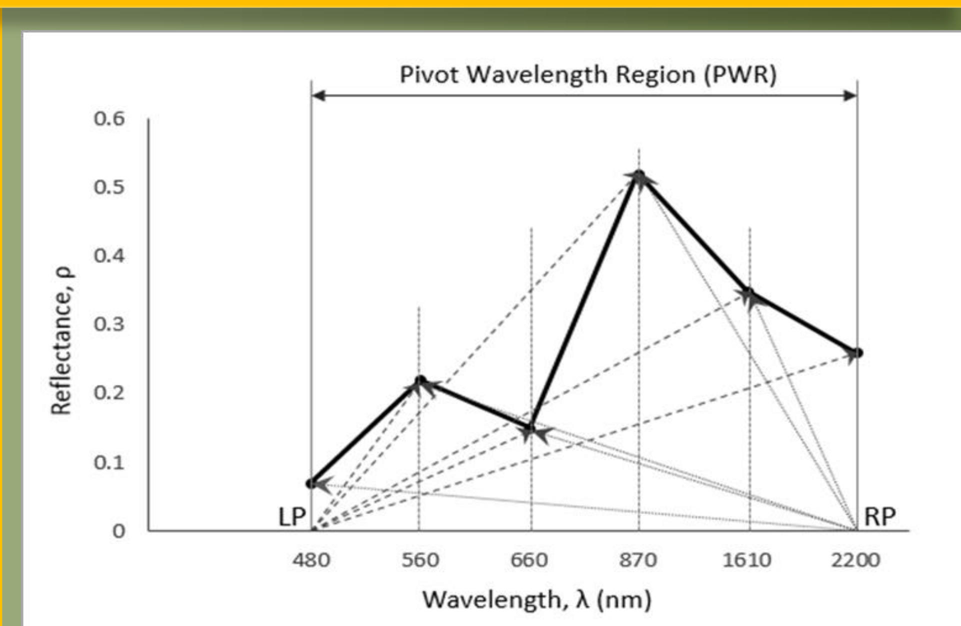
There have always been some challenges within the remote sensing community related to the processing of contiguous spectral bands contained in hyperspectral datasets. Most approaches would resort to using averaged spectral information over wide bandwidths resulting in loss of crucial information available in those contiguous bands. The loss of information could mean a drop in the discriminative power when it comes to land cover classes with comparable spectral responses, as in the case of cultivated fields versus fallow lands. In this study, we proposed and tested three optimized novel algorithms based on Moment Distance Index (MDI) that characterizes the whole shape of the spectral curve. The image classification tests conducted on two publicly available hyperspectral data sets (AVIRIS 1992 Indian Pine and HYDICE Washington DC Mall images) showed the robustness of the optimized MDI algorithms in terms of classification accuracy. We achieved an overall accuracy of 98% and 99% for AVIRIS and HYDICE, respectively using the optimized MDI algorithms. The optimized indices were also time efficient as it avoided the process of band dimension reduction, such as those implemented by several well-known classifiers. Our results showed the potential of the optimized shape indices to discriminate between grass/pasture and grass/trees, tree and grass, and between types of tillage (corn-min and corn-notill) under object-based random forest approach. The results highlight the importance of MDI that completely utilizes the contiguous spectral bands to define the gradient of the curve and improve image classification accuracy.

Characterizing the Gradient of the Spectral Curve for Object-Based Random Forest Image Classification Using Airborne Hyperspectral Datasets

Eric Ariel L. Salas[1] and Sakthi Subburayalu[1]
[1] Agricultural Research and Development Program, College of Science and Engineering, Central State University, Wilberforce, Ohio, USA, 45384

1 Overview

- There are challenges related to the processing of contiguous spectral bands contained in hyperspectral datasets.
- Approaches resort to using averaged spectral information over wide bandwidths resulting in loss of crucial information available in those contiguous bands.
- So, how to discriminate comparable spectral responses, as in the case of cultivated fields versus fallow lands?
- We proposed and tested three optimized novel algorithms based on Moment Distance Index (MDI) that characterizes the whole shape of the spectral curve.
- We used two publicly available hyperspectral data sets (AVIRIS 1992 Indian Pines and HYDICE Washington DC Mall images).
- Optimized MDI algorithms achieved an overall classification accuracy of 98% and 99% for AVIRIS and HYDICE.
- Results showed the potential of the optimized shape indices to discriminate between grass/pasture and grass/trees, tree and grass, and between types of tillage (corn-min and corn-notill) under object-based random forest approach.



$$MD_{LP} = \sum_{i=\lambda_{LP}}^{\lambda_{RP}} ({\rho_i}^2 + (i - \lambda_{LP})^2)^{0.5} \quad [1]$$

$$MD_{RP} = \sum_{i=\lambda_{RP}}^{\lambda_{LP}} ({\rho_i}^2 + (\lambda_{RP} - i)^2)^{0.5} \quad [2]$$

$$MDI = MD_{RP} - MD_{LP} \quad [3]$$

2 Methodology

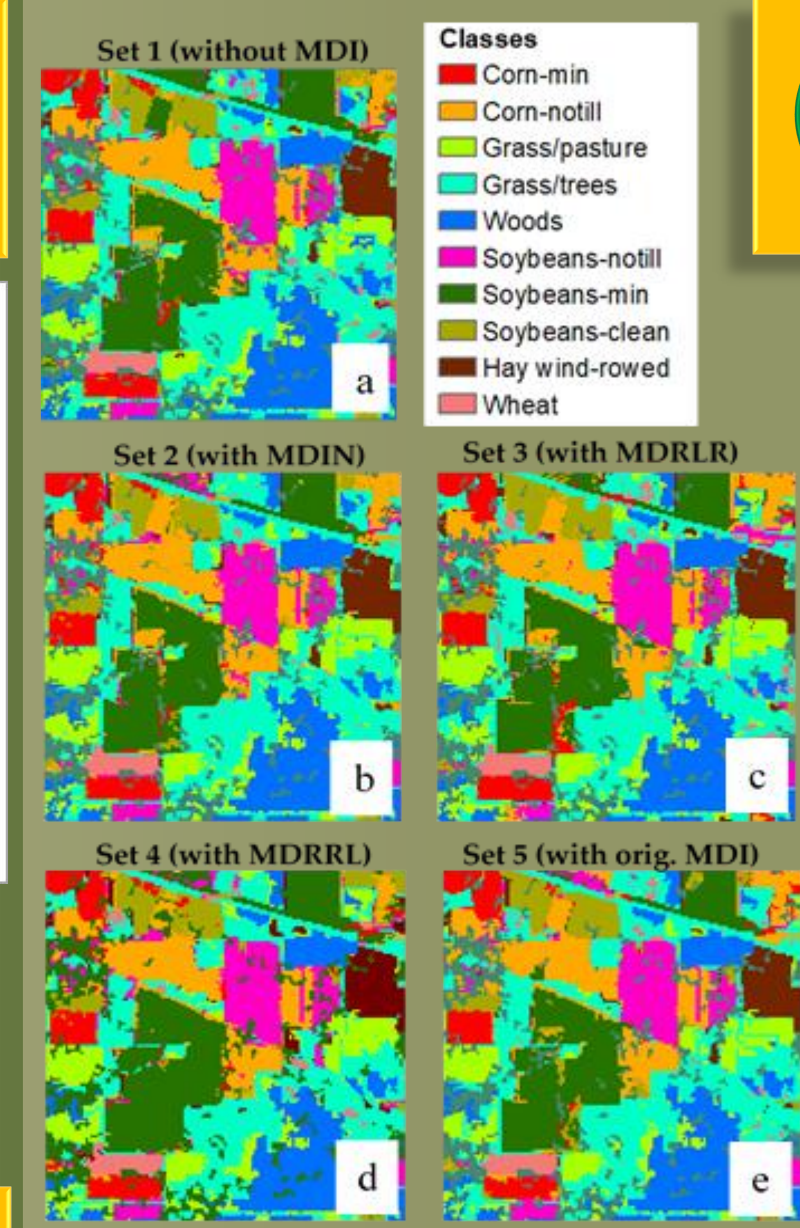
- Used Airborne AVIRIS Indian Pines and HYDICE Washington DC Mall images
- Derived spectral indices and image textures: variance (VAR), entropy (ENT), correlation (COR), contrast (CON), and angular second moment (ASM)
- Performed Object-based image analysis (OBIA)
- Ran different scale levels during segmentation process to incorporate the scale effects on prediction accuracy
- 63 segmented variables resulting from three different scales
- Applied Random Forest classifier to five sets of data, with and without MDI
- Evaluated classification accuracy and tested significant differences in classification

Table 1: Spectral indices used as input predictor variables.

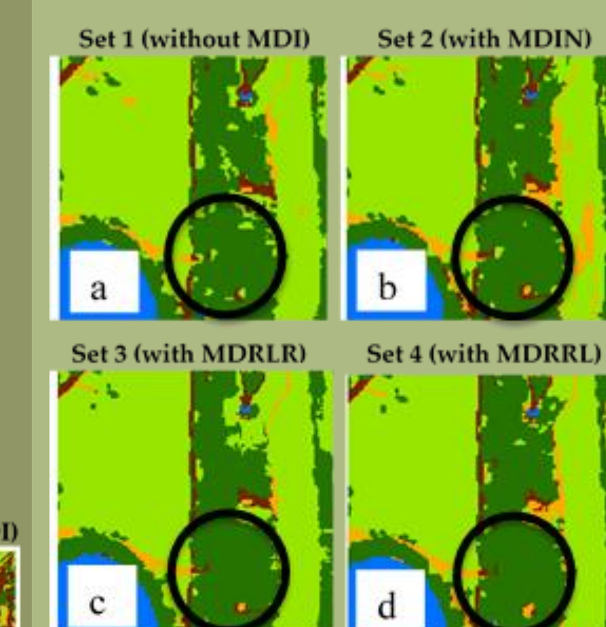
VARIABLES	EQUATION
NDVI	$\frac{Red - NIR}{Red + NIR}$
EVI	$2.5 * \frac{NIR - Red}{1 + NIR + 6 * Red - 7.5 * Blue}$
NDII	$\frac{\lambda 819 - \lambda 1649}{\lambda 819 + \lambda 1649}$
NRI	$\frac{Green - Red}{Green + Red}$
PSRI	$\frac{Red - Blue}{NIR}$
PRI	$\frac{\lambda 529 - \lambda 580}{\lambda 529 + \lambda 580}$
MDIN	$\frac{MD_{RP} - MD_{LP}}{MD_{RP} + MD_{LP}}$
MDRLR	$\frac{MD_{LP}}{MD_{RP}}$
MDRRL	$\frac{MD_{RP}}{MD_{LP}}$

Table 2: Five sets of data separately used as inputs in the object-based random forest classification.

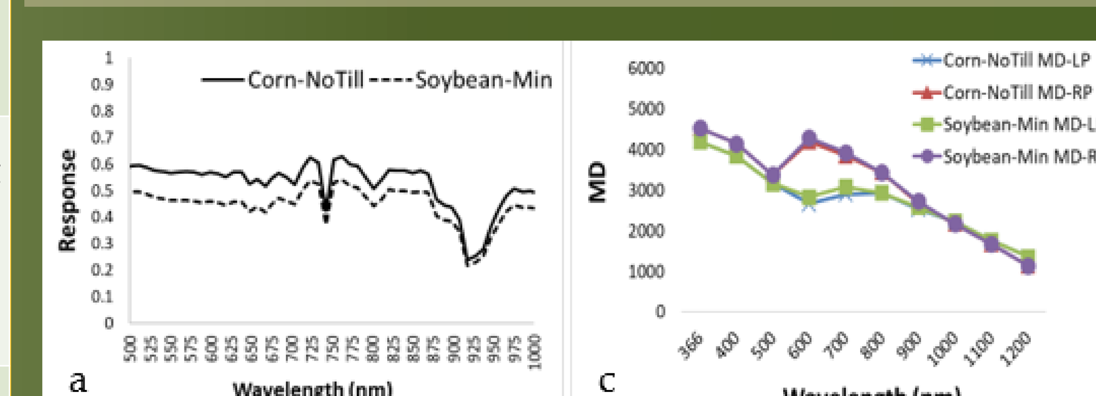
SETS	VARIABLE INPUTS	TOTAL SEGMENTED VARIABLES
1	NDVI, EVI, NDII, NRI, PSRI, PRI, Texture (VAR, ENT, COR, CON, ASM)	63 segmented variables resulting from three different scales
2	NDVI, EVI, NDII, NRI, PSRI, PRI, MDIN, Texture (VAR, ENT, COR, CON, ASM)	66 segmented variables resulting from three different scales
3	NDVI, EVI, NDII, NRI, PSRI, PRI, MDRLR, Texture (VAR, ENT, COR, CON, ASM)	66 segmented variables resulting from three different scales
4	NDVI, EVI, NDII, NRI, PSRI, PRI, MDRRL, Texture (VAR, ENT, COR, CON, ASM)	66 segmented variables resulting from three different scales
5	NDVI, EVI, NDII, NRI, PSRI, PRI, Original MDI, Texture (VAR, ENT, COR, CON, ASM)	66 segmented variables resulting from three different scales



AVIRIS Indian Pines: Difference of the performances of using a dataset (a) without MDI, (b) with MDIN, (c) with MDRLR, and (d) with MDRRL for classes soybean-min and corn-notill.



HYDICE Washington DC Mall: Difference of the performances of using a dataset (a) without MDI, (b) with MDIN, (c) with MDRLR, and (d) with MDRRL for classes tree and grass.



How MDI values vary moving the pivot from left to right (and vise versa) for quite similar spectral responses of corn-notill and soybean-min.

3 Results

Class	Set 1 (no MDI)		Set 2 (MDIN b)		Set 3 (MDRLR b)		Set 4 (MDRRL a,b)		Set 5 (orig. MDI b)	
	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA
Corn-min	80.6	80.0	95.6	92.1	98.2	94.9	97.0	94.3	92.1	89.1
Corn-notill	82.7	85.9	90.3	94.9	86.7	95.7	94.4	96.1	89.3	85.3
Grass/pasture	97.4	91.7	97.2	93.6	98.1	94.6	99.2	99.2	97.5	99.2
Grass/trees	97.8	98.8	99.7	96.0	97.9	98.9	100	98.4	99.0	98.5
Woods	98.8	99.4	100	100	99.3	100	100	100	99.7	100
Soybean-notill	99.2	89.6	99.2	97.3	98.8	95.8	98.5	98.7	99.6	93.5
Soybean-min	89.5	97.0	94.5	99.1	95.2	97.6	98.5	99.6	87.8	99.4
Soybean-clean	90.6	72.7	98.1	93.2	99.4	81.7	95.9	93.0	98.2	81.2
Hay	97.8	99.2	100	99.2	97.1	100	100	97.5	100	98.4
Wheat	100	98.1	95.4	100	100	89.5	100	100	100	96.6

Table 3: Summary of classification accuracies (%) from five sets of data using AVIRIS Indian Pines image. The marks a,b signify that the set produces significant differences at the 5% level against set 5 and set 1, respectively.

Class	Set 1 (no MDI)		Set 2 (MDIN b)		Set 3 (MDRLRa,b)		Set 4 (MDRRLa,b)		Set 5 (orig. MDI b)	
	PA	UA	PA	UA	PA	UA	PA	UA	PA	UA
Water	100	100	100	100	100	100	100	100	100	100
Tree	86.4	89.7	90.5	99.9	99.8	99.8	99.8	100	91.2	96.4
Grass	90.6	94.5	93.8	96.6	100	99.9	100	99.9	100	93.8
Road	87.5	68.6	95.0	67.3	99.7	100	98.7	100	78.7	93.0
Pathway	77.3	63.0	96.4	97.2	100	98.7	100	97.8	85.9	100

Table 4: Summary of classification accuracies (%) from five sets of data using HYDICE Washington DC Mall image. The marks a,b signify that the set produces significant differences at the 5% level against set 5 and set 1, respectively.

Rank	Set 3 (AVIRIS) (with MDRLR)	Set 4 (AVIRIS) (with MDRRL)	Set 3 (HYDICE) (with MDRLR)	Set 4 (HYDICE) (with MDRRL)
	Variable	Variable	Variable	Variable
1	NRI@10	MDRRL@5	PSRI@10	MDRRL@20
2	MDRLR@20	EVI@20	MDRLR@20	NDVI@20
3	EVI@5	NRI@10	ENT@10(PCA1)	NDVI@10
4	PRI@10	NDVI@10	NDVI@20	MDRRL@10
5	ENT@10(PCA1)	MDRRL@10	PSRI@20	PSRI@10
6	EVI@10	PRI@5	NDVI@10	NDVI@5
7	MDRLR@5	NDVI@20	PSRI@5	PSRI@20
8	NDVI@20	PRI@10	MDRLR@5	MDRRL@5
9	NDVI@5	MDRRL@20	MDRLR@10	PSRI@5
10	NDVI@10	NDII@20	PRI@20	NDII@20

Table 5: Rankings of the top 10 object features with maximum importance across classes in the RF model using AVIRIS Indian Pines and HYDICE Washington DC Mall images. The segmentation scales are also listed after the "@" symbol.

4 Conclusion

- We achieved an overall accuracy of 98% and 99% for AVIRIS and HYDICE, respectively.
- Differences in absorption features from two tillage systems became magnified and highlighted the shape differences of the individual spectral curve.
- We showed the potential of optimized shape indices, specifically the Moment Distance Ratio Right/Left (MDRRL), to discriminate between types of tillage (corn-min and corn-notill) and between grass/pasture and grass/trees, tree and grass under object-based random forest approach.