REGULARIZATION OF DISCRIMINANT ANALYSIS FOR THE STUDY OF BIODIVERSITY IN HUMID TROPICAL FORESTS

Jean-Baptiste FERET[1], Gregory P. ASNER[1], Stéphane JACQUEMOUD[2]

1 Department of Global Ecology, Carnegie Institution for Science, 260 Panama Street, Stanford, CA 94305, USA (jbferet@stanford.edu & gpa@stanford.edu)

2 Institut de physique du globe de Paris - Sorbonne Paris Cité, Université Paris Diderot, UMR CNRS 7154, Case 7011, 35 rue Hélène Brion, 75013 Paris, France (jacquemoud@ipgp.fr)

ABSTRACT

The performance of two supervised classifiers, linear and regularized discriminant analysis (LDA and RDA), is compared here for canopy species discrimination in humid tropical forest, based on airborne hyperspectral imagery acquired with the sensor Carnegie Airborne Observatory Alpha System (CAO-Alpha). Classification is performed to identify 13 species at pixel scale, crown scale, and using an object-based approach. The results show that for each scale of study, 70% to 75% overall accuracy is obtained with LDA. RDA allows improved classification for more than half species, and 5% increase of overall accuracy compared to LDA. The extended spectral range of the forthcoming CAO AToMS system (380-2500 nm) will allow for even more accurate classifications of tropical canopy species.

Index Terms— Image Classification, Humid Tropical Forests, Discriminant Analysis, Regularization, CAO

1. INTRODUCTION

Study and monitoring of biodiversity is a critically important contribution to both scientific and conservation efforts. Due to the spatially very extensive task, remote sensing appears as a particularly suitable tool. Hyperspectral sensors have proven useful in providing information to study biodiversity [1] and in performing accurate species classification with linear classifiers. Monitoring the high biodiversity found in humid tropical forests is particularly important but studies of these ecosystems have focused on just a few species [2,3]. Moreover, the lack of ground observations, used to train the classifier, can lead to serious inaccuracies. The size of this training dataset depends directly on the number of spectral features used for discrimination [4,5]. Consequently, the improved quality of hyperspectral sensors, providing more wavelengths of significant interest for classification, suggests an increasing need for ground observations, and methods have been employed to reduce the size of training datasets. Regularization of the linear classifier appears as an alternative to features selection, by introducing additional information in order to solve an ill-posed problem [6].

In this study, we compare the performance of two classifiers for species identification based on hyperspectral imagery: linear discriminant analysis (LDA) and regularized discriminant analysis (RDA). Thirteen species are discriminated with reflectance measurements acquired over Hawaiian tropical rainforest with the hyperspectral airborne sensor Carnegie Airborne Observatory Alpha (CAO-Alpha). This study is performed at the pixel scale, tree crown scale, and using reflectance averaged over one tree crown (object-based approach). The originality of this work lies in the number of species studied and the classifier applied to this type of discrimination.

2. MATERIAL & METHODS

2.1. Study site and hyperspectral imagery

This study was conducted at the Nanawale Forest Reserve, Hawaii. Thirteen common canopy species covering the high majority of the surface under study have been identified in our study site. Field work provided the location and identity of tree crowns at the top of the canopy.

Hyperspectral imagery was acquired with the CAO-Alpha in September 2007. The image used in this study is a 1980 by 1420 pixel image with 0.56 m spatial resolution, and the spectral information is measured on 24 spectral bands evenly spaced between 390 nm and 1044 nm. Atmospheric corrections were applied, as well as a mask based on LiDAR measurements to avoid pixels with vegetation less than one meter and pixels shaded by their neighbors.

2.2. Discriminant analysis

LDA already showed good performance for the identification of tropical species based on their optical properties but to date, the biological diversity investigated using airborne or spaceborne hyperspectral imagery never exceeded seven species [2,3]. However, very promising results were obtained, and this type of research needs to be conducted further in order to test the performance of classifiers with increasing biodiversity.

978-1-4577-2201-1/11/$26.00 ©2011 IEEE

The well-known problem called Hughes phenomenon is inherent to classification in hyperspectral imagery analysis, and occurs when the ratio between the number of samples collected for each class to identify is too small compared to the number of spectral features [4,7,8], leading to singularity of the covariance matrices corresponding to the different classes. Training samples can be very expensive to collect, particularly in humid tropical regions. This factor, combined with the technical progress observed for the development of high spectral resolution sensor makes Hughes phenomenon particularly problematic. The introduction of a regularization term in the class sample covariances is considered here to decrease the risk of occurrence of this phenomenon. Different algorithms exist to perform this regularization. Recent studies showed that RDA outperforms other linear classifiers and is competitive with nonlinear classifiers in situation of ill-posed hyperspectral image classification [7]. The algorithm used in our study [6] assesses two regularization parameters λ and γ. λ controls the degree of shrinkage of the individual class covariance matrix towards a pool estimate, whereas γ controls shrinkage towards a multiple of the identity matrix. LDA corresponds to the case ($\lambda = 1, \gamma = 0$) and the case ($\lambda = 0, \gamma = 0$) corresponds to quadratic discriminant analysis (QDA), where each class covariance matrix is different, contrary to LDA.

2.3. Training and validation

The training and validation datasets used for LDA and RDA are identical at each scale studied: pixel, tree crown, and object. At the pixel scale, the classification gives the posterior probability for the pixel to correspond to each of the studied species. The sum of all the posterior probabilities associated to a given pixel equals 1. Tree crowns were located during a field campaign, and the classification is based on the average posterior probability of the pixels within each of them. Object-based approach uses the mean reflectance spectrum of the pixels included in one tree crown. The classification is processed for this unique spectrum to identify each corresponding object.

The minimal sample size used to train a classifier is linked to the quantity of measured spectral features. In theory, the number of samples per class should be superior to the number of spectral features to avoid ill-posed problem. In practice, the appropriate number of training samples for each class is twice the number of spectral features for linear classifiers [5]. In order to meet conditions for ill-posed problem, 20 spectra were randomly selected from each species among the tree crowns located in the field, resulting in a training dataset of 260 spectra. When training the classifier, the weight of each pixel and the prior probability for each species were set to a unique value.

LDA was trained straightforwardly while the optimal (λ, γ) parameters were assessed prior to train RDA. Optimal values were investigated between 0 and 1, with 0.01 steps.

Pixel-based validation dataset was performed using 100 spectra randomly selected from each species, excluding the spectra already used for training. The classification of the whole image was also performed after the training stage. Tree-crown and object-based classification were performed on all the data available. A sigmoid filter was applied to each pixel probability before computing the tree crown posterior probability, in order to enhance the importance of pixels matching a species with more than 50% probability, and to lower the importance of the other ones.

2.4. Segmentation for automated tree crown location

Segmentation aims at providing a rough estimation of individual tree crowns (ITC). The segments delineated give spatial information, combined with spectral information derived from the pixel scale discriminant analysis. One species is assigned to each segment, adding coherence to the map. Segmentation of dense close canopy is very complex because of the absence of clearly defined edges between trees, but the mean shift clustering method [9] applied to a 3-channels subimage performed better than other algorithms. A slight over-segmentation was observed, meaning that one tree crown is cut into several segments. This situation is preferred to situations where several species are included into one segment. Once the segmentation is performed, the posterior probability was computed for each segment, following the same method as for tree-crown study.

3. RESULTS & DISCUSSION

The results of the classification using LDA and RDA in terms of user's accuracy, producer's accuracy for each species and overall accuracy (OA) are compared. The optimization of the regularization parameters are the same for all scales: $\lambda = 0.14$, and $\gamma = 0$. RDA is then closer to QDA than LDA. However, QDA performed very poorly due to the lack of training data (results not shown).

3.1. Pixel scale

The results for pixel-scale classification are summarized in Table 1. Performance strongly varies depending on the species. The producer's accuracy obtained with LDA is inferior to 70% on four species, and superior to 85 % on five. User's accuracy shows that the proportion of false positives is important for five species. OA is close to 75%, which can be compared to the results obtained by [2] when selecting 20 to 30 spectral features between 400 nm and

2500 nm. These results are satisfactory for different reasons: first, our classification takes into account twice as much species as in [2], moreover the spectral domain used does not include short wave infrared (SWIR), and finally the size of the training dataset is much smaller than [2] (20 samples per species to compare to 300 samples per species).

When applying RDA, the producer's accuracy is improved of 5% to 22% for more than half of the species, and equivalent for three species already well identified by LDA. The OA is close to 80%: RDA performs better than LDA, but the latter shows very decent results even though the limited size of the training dataset. RDA clearly outperforms LDA for even smaller size of training samples, but the OA consequently decreases (e.g. using 3 samples per class, OA = 25% for LDA, and OA = 41% for RDA).

Table 1: Pixel-scale performance of LDA and RDA.

Species	LDA		RDA	
	Producer	User	Producer	User
common guava	37.0	58.7	51.0	68.0
strawberry guava	54.0	80.6	76.0	77.6
kukui	89.0	93.7	90.0	92.8
mango	88.0	76.5	93.0	80.2
gunpowder	79.0	79.8	71.0	93.4
eucalyptus	95.0	93.1	94.0	99.0
ohia	80.0	88.9	91.0	83.5
coconut	75.9	98.4	84.34	100.0
avocado	88.0	83.8	82.0	90.1
monkeypod	100.0	84.8	100.0	94.3
cecropia	70.0	49.3	64.0	53.8
hala	61.0	46.9	74.0	53.6
melochia	41.0	44.1	51.0	54.8
Overall accuracy	**74.7**		**79.6**	

3.2. Tree-crown scale and Object-based approach

Tree crown scale and object-oriented approach share the same analysis because no significant differences were obtained between these two methods, contrary to the results of [2]. Table 2 summarizes the results obtained specifically for tree crown scale classification. Some common points with Table 1 can be noticed: three species show a low producer's accuracy at both scales. However, at the crown scale, the producer's accuracy associated to melochia is good, suggesting that the pixels successfully identified as this species are clearly identified (more than 50% posterior probability), because of the sigmoid filter. However, the user's accuracy is very low, showing that tree crowns were wrongly identified as melochia. On the contrary, the successful identification for gunpowder decreased at the tree crown scale, suggesting that the posterior probability associated to successfully identified pixel of these species is not clearly superior to the posterior probability associated with the other species. Shifting the sigmoid filter can be applied in order to enhance posterior probabilities lower than 50% and improve species identification.

For both tree crown scale and object-based approach, the OA decreased compared to the identification at the pixel scale, contrary to the results of [2], but the improvement of the OA induced by RDA compared to LDA is the same order of value, about 5%. However, as for pixel scale identification, the improvements allowed by the study of the whole spectral domain let us expect a dramatic improvement of species identification if including SWIR for classification.

Table 2: Tree-crown scale performance of LDA and RDA.

Species	LDA		RDA	
	Producer	User	Producer	User
common guava	20.9	19.2	30.2	24.1
strawberry guava	56.4	73.8	76.4	72.4
kukui	74.8	99.2	86.2	99.3
mango	91.4	96.5	90.1	96.5
gunpowder	56.5	74.3	58.7	87.1
eucalyptus	95.5	87.5	95.5	100.0
ohia	73.9	70.8	87.0	69.0
coconut	100.0	81.8	100.0	100.0
avocado	100.0	66.7	100.0	66.7
monkeypod	100.0	10.0	100.0	33.3
cecropia	69.1	70.2	66.9	76.5
hala	47.5	36.5	65.0	47.3
melochia	83.3	10.2	66.7	12.1
Overall accuracy	**70.9**		**76.3**	

3.4. Combining spectral and spatial information

The comparison between LDA and RDA showed the better performance of RDA. In order to obtain a map showing the spatial distribution of all the species on the whole map, RDA was applied at the pixel scale. Each polygon produced by the segmentation (see section 2.5) was considered as a tree crown, and a crown scale posterior probability was computed. The initial image is presented in Figure 1 (left), and includes region of interest (ROIs) corresponding to the different species. Figure 1 (right) shows the result of RDA combined with mean shift clustering. The agreement between ground truth and species classification is very good, and this validation with ground observations will continue.

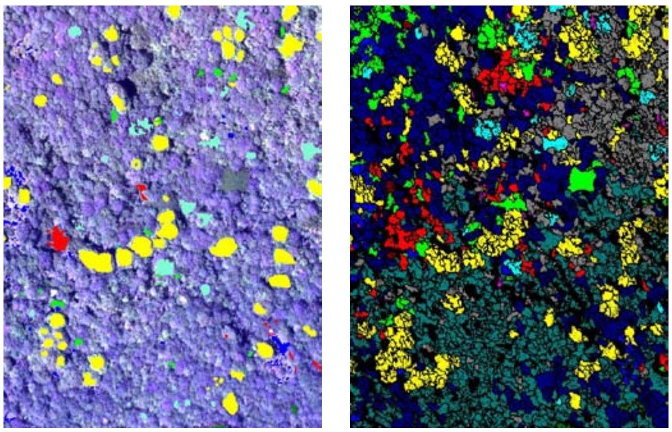

Figure 1: subimage acquired by CAO-Alpha over Nanawale Forest Reserve (HI). ROIs show ground data (left); Species distribution obtained using RDA combined with segmentation (right).

4. CONCLUSION

This study demonstrated that species discrimination based on hyperspectral remote sensing can be successfully applied to humid tropical forests with, to date, unequalled biodiversity, as we obtained good results for the classification of 13 species. RDA proved to perform better than LDA after training with a small dataset, at each scale studied, and its advantage could appear more clearly if more spectral features and species are included in the analysis.

Very promising methods for supervised and unsupervised classification have been developed recently, combining full spectral information with spatial information [10,11]. Their performance for this type of classification will also be studied soon.

The conditions of our study are not optimal because of the spectral domain investigated, which means that using a full range sensor may dramatically improve species discrimination. AToMS, the next generation of full range hyperspectral airborne sensors developed to study humid tropical rainforests, is planned to be operational in 2011, and will allow analysis of images with higher biodiversity.

5. REFERENCES

[1] K.C. Carlson, G.P. Asner, R.F. Hughes, R. Ostertag, and R.E. Martin, "Hyperspectral remote sensing of canopy biodiversity in Hawaiian lowland rainforests", *Ecosystems*, vol. 10, pp. 536–49, 2007

[2] M.L. Clark, D.A. Roberts, and D.B. Clark, "Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales". *Remote Sensing of Environment*, vol. 96, pp. 375−398, 2005.

[3] M. Papeş, R. Tupayachi, P. Martínez, A.T. Peterson, and G.V.N. Powell, "Using hyperspectral satellite imagery for regional inventories: a test with tropical emergent trees in the Amazon Basin", *Journal of Vegetation Science*, vol. 21, pp. 342–354, 2010.

[4] G.F. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Transactions on information theory*, vol. IT-14, pp. 55–63, 1968.

[5] T.G. Van Niel, T.R. McVicar, B. Datt, "On the relationship between training sample size and data dimensionality: Monte Carlo analysis of broadband multi-temporal classification", *Remote Sensing of Environment*, vol. 98, pp. 468–480, 2005.

[6] J. Friedman, "Regularized discriminant analysis", *Journal of the American Statistical Association*, vol. 84, pp. 165–175, Mar. 1989

[7] T.V. Bandos, L. Bruzzone and G. Camps-Valls, "Classification of Hyperspectral Images With Regularized Linear Discriminant Analysis", *IEEE Transactions on geosciences and remote sensing*, vol. 47, pp. 862-873, 2009.

[8] Landgrebe, D.A., *Signal theory methods in multispectral remote sensing*. Hoboken, New Jersey, 2003.

[9] D. Comaniciu, and P. Meer, "Mean shift: a robust approach toward feature space analysis", *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 5, 2002

[10] N. Gorretta, J.M. Roger, G. Rabatel, V. Bellon-Maurel, C. Fiorio, and C. Lelong, "Hyperspectral image segmentation: the butterfly approach", *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing,* France, 2009.

[11] Y. Tarabalka, M. Fauvel, J. Chanussot, J. A. Benediktsson, "SVM and MRF-based method for accurate classification of hyperspectral images," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, pp. 736-740, 2010.