Improving classification accuracy of spectrally similar tree species: A complex case study in the Kruger National Park

Pravesh Debba

CSIR Built Environment,
Logistics and Quantitative Methods,
South Africa

Presented at Rhodes University 2009

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

Outline

What is classification?

- The aim of classification is to assign an object $\mathbf{x}$ into one class ω_i of a set of c given classes $\{\omega_1, \omega_2, \ldots, \omega_c\}$.
- Clustering — natural grouping for eg KNN, K-Means
- Classification — predicts categorical class labels for eg MLC, DT, NN
- Clustering – unsupervised learning — no training data or ground truth data — no predefined classes or no examples that would show the desired relationships
- Classification — supervised learning — have training data or ground truth data — have predefined classes

Common classification techniques

- Statistical
  - Parametric eg Naive Bayes, MLC
  - Non-parametric eg k-NN, Parzen
- Artificial Neural Networks
- Decision Trees
- Support Vector Machines

How to define similar?

- The definition of similarity is subjective.
- Similarity measures d_{ij}:
  - Squared Euclidean distance $d(\mathbf{x}_i, \mathbf{x}_j)^2 = (\mathbf{x}_i - \mathbf{x}_j)^T(\mathbf{x}_i - \mathbf{x}_j)$
  - Spectral angle/correlation
  - Spectral Information Divergence, etc.

If $d_{ij} < T$, (T: user defined threshold), the two pixel vectors are regarded as similar.

Overview of hyperspectral remote sensing

Hyperspectral sensors

- record the reflectance in many narrow contiguous bands
- various parts of the electromagnetic spectrum (visible - near infrared - short wave infrared)
- at each part of the electromagnetic spectrum results in an image

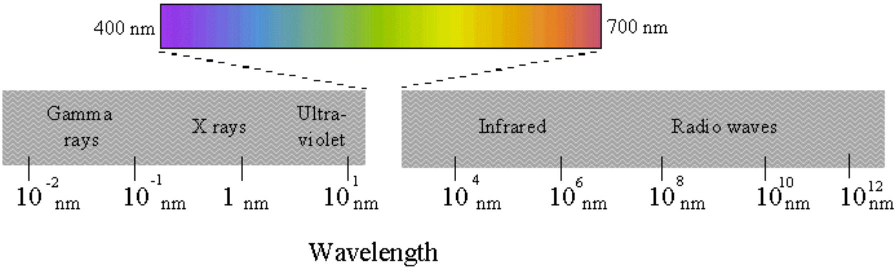

Figure: Spectral Range

Overview of hyperspectral remote sensing (cont. . .)

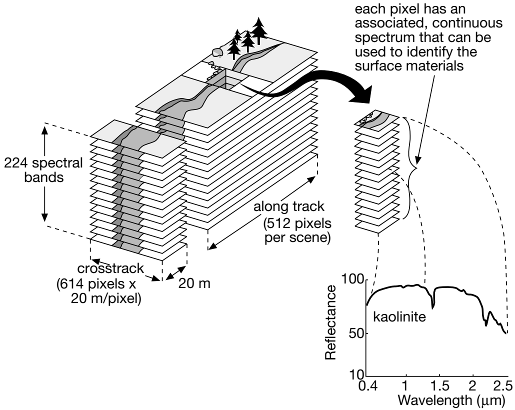

Figure: Hyperspectral cube

Overview of hyperspectral remote sensing (cont. . .)

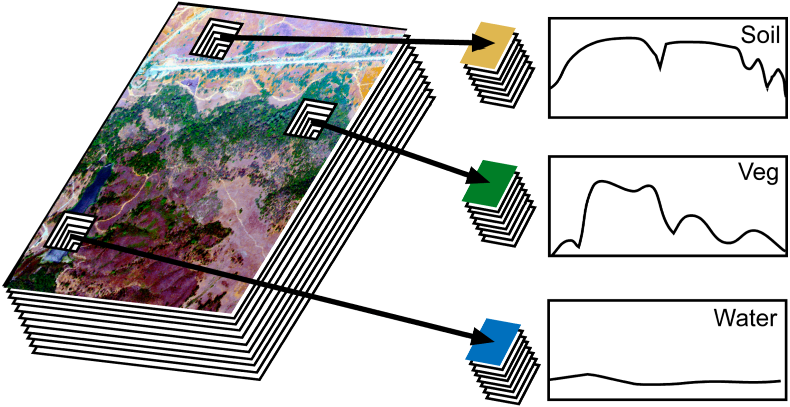

Figure: Pixels in hyperspectral image

Overview of hyperspectral remote sensing (cont. . .)

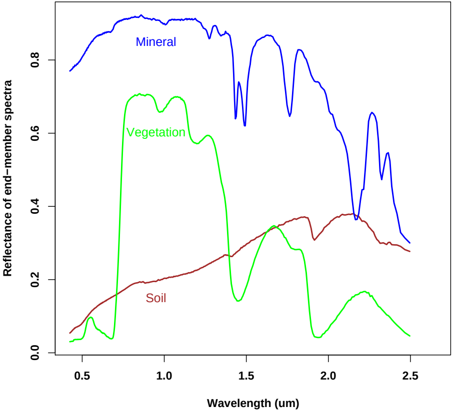

Figure: Example of 3 different spectral signatures

Iterated Conditional Modes (ICM) Algorithm

- Adequate image segmentation takes into account both spectral features and spatial information.
- Markov Random Fields (MRF) have been useful in this respect.
- $$\arg\min_k \left\{ \left(f_{ij} - \mu_k^{(\alpha)}\right)^T \left(f_{ij} - \mu_k^{(\alpha)}\right) - \beta \nu^{(\alpha)} N_{ij}^{(\alpha)}(k) \right\} \tag{1}$$

$$\nu^{(\alpha)} = \frac{1}{N} \sum_{k=1}^{K} \sum_{(i,j)\in \mathbf{C}_k^{(\alpha)}} \left(f_{ij} - \mu_k^{(\alpha)}\right)^T \left(f_{ij} - \mu_k^{(\alpha)}\right). \tag{2}$$

Second order MRF for ICM

A second order MRF was applied in which the neighbors of each pixel consists of its eight adjacencies, with border pixels adjusted appropriately.

1	2	2
1	(i,j)	2
3	3	5

$$N_{ij}^{(\alpha)}(1) = 2$$
$$N_{ij}^{(\alpha)}(2) = 3$$
$$N_{ij}^{(\alpha)}(3) = 2$$
$$N_{ij}^{(\alpha)}(4) = 0$$
$$N_{ij}^{(\alpha)}(5) = 1$$

Figure: Calculation of $N_{ij}^{(\alpha)}(k)$ for an arbitrary interior pixel (i,j) belonging to category k.

Study Site

- Study site – Tedej – Hungary.
- Crops: barely, maize, sugar beet, sunflower, alfalfa.
- Digital Imaging Spectrometer – DAIS-7915 – 79 channel hyperspectral image.
- Spectral range from visible (0.4 μm) to thermal infrared (12.3 μm).
- Spatial resolution 3–20 m depending on the carrier aircraft altitude.

Study Site

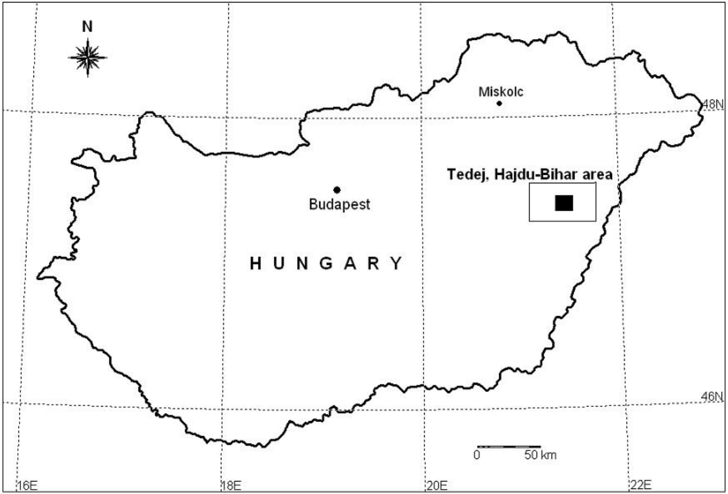

Figure: Study area in Tedej, Hajdu-Bihar area, Hungary.

Study Site

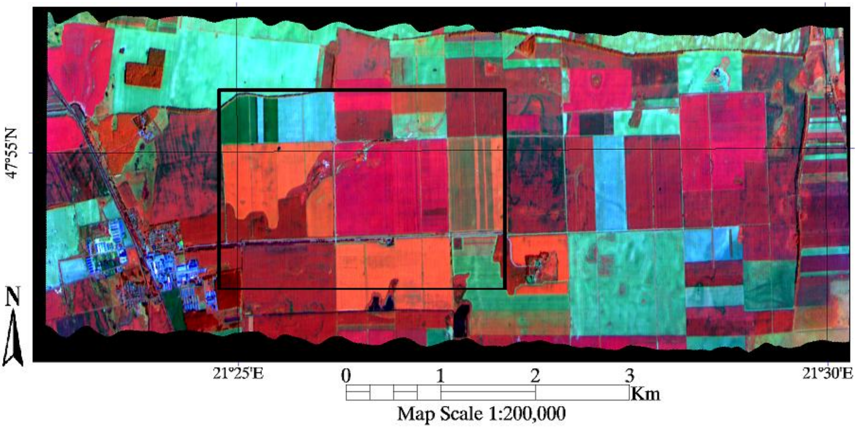

Figure: Hyperspectral image of study area in Tedej, Hajdu-Bihar area, Hungary. Reflectance values for bands 29 (0.988 μm), 39 (1.727 μm) and 1 (0.496 μm).

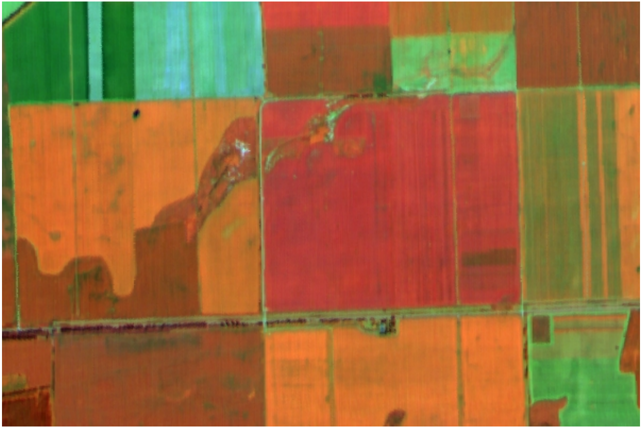

Figure: Original hyperspectral image. Reflectance values for bands 29 (0.988 μm), 39 (1.727 μm) and 1 (0.496 μm).

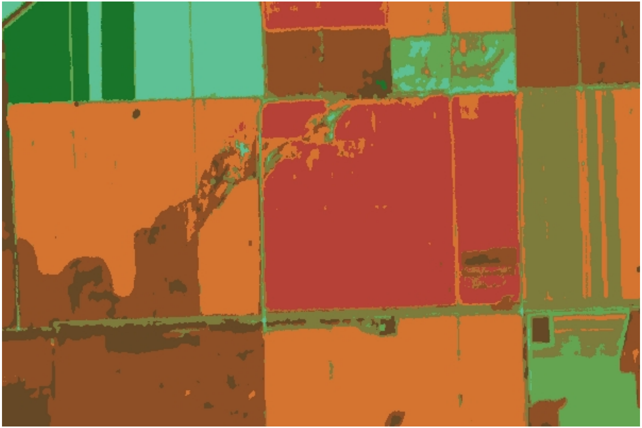

Figure: ICM Segmented image with eight categories.

Endmember Spectra

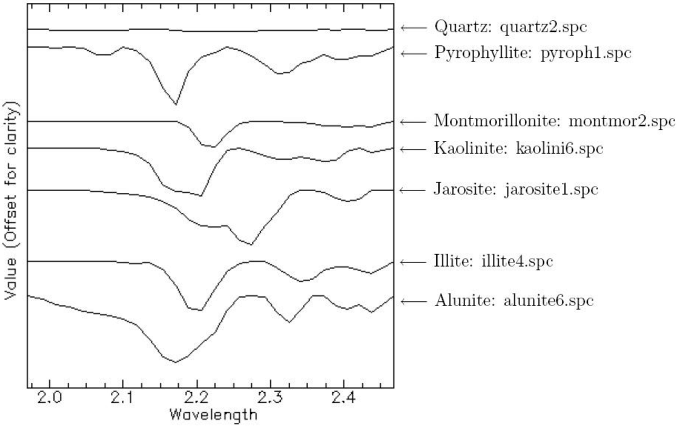

Figure: Plot of 7 endmembers from USGS spectral library for the 30 selected bands, enhanced by continuum removal.

Spectral Angle Mapper (SAM) Classifier

- SAM – pixel based supervised classification technique
- Measures the similarity of an image pixel reflectance spectrum to a reference spectrum
- Spectral angle (in radians) between the two spectra

$$\theta(\overrightarrow{\mathbf{x}}) = \cos^{-1}\left(\frac{f(\lambda)\cdot e(\lambda)}{||f(\lambda)||\cdot||e(\lambda)||}\right) \quad , \qquad (3)$$

$f(\lambda)$ – image reflectance spectrum and $e(\lambda)$ – reference spectrum.

- Results in a gray-scale rule image – values are the angles

Spectral Angle Mapper (SAM) Classifier

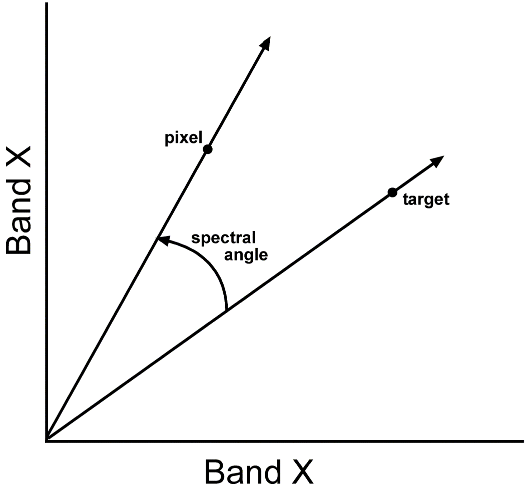

Figure: Spectral angle.

Study Site

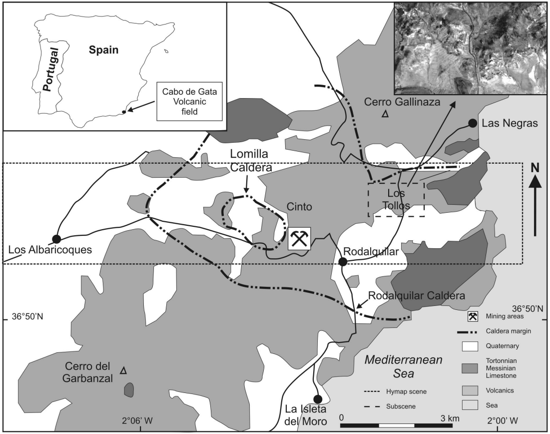

Figure: A generalized geological map of the Rodalquilar study area showing the flight line and the hyperspectral data

Data Used

- HyMap: 126 bands – 0.4–2.5 μm
- Geology: 30 bands – 1.95–2.48 μm
- Distinctive absorption features at wavelengths near 2.2 μm
- We collected field spectra during the over-flight using the Analytical Spectral Device (ASD) fieldspec-pro spectrometer – 0.35–2.50 μm

SAM Rule Image for Alunite

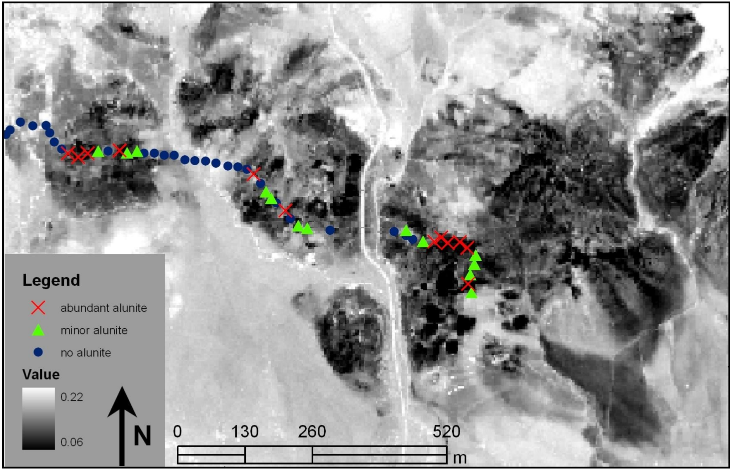

Figure: SAM classification rule image for alunite. Dark areas indicate smaller angles, hence, greater similarity to alunite.

Continuum Removal

Spectra are normalized to a common reference using a continuum formed by defining high points of the spectrum (local maxima) and fitting straight line segments between these points. The continuum is removed by dividing it into the original spectrum.

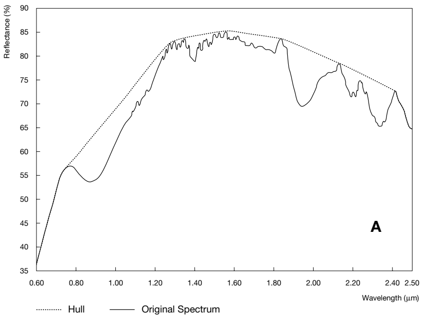

Figure: Concept of the convex hull transform; (A) a hull fitted over the original spectrum; (B) the transformed spectrum.

Continuum Removal

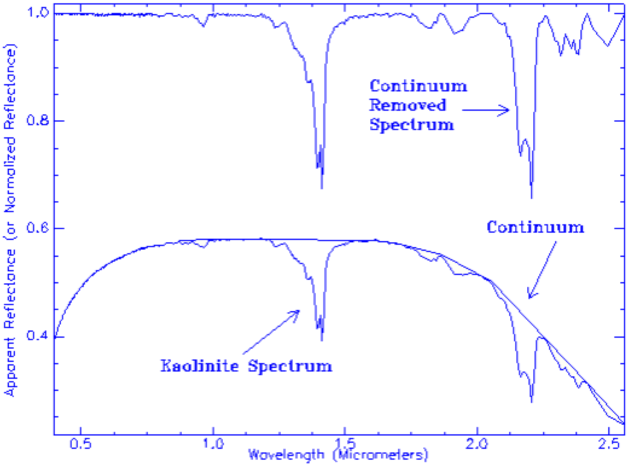

Figure: Original and continuum removed spectra.

Spectral Feature Fitting (SFF)

- SFF – pixel based classification technique.
- Remove the continuum from both the reference and unknown spectra.
- SFF produces a scale image for each endmember selected for analysis by first subtracting the continuum-removed spectra from one (inverting it), and making the continuum zero.
- SFF determines a single multiplicative scaling factor that makes the reference spectrum match the unknown spectrum.

Spectral Feature Fitting (SFF)

- SFF then calculates a least-squares-fit, band-by-band, between each reference endmember and the unknown spectrum.
- The total root-mean-square (RMS) error is used to form an RMS error image for each endmember.
- Scale/RMS provides a fit image that is a measure of how well the unknown spectrum matches the reference spectrum on a pixel-by-pixel basis.

SFF Rule Image for Alunite

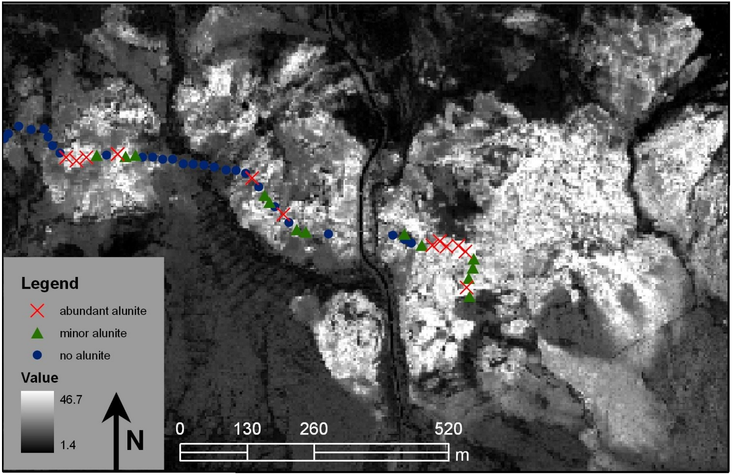

Figure: SFF fit image for alunite. Lighter areas indicate better fit values between pixel reflectance spectra and the alunite reference spectrum.

The Problem

- The 2 main types of variability, necessary for any image classification and/or spectral unmixing techniques are (i) the variability within a species class, and (ii) the similarity between the species classes.
- When the variability within a species class is small compared to the variability between the species classes, this results in relatively good accuracy for image classification and/or spectral unmixing.
- When the species spectra is similar, the within-species variability can be large compared to the between-species class variability — prominent in vegetation studies — producing poor results for image classification and/or spectral unmixing techniques.
- This research studies the variability within a species class and the variability between the species classes of seven spectrally similar tree species and presents ways in which the within-species class variability can be reduced compared to the between-species class variability.

Data description

- ASD spectrometer used to record hyperspectral measurements of leaf samples taken from several different savannah trees in the Kruger National Park in South Africa, in an attempt to assess tree species diversity in the park.
- The hyperspectral data consist of 2151 spectral bands at a spectral resolution of 1 nm for seven common plant tree species in the area.
- The seven tree species include Lonchocarpus capassa, Combretum apiculatum, Combretum heroense, Combretum zeyherrea, Gymnospora buxifolia, Gymnospora senegalensis, and Terminalia sericia.
- Each tree species has 10 measurements recorded with the exception of Gymnospora Buxifolia, which has only seven. The total data set therefore had 67 observations for the species measurements.

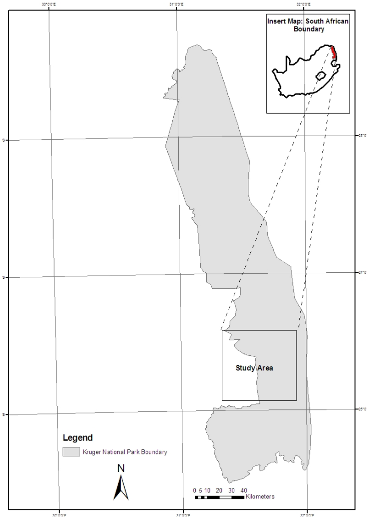

Figure: Study Area: Kruger National Park, South Africa

Figure: Variation of tree species in the Kruger National Park, South Africa

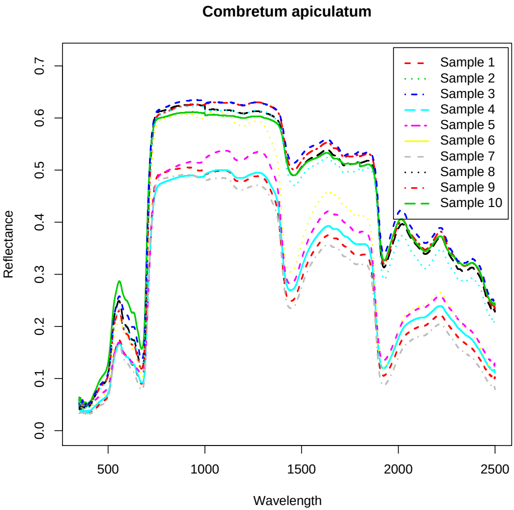

Figure: Reflectance spectra of the 10 samples for Combretum apriculatum.

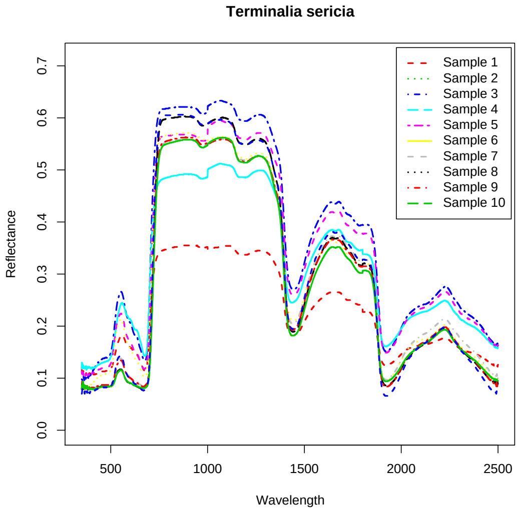

Figure: Reflectance spectra of the 10 samples for Terminalia sericia.

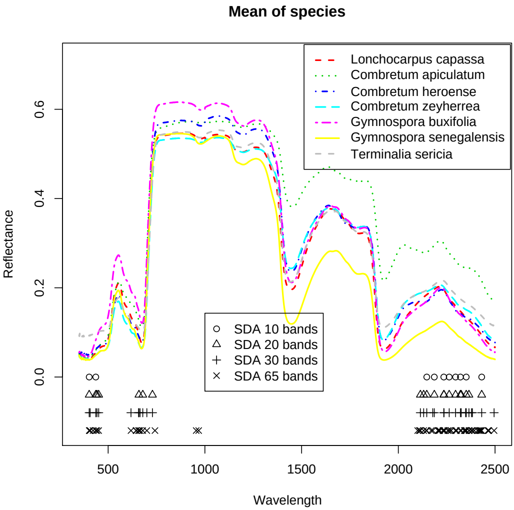

Figure: Mean spectral reflectance for all seven species. Also, band selection using stepwise discriminant analysis. For SDA the results for the best 10, 20, 30 and 65 selected bands are shown.

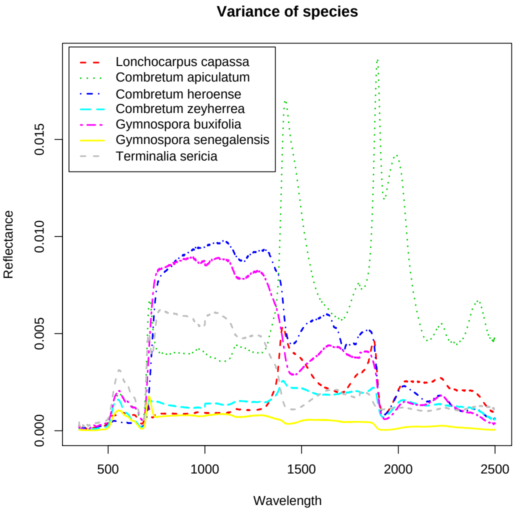

Figure: Variance of the spectral reflectance for all seven species.

Figure: Causes of high intra-species variability.

Method

- Let y_i^k denotes the d-dimensional feature vector (d represents the number of bands) selected from the i^{th} sample of the k^{th} class, c_k, with n_k samples in the k^{th} class.
- Also, let μ_k $(k = 1, \ldots, c)$ be the mean vector of k^{th} class and μ be the total mean vector in this d-dimensional feature space.
- The within-class variability, S_w and between-class variability, S_b:

$$S_w = \frac{1}{c} \sum_{k=1}^{c} \left[\frac{1}{n_k} \sum_{i=1}^{n_k} \left(y_i^k - \mu_k \right)^T \left(y_i^k - \mu_k \right) \right] \tag{4}$$

$$S_b = \frac{1}{c} \sum_{k=1}^{c} \left(\mu_k - \mu \right)^T \left(\mu_k - \mu \right). \tag{5}$$

- The ratio of the between-class variability to the within-class variability, commonly known as Fisher's criterion ratio, is a measure for class separation, with high values indicating greater class separation.

A comparison is made through evaluating the within-class species variability and the between-class species variability using:

- the original, first and second derivative spectra.
- for each, of the above, the experiment was conducted
  1. over the entire electromagnetic spectrum (EMS) (0.350–2.500 μm),
  2. the visible (VIS) (0.400–0.740 μm) region,
  3. the near infrared (NIR) (0.741–1.300 μm) region,
  4. the short wave infrared (SWIR) (1.301–2.500 μm) region,
  5. using band selection, for example, best 10, 20, 30 and 65 bands selected, through linear stepwise discriminant analysis (SDA),
  6. using sequential selection of bands, for example, every 5th, 9th, 15th, 19th or 25th band selected and
  7. spectral degradation of the spectral bands by averaging the reflectance values for every 5th, 9th, 15th, 19th or 25th band.

Table: Within- and between-class variability for various regions of the EMS.

Bands	Within-class var	Between-class var	Ratio
All			
Original	5.574	5.030	0.902
1st derivative	9.007×10^{-3}	4.000×10^{-3}	0.444
2nd derivative	1.522×10^{-2}	3.582×10^{-2}	0.235
VIS			
Original	0.316	0.291	0.920
1st derivative	2.220×10^{-4}	1.160×10^{-4}	0.523
2nd derivative	1.787×10^{-4}	2.797×10^{-5}	0.157
NIR			
Original	2.090	0.481	0.230
1st derivative	1.163×10^{-4}	4.254×10^{-4}	0.366
2nd derivative	2.557×10^{-4}	7.420×10^{-5}	0.290
SWIR			
Original	3.162	4.241	1.341
1st derivative	3.594×10^{-4}	1.568×10^{-4}	0.436
2nd derivative	7.013×10^{-4}	8.371×10^{-5}	0.119

Table: Within- and between-class variability for selected bands using SDA.

Bands	Within-class var	Between-class var	Ratio
SDA10			
Original	0.013	0.021	1.600
1st derivative	6.621×10^{-8}	1.445×10^{-7}	2.183
2nd derivative	4.763×10^{-11}	1.273×10^{-10}	2.672
SDA20			
Original	0.026	0.038	1.463
1st derivative	1.339×10^{-6}	8.253×10^{-7}	0.616
2nd derivative	2.061×10^{-7}	3.661×10^{-8}	0.178
SDA30			
Original	0.037	0.055	1.473
1st derivative	4.520×10^{-6}	3.194×10^{-7}	0.707
2nd derivative	2.061×10^{-7}	6.138×10^{-8}	0.247
SDA65			
Original	0.095	0.135	1.428
1st derivative	7.298×10^{-6}	4.482×10^{-6}	0.614
2nd derivative	3.271×10^{-6}	3.130×10^{-7}	0.096

Table: Within- and between-class variability for sequentially selected bands.

Bands	Within-class var	Between-class var	Ratio
Every 5th spectrum			
Original	1.114	1.006	0.902
1st derivative	6.060×10^{-5}	5.062×10^{-5}	0.835
2nd derivative	1.135×10^{-6}	5.649×10^{-7}	0.498
Every 9th spectrum			
Original	0.619	0.559	0.903
1st derivative	3.002×10^{-5}	2.636×10^{-5}	0.878
2nd derivative	1.755×10^{-7}	1.327×10^{-7}	0.756
Every 15th spectrum			
Original	0.371	0.335	0.902
1st derivative	1.658×10^{-5}	1.480×10^{-5}	0.893
2nd derivative	6.056×10^{-8}	5.305×10^{-8}	0.876
Every 19th spectrum			
Original	0.293	0.264	0.902
1st derivative	1.232×10^{-5}	1.113×10^{-5}	0.904
2nd derivative	3.592×10^{-8}	3.265×10^{-8}	0.909
Every 25th spectrum			
Original	0.223	0.201	0.902
1st derivative	8.837×10^{-6}	7.753×10^{-6}	0.877
2nd derivative	2.038×10^{-8}	1.702×10^{-8}	0.835

Table: Within- and between-class variability for spectrally degraded bands.

Bands	Within-class var	Between-class var	Ratio
Every 5th averaged			
Original	1.114	1.005	0.902
1st derivative	5.514×10^{-5}	4.819×10^{-5}	0.874
2nd derivative	5.582×10^{-7}	3.339×10^{-7}	0.598
Every 9th averaged			
Original	0.619	0.559	0.902
1st derivative	2.799×10^{-5}	2.511×10^{-5}	0.897
2nd derivative	1.260×10^{-7}	1.089×10^{-7}	0.864
Every 15th averaged			
Original	0.371	0.334	0.901
1st derivative	1.504×10^{-5}	1.356×10^{-5}	0.902
2nd derivative	4.713×10^{-8}	4.219×10^{-8}	0.895
Every 19th averaged			
Original	0.293	0.264	0.902
1st derivative	1.090×10^{-5}	9.880×10^{-6}	0.906
2nd derivative	2.737×10^{-8}	2.478×10^{-8}	0.905
Every 25th averaged			
Original	0.222	0.201	0.902
1st derivative	7.383×10^{-6}	6.591×10^{-6}	0.893
2nd derivative	1.399×10^{-8}	1.193×10^{-8}	0.853

Conclusions and future research

- For this data set, there are important bands from the original spectra, the first and second derivative spectra and from various regions of the EMS (VIS, NIR, SWIR) for species separability. We recommend further research and improvement to selecting spectral bands from a broader combined set of the original, first and second derivative spectra.
- There seem to be a number of sub-classes within each of the seven classes. We recommend further investigations to cluster spectra within each species by also incorporating their geographical location (spatially – Jun and Ghosh, 2009) and temporally when data is collected at different time.

Conclusions and future research

- Furthermore, for this data set, there does not seem to be any decrease in species separability by degrading the spectral bands through averaging the reflectance. This implies that hyperspectral (extremely high spectral) measurements did not prove useful in species separability compared to a lower spectral resolution data.

More details: Debba *et. al.* (2009). Within- and between-class variability of spectrally similar tree species. *In Proceedings of 2009 IEEE International Symposium on Geoscience and Remote Sensing*. July 13–17, 2009 Cape Town, South Africa. Accepted.

Further exploration

- We have explored two classification approaches with spectral angle mapper: (i) using a spectral library composed of one spectrum (endmember) per species and (ii) a multiple endmember approach conventionally called K-nearest neighbour classifier.

Data description

- Eight sites were chosen for the study including two sites in the KNP, two sites in private game reserves and four sites in communal lands.
- The species data used in this study consist of tree species generally more than 2 m tall identified and geo-registered using a Leica differential global positioning system (GPS).
- Eighteen dominant species are examined in the study. These include Acacia gerradii, Acacia nigrescens, Combretum apiculatum, Combretum collinum, Combretum hereroense, Combretum imberbe, Combretum zeyheri, Dichrostachys cinerea, Euclea sp (E. divinurum and E. natalensis, Gymnosporia sp (G. buxifolia and G. senegalensis), Lonchocarpus capassa, Peltoforum africanum, Piliostigma thonningii, Pterocarpus rotundifolia, Sclerocarya birrea, Strchynos sp (S. madagascariensis, S. usambarensis), Terminalia sericea and Ziziphus mucronata.

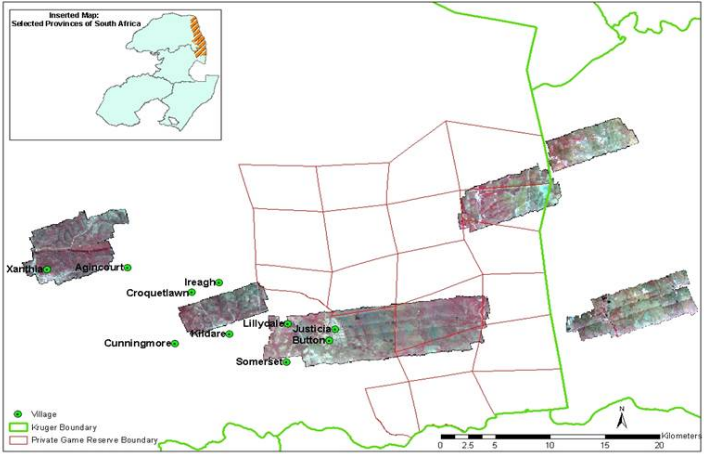

Figure: Study area showing Carnegie Airborne Observatory (CAO) image scenes in the Kruger National Park, South Africa

A

	AG	AN	CA	CC	CH	CI	CZ	DC	ED	GY	LC	PA	PT	PR	SB	SY	TS	ZM	sum	User accuracy(%)
AG	7	5	2	2	0	4	0	0	3	1	2	2	1	0	0	1	4	4	38	18.4
AN	0	1	1	0	1	1	0	1	1	0	2	0	0	0	0	0	0	1	9	11.1
CA	1	1	0	0	1	1	3	2	2	1	0	1	0	0	6	0	0	1	20	0.0
CC	11	3	0	12	0	3	0	4	3	2	6	2	5	0	8	1	2	3	65	18.5
CH	0	0	1	0	1	0	0	0	0	0	0	0	0	0	4	0	0	0	6	16.7
CI	2	3	1	0	1	3	0	1	1	0	1	0	0	0	6	0	0	0	19	15.8
CZ	0	0	1	0	0	3	15	0	0	0	0	1	0	1	12	1	4	0	38	39.5
DC	0	1	2	1	1	1	0	0	4	0	0	0	0	3	1	0	0	0	14	0.0
ED	4	1	0	2	0	1	1	3	6	0	3	0	0	0	8	5	1	2	37	16.2
GY	1	9	1	3	2	8	5	5	3	13	4	3	0	0	20	0	22	2	101	12.9
LC	1	0	0	3	2	2	5	2	0	0	1	0	0	0	1	1	9	0	27	3.7
PA	1	3	1	1	0	0	0	1	2	0	0	0	0	1	0	0	0	0	10	0.0
PT	0	0	0	18	0	0	0	0	13	2	2	2	11	0	12	1	3	0	64	17.2
PR	1	1	3	0	0	0	0	2	0	4	0	1	0	14	0	0	2	0	28	50.0
SB	0	2	0	1	1	0	1	0	2	3	2	0	0	0	1	0	2	1	16	6.3
SY	3	3	1	2	1	0	0	2	0	0	1	0	1	0	2	1	3	0	20	5.0
TS	0	0	0	0	0	1	0	0	1	0	0	0	0	0	2	0	0	0	4	0.0
ZM	1	4	0	1	1	0	0	1	2	1	0	0	0	0	3	1	1	1	17	5.9
sum	33	37	14	46	12	28	30	24	43	27	24	12	18	19	86	12	53	15	533	
Producer accuracy(%)	21.2	2.7	0.0	26.1	8.3	10.7	50.0	0.0	14.0	48.1	4.2	0.0	61.1	73.7	1.2	8.3	0.0	6.7		
Overall accuracy(%)	16																			

Figure: Using the mean spectra of the training sets as reference spectra.

B

	AG	AN	CA	CC	CH	CI	CZ	DC	ED	GY	LC	PA	PT	PR	SB	SY	TS	ZM	sum	User accuracy(%)
AG	12	4	1	2	1	2	0	1	0	1	0	0	0	4	2	1	1	0	32	37.5
AN	3	18	1	0	1	2	0	0	2	2	0	2	0	0	3	0	1	0	35	51.4
CA	1	0	4	0	1	0	2	0	1	0	0	0	0	0	0	0	0	0	9	44.4
CC	1	0	0	26	1	2	0	3	5	1	2	3	3	0	3	0	0	2	52	50.0
CH	0	1	0	1	2	0	1	1	0	0	0	0	0	0	1	0	2	0	9	22.2
CI	1	1	1	3	2	10	0	0	2	0	1	1	0	0	2	0	1	1	26	38.5
CZ	0	0	2	0	0	0	20	1	0	0	0	2	0	0	10	0	2	0	37	54.1
DC	5	1	0	1	0	3	0	6	3	0	0	0	0	0	1	3	3	0	26	23.1
ED	3	2	0	2	1	2	0	4	23	0	2	0	0	0	3	0	0	1	43	53.5
GY	0	2	0	0	0	0	0	1	1	14	0	0	0	0	1	0	0	1	20	70.0
LC	0	1	0	1	0	2	0	1	2	0	8	0	3	0	6	1	0	0	25	32.0
PA	1	0	0	0	0	1	1	0	0	1	0	3	0	0	0	0	0	0	7	42.9
PT	4	0	0	3	0	0	0	0	0	1	2	0	8	0	2	0	0	1	21	38.1
PR	0	1	0	0	0	0	0	0	0	1	0	0	0	12	0	0	0	2	16	75.0
SB	0	3	2	4	1	2	2	4	2	2	5	0	2	1	45	1	4	1	81	55.6
SY	1	2	0	2	2	1	0	0	0	0	1	1	2	0	2	3	0	0	17	17.6
TS	0	1	3	1	0	1	4	1	2	4	2	0	0	2	2	3	39	1	66	59.1
ZM	1	0	0	0	0	0	0	1	0	0	1	0	0	0	3	0	0	5	11	45.5
sum	33	37	14	46	12	28	30	24	43	27	24	12	18	19	86	12	53	15	533	
Producer accuracy(%)	36.4	48.6	28.6	56.5	16.7	35.7	66.7	25.0	53.5	51.9	33.3	25.0	44.4	63.2	52.3	25.0	73.6	33.3		
Overall accuracy(%)	48																			

Figure: Using all training spectra for each species as reference spectra (K-nearest neighbour classifier, k=1).

Conclusions

- Intra-species spectral variability and the reference data sample size are two important factors that affect tree species differentiation in the savanna ecosystem.
- We recommend the utilisation of the multiple endmembers SAM approach as opposed to the traditional SAM classifier involving single spectrum endmember per species for mapping of Kruger National Park species.
- The training endmembers should be truly representative of the different distributions in the population.
- The classification of the species could be limited to the dominant species.

More details: Cho *et. al.* (2009). Spectral variability within species and its effects on savanna tree species discrimination. *In Proceedings of 2009 IEEE International Symposium on Geoscience and Remote Sensing*. July 13–17, 2009 Cape Town, South Africa. Accepted.

Future research

- Feature selection and feature extraction (original, higher order derivatives)
- Spatial and temporal sub-classes.
- Improved classification techniques.
- Classification of hyperspectral images.
- SNR not constant throughout EMS.
- Classification on continuum removed spectra.