

MAPPING AND SPATIAL CHARACTERIZATION OF NONNATIVE GRASSES IN THE BIG ISLAND, HAWAII USING HYPERSPECTRAL IMAGERY

Sunyurp Park, Assistant Professor
Department of Geography and Environmental Studies
University of Hawaii-Hilo, Hilo HI 96720
sunypark@hawaii.edu

ABSTRACT

A cross-sensor (hyperspectral and high-resolution data sets) hybrid approach was used to map grass species in the coastal lowland area of the Hawaii Volcanoes National Park. AVIRIS imagery was selected for hyperspectral data and its 20-meter resolution was compensated with IKONOS 1-meter resolution data. Three main native and nonnative grass species were focused in the study, including broomsedge (*Andropogon virginicus*), natal redtop (*Melinis repens*), and pili grass (*Heteropogon contortus*). A 3-step, hybrid approach, combining an unsupervised and a supervised classification schemes, was applied to grass mapping. First, the IKONOS 1-m high-resolution data were classified to create a binary image (vegetated vs. non-vegetated) and converted to 20-meter resolution percent cover vegetation data to match AVIRIS data pixels. Second, the minimum noise fraction (MNF) transformation was used to extract a coherent dimensionality from the original AVIRIS data. Since the grasses were sparsely distributed and most image pixels were intermingled with lava surfaces, the reflectance component of lava was filtered out with a binary fractional cover analysis assuming that the total reflectance of a pixel was a linear combination of the reflectance spectra of vegetation and the lava surface. Finally, a supervised approach was used to classify the grass species based on the maximum likelihood algorithm. The classification result showed that there was much confusion between the grasses, especially between broomsedge natal redtop. Knowing that there was co-occurrence of one or more grass species, more accurate sampling schemes and additional phenology characteristics of the species would be needed to better define training sites.

INTRODUCTION

Vegetation mapping and classification have been conducted successfully using various types of remotely sensed data (Schmidt and Skidmore 2003; Schmidtlein and Sassini 2004; Ustin and Xiao 2001). Multi-spectral remote sensing has been the most common approach in estimating land cover types. Although most vegetation mapping studies reported reasonably high classification accuracy at general plant-type levels, such as biomes, life zones, and life forms, species-level mapping effort has been challenging due to limitations of sensor characteristics. Limitations in the number of bands and the number of temporal data sets and their spatial resolving power, or spatial resolution have restricted the level of vegetation classification.

In addition, cloud contamination makes it difficult for remote sensing communities to map vegetation in Hawaii. It is not unusual to have less than 3 different temporal scenes for a given area throughout the entire year. This problem significantly reduces the multi-temporal capability of most multi-spectral remote sensors. As a result, knowing that species-level vegetation mapping needs spectral information from as many spectral bands as possible, it requires hyperspectral remote sensing data to statistically separate plant species from one another. The NASA Airborne Visible and Infrared Imaging Spectrometer (AVIRIS) is one of a few hyperspectral remote sensing systems available to the scientific community. This system acquires optical radiance data in 224 contiguous wavelengths ranging from 400 to 2500nm and provides high quality signature measurement of solar radiation reflected from the surface (Green et al. 1998). Since the first measurement of the system in 1987, its applications have been widely spread into diverse scientific fields (Asner et al. 2005; Goetz et al. 1985; Li et al. 2005; Okin et al. 2001)

Recently, groups of researchers have used aircraft-based hyperspectral approaches in species-level vegetation mapping (Gong et al. 1997; Hirano et al. 2003; Schmidt and Skidmore 2001; Schmidt and Skidmore 2003; Underwood et al. 2003). Previous study results showed that hyperspectral applications could be promising in detailed vegetation classification and our ability to map the extent of invasive species and their spatial distribution would be improved. A sub-pixel fractional cover analysis, called linear spectral unmixing or spectral mixture analysis (SMA), has been commonly used in image classification (Adams et al. 1995; Roberts et al. 1993; Smith et al. 1990). Although SMA is a useful and theoretically appealing approach for sub-pixel image classification

breaking down mixture pixels into the major components, its usage can be questionable if the surface is extremely heterogeneous, atmospheric conditions change frequently, topography is highly rugged, image endmembers do not exist, or pure reference samples cannot be collected at the same time of image acquisition. In fact, it is often difficult to identify necessary, pure endmembers in an image, especially on sparsely vegetated surfaces. One or more of these conditions has made it difficult to perform detailed vegetation classification in Hawaii. Integrating image data from two different sensors, AVIRIS and IKONOS, this study investigated the potential of a binary fractional cover analysis for mapping nonnative grass species vegetated on lava surfaces.

METHODOLOGY

Study Area

The study area is located at the southeastern edge of Big Island of Hawaii, which is the youngest island of the Hawaiian archipelago. This area is part of the eastern coastal lowlands within the Hawaii Volcanoes National Park (HAVO), and it occupies an area between the coastline and the base of the major fault scarps (Figure 1). Elevations of the study area range approximately from 20 to 90 meters above the sea level. The climate is seasonal with hot, dry summers and wet winters. Mean annual temperature is 23-26 °C and annual precipitation is usually less than 1,000 mm (Giambelluca and Schroeder 1998; Wegner et al. 1990). Most native and nonnative grasses are found on thin soils and old lava flows. The substrates of the Big Island are derived from one of five volcanoes: Kohala, Mauna Kea, Hualalai, Mauna Loa, and Kilauea. The entire island is dissected by a various lava flows of different ages, and this mosaic of diverse substrates spawned wide gradients of soil and ecosystem development (Vitousek et al. 1992).

The study area is surrounded by lava flows in 1967-74 and 1992-2002 and sparsely vegetated primarily by native and nonnative grasses, shrubs, ferns, and occasional Ohia trees. Pili grass (*Heteropogon contortus*) is the most common native grass, and major nonnative grasses include broomsedge (*Andropogon virginicus*), and natal redtop (*Melinis repens*). The range of broomsedge has increased significantly in HAVO as a result of volcanoes-derived or human-caused fires (Wegner et al. 1990). Among the most common herbs and shrubs are sleeping grass (*Mimosa pudica*), lantana (*Lantana camara*), pukiawe (*Styphelia tameiameia*), a'ali'i (*Dodonaea viscosa*), and uhaloa (*Watherria indica*). A'ali'i, pukeawe, and uhaloa are native species while sleeping grass and lantana are introduced species.

Data Sets and Pre-processing

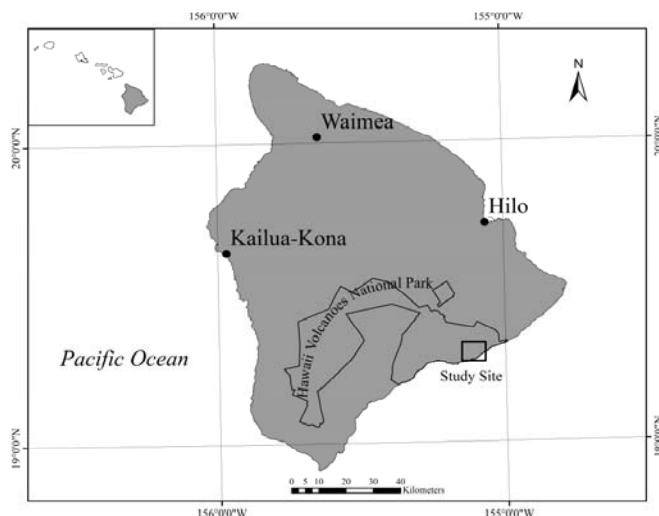


Figure 1. Study area. It is located in the Hawaii Volcanoes National Park (HAVO).

AVIRIS data covering the study area were obtained from Jet Propulsion Laboratory (JPL), NASA and the image was acquired on October 30, 2001. The sensor was flown aboard the NASA ER-2 aircraft at 20 km above ground level and the typical instantaneous field of view (IFOV) of 1.0 mrad had a 20 20 m ground resolution. AVIRIS acquires images in 224 contiguous bands from 400nm to 2500nm with 10 nm intervals (Green et al. 1998). Very-high-resolution IKONOS data (Space Imaging, Inc.) were provided by the Hawaii IKONOS Consortium formed by the Hawaii Natural Heritage Program (HNHP) at the University of Hawaii. The data consisted of a 1-meter pan-sharpened color and 4-meter multispectral imagery bundle.

For radiometric calibration, a dark-object subtraction method was used to the IKONOS data. This method was believed to be ideal because the study area's substrates are virtually dark objects and patches of the lava flows were shadowed by cloud cover. Therefore, the pixel value of the shadow area was simply subtracted from the other pixels' values for each band. Atmospheric correction is even more important for hyperspectral data. The Fast Line-

of-sight Atmospheric Analysis of Spectral Hypercubes (FLAASH) was applied to the AVIRIS data. This algorithm is based on the MODTRAN code and available as a plug-in to the ENVI image processing software (Research systems, Inc.). The correction process calculates the amount of molecular scattering and absorption present, which is primarily caused by seven gases including water vapor (H₂O), carbon dioxide (CO₂), ozone (O₃), nitrous oxide (N₂O), methane (CH₄), carbon monoxide (CO), and oxygen (O₂), and remove the effects of the gases from each pixel. Since the AVIRIS product was geometrically distorted, it needed to be rectified using the georectified IKONOS image. For this process, ground control points were selected throughout the imagery for mathematical transformation, and the input image pixels were transformed to UTM Zone 5 projection with the root-mean-square error (RMSE) of 0.38.

Image Classification

Field sample collection. For ground-truth sampling, a comprehensive field survey was conducted during the summer (June 8-17) in 2005. A systematic transect sampling method was used. Fifteen sampling paths were used and these paths were spaced with a fixed interval of 250 m except two paths (500 m). Three to fifteen transects were surveyed along each sampling path with an interval of 200 m, depending on the lengths of the sampling paths. This sampling design generated 133 individual transects in total throughout the study area. Each transect was 10 meters long, and plant species or a substrate type that touched a transect line was recorded every 20 cm along the transect. Supplementary information such as soil pH, soil color, and soil moisture content (%) was also obtained if it was available. For accurate species identification, the field crew consulted local botanists, visiting a herbarium of the USGS Biological Resources Division in the Hawaii Volcanoes National Park before and after the survey. Locational information for the transects was acquired using handheld GPS units (Garmin Rino 130 handheld/2-way radio and GPS). In situ reflectance spectra of the major plant species described above were also acquired in the field using a portable spectroradiometer (Fieldspec Pro Jr., ASD). Data were acquired separately on November, 2 in 2005, which was an approximate anniversary date of the AVIRIS imagery acquisition date. The spectroradiometer measures reflectance values from 350 to 2,500 nm. Reflectance spectra were repetitively collected 10 times and averaged for each species.

Unsupervised classification. Visual examination of the IKONOS imagery showed that 4-meter multispectral data were still too coarse to identify clumps of plants distributed across the study area. Therefore, only 1-meter pan-sharpened color imagery was used to conduct an unsupervised classification using the Erdas Imagine 8.7 software package (Leica Geosystems, GIS and Mapping, LLC). Vegetated- and non-vegetated pixels were separated from each other successfully after the classification. These resulting binary (vegetated or non-vegetated pixels) 1-meter data were degraded to 20-meter data so that spatial resolution of the classified imagery could match that of the AVIRIS imagery. Since 20 20 pixels of the 1-meter classified data were collapsed to one 20 20 meter pixel, the percentage of vegetated area per each 20-meter pixel was conveniently computed.

Minimum noise fraction transformation. The minimum noise fraction (MNF) transformation is two cascaded principal components analysis to separate noise from the sensor signal, to compress the spectral information into a few components, and determine the inherent dimensionality of image data for subsequent processing (Green et al. 1998). MNF transformation is implemented in ENVI software package, and a standard processing technique was applied to the AVIRIS imagery. An eigenvalue is a statistical measure that evaluates the dimensionality of the data, and normally the first few MNF components explain most of the variance of the data. In the case of AVIRIS data, the process significantly reduces the number of bands or components that need to be further processed typically to < 20 useful MNF bands. MNF bands that contain meaningful information have an eigenvalue an order of magnitude greater than those that contain mostly noise. ENVI provides both eigenvalues and output MNF eigenimages for determination of data dimensionality. Generally, the more spatially coherent (or continuous) the image, the less noise and greater the information content (Jensen 2005). For MNF band selection, the eigenvalues of MNF bands were plotted and examined in this study. The eigenvalues dramatically as MNF band number increased, and no significant change was observed after MNF band 6. Therefore, the first 6 MNF bands were used for following image classification (Figure 2).

Binary fractional cover analysis and supervised classification. Most of the AVIRIS image pixels contained reflectance signals from more than one material on the ground. To extract only the reflectance component of vegetation from the mixed signal of each pixel, it was assumed that the hyperspectral reflectance of a pixel was a linear combination of the reflectance spectra of vegetation and the lava surface. To define and filter out the reflectance component of lava, "pure" lava pixels were identified from the unsupervised classification result and their digital values from each of the 6 MNF bands were computed. To simply state, the proportion of lava's contribution to a MNF band was subtracted from each pixel (DN_T) of that MNF band to extract the component of vegetation in each MNF band. The proportion of lava's contribution per pixel for each MNF band was calculated by

multiplying per cent cover lava (1 - % vegetation) and the mean of the pure lava pixels (DN_L) of each MNF band. Since the term calculated so far ($DN_L (1 - \% \text{ vegetation})$) was biased by the area of vegetation cover in each pixel, the vegetation component of each pixel of MNF bands needed to be normalized for the pixel area ($20 * 20 = 400 \text{ m}^2$) by inversely weighting the per cent vegetation cover. Therefore, a vegetation component in each pixel was obtained

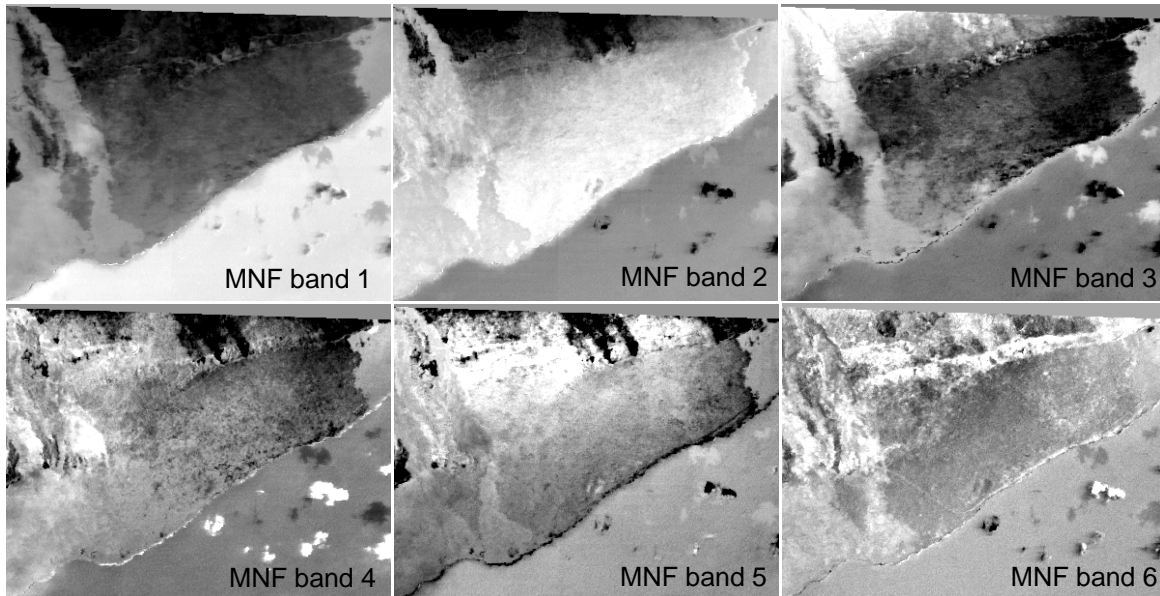


Figure 4. The first 6 MNF eigenimages created by ENVI. These images are spatially coherent and have the highest eigenvalues of all with the least noise.

applying the following formula to each MNF band:

$$DN_v = \frac{DN_T - DN_L (1 - \% \text{ vegetation})}{\% \text{ vegetation}}$$

These vegetation component MNF bands were used as input to supervised classification. Randomly selected, half of the ground samples were used for training sites, and these samples were categorized into one of plant classes, which were determined by a dominant plant species. As a result, each ground sample was classified as broomsedge, natal redtop, pili, shrub, or mixed class based on the plant species identification survey result. To reduce classification errors, only the area that is dominated by grasses was subset from the vegetation component MNF data before supervised classification was performed. The maximum likelihood classifier algorithm was used for image classification based on the training samples.

RESULTS & DISCUSSION

Classification accuracy evaluation was conducted using randomly selected field samples. A typical error matrix was created based on 5 major classes.

Spatial Distribution of Grasses

Per cent cover vegetation derived from IKONOS 1-m resolution data showed the overall distribution of plants well in the study area. Vegetated and lava surfaces were easily distinguishable from each other, and detailed examination of the classification result with comparison to the IKONOS imagery confirmed that the separation of vegetated pixels from lava pixels was successful. This binary data were converted to per cent cover vegetation data

degrading the spatial resolution to with 20-m (Figure 3). The per cent cover vegetation increased as it went from the southern coastal line to higher elevation. Most of the area was dominated by per cent cover of 60% or less (Table 1). Only 9% of the pixels had per cent cover of 60% or more. Generally, per cent cover increased with elevation, but the direction of per cent cover did not precisely correspond with the direction of elevation. The direction of per cent cover change was north-south, while elevation contour lines were roughly parallel to the coast line, which runs southwest to northeast. Ages of the substrates and strong wind effects might have influenced this spatial pattern.

The final classification result is shown in Figure 4. Overall, broomsedge was the most dominant species with widely spread pattern across the area. Although there was no particular spatial structure of the species' distribution, higher concentration of pili grass, the native species, was observed in the northeastern part of the area. Another noticeable feature was that shrubs occupied central portion of the area while natal redtop relatively heavily populated along the boundary area or closer to the road. Possible reasons for this outcome stay unexplained, and it is difficult to analyze the observed patterns because the resulting map is not accurate enough for further spatial analyses. However, it is reasonable to state that the two nonnative species, broomsedge and natal redtop, were the prevalent species in the coastal lowlands area in the park based on the preliminary image classification and the error matrix figures.

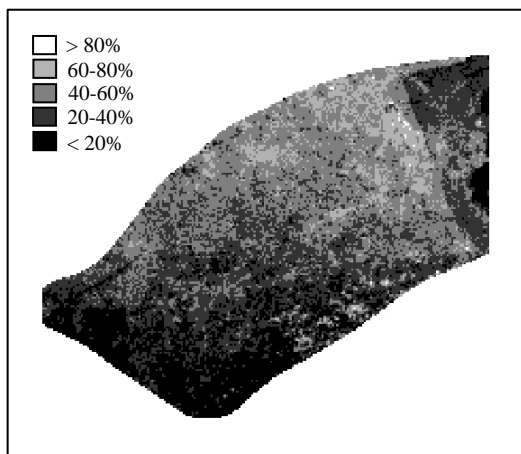


Figure 3. Per cent cover vegetation. Lavas are the main background substrates in the area. Vegetated and non-vegetated surfaces were separated by unsupervised classification of high-resolution IKONOS data.

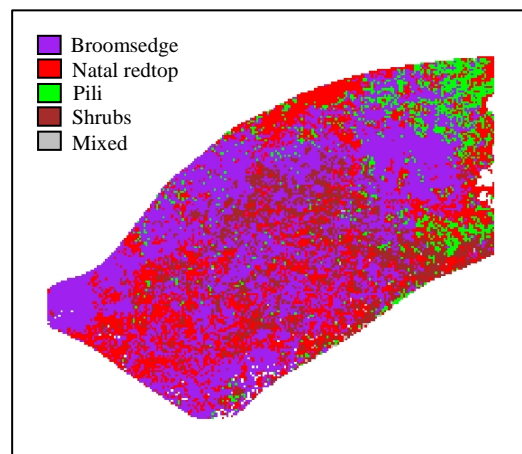


Figure 4. The result of unsupervised classification. Each class was defined by a dominant plant species

Table 1. Categorization of per cent cover vegetation derived from IKONOS 1-m resolution.

Per cent cover (%)	Number of pixels	Area (ha)	Proportion (%)	Cumulative (%)
0-20	4,298	171.9	20.5	20.5
20-40	7,363	294.5	35.1	55.6
40-60	7,449	298.0	35.5	91.1
60-80	1,837	73.5	8.8	99.9
80-100	28	1.1	0.1	100.0
Total	20,975	839.0	100.0	

Classification Accuracy

An error matrix is a simple array of numbers laid out in rows and columns that express the number of image pixels assigned to a particular classification category relative to the actual category as verified in the field. The agreement between the field sample data and the classified image is shown as a co-occurrence table (Table 2). As reported in the table, the three different grass species were not clearly differentiated from each other in the classification process. Especially, broomsedge was significantly confused with Natal redtop. Heterogeneous classes such as shrubs and mixed plants were substantially confused with grass species. As a result, classification accuracy

for any of the grass species was no higher than 57%, and those of shrubs and mixed classes were lower than 30%.

Table 2. Producer's, user's, and overall accuracy measures for the five classification categories.

	Broomsedge	Natal redtop	Pili	Shrub	Mixed	Sum	Producer's accuracy
Broomsedge	104	36	11	31	1	183	56.8
Natal redtop	45	74	9	16	0	144	51.4
Pili	36	11	20	5	0	72	27.8
Shrub	30	24	1	19	1	75	25.3
Mixed	9	9	4	9	3	34	8.8
Sum	224	154	45	80	5	508	
User's accuracy	46.4	48.1	44.4	23.8	60.0		Overall accuracy=43.3

Field survey archived in the herbarium at HAVO and our field identification both indicated that the native grass (pili grass, *Heteropogon contortus*) grew with the nonnative grasses (broomsedge, *Andropogon virginicus* and natal redtop, *Melinis repens*) throughout the study area. Observing that there was significant amount of confusion between the three grass species, it is believed that the likeliness of co-occurrence of two or more different species has made the separation of the three species difficult.

Pure, dark lavas might have had a strong effect on the classification process because it is believed that the contribution of lava surfaces to pixel reflectance values must have dampened unique signals from the grass species. Especially, for those pixels that were dominated by lava, the reflectance components of the grasses might have been weakened or removed when the component of the lava surface was subtracted from each pixel. For this study, training sites were selected from pixels that had per cent cover of 25% or more. Better-defined training site selection with higher percent cover may improve accuracy of the classification reducing the potential influence of lavas.

Another factor for low classification accuracy may be the lack of uniqueness of the reflectance characteristics among the grass species. The field-measured reflectance curves of the three grass species were not dramatically different from each other. Although any detailed study of the phenology of these grass species has not been conducted in this area, local botanists advised that seasonal changes of their lives might be more noticeable than annual changes (personal communication). The structures of leaf blades and stems are unique for grass species, but their inflorescences and flowers are more characteristic. Therefore, seasonal reflectance characteristics of the species are expected to provide phenological information to image classification process.

Other Physical Factors

There was no substantial variation of soil acidity throughout the area. The range of soil pH was from 5.4 to 6.8, and most observations were between 6.0 and 6.8. This result indicates that soil acidity was fairly close to neutral and did not play a major role in the species' establishment and spread. Most soils were underdeveloped and did not represent any variation in soil color. Soil color determination based on Munsell soil color charts was primarily 2.5Y/2.5/1 (hue/value/chroma).

Soil moisture content showed differences between broomsedge, natal redtop, and pili grass-occupied areas. It is known that native species, such as pili grass, can establish more successfully on thin soils or harsh old lava flows because they have been adapted to the local environment. Sampling sites, where pili grass dominated, had the lowest mean soil moisture content, 30%. The two nonnative species dominated on two contrasting moisture environments. Soil moisture content of broomsedge-dominating sites was higher (35.7%), while that of natal redtop-dominating sites was lower (26.3%), respectively. The relationship between soil moisture and the distribution of these species is uncertain in this study because most alien species are strongly fire-adapted, or fire-tolerant and seem to replace the native species in this dry environment.

CONCLUSIONS

Species-level plant classification is still very challenging work even with hyperspectral data. Reflectance signals of sparsely vegetated short grasses with dark background substrates such as lavas may not be strong enough to overcome the impurity of mixed pixels with 20-meter spatial resolution of AVIRIS data. Although mixed plants or shrubs were not the focus of this study, their classification accuracy result showed that pixels occupied by multiple plant species intermingled with lava background, such as shrubs and mixed classes, were not classifiable with the approach used in the study. It is encouraging, however, to know that grass classes dominated by a native or a

nonnative species were classified with much higher accuracy even though the accuracy figure was still low only reaching up to 57%. Higher-quality field sampling with more transects, or multi-directional transects, and more precise locational data attached to field sampling sites will reduce classification errors and improve grassland classification. Obviously, the usage of multi-seasonal data will better define the reflectance characteristics of different species, and therefore increase the capability of image classification. If this classification approach turns out to be useful and more successful in the future, sparsely distributed plants with relatively pure background, such as species on dry soils or hydrophytes floating on water bodies, can be targeted for species-level image classification.

REFERENCES

- Adams, J.B., D.E. Sabol, V. Kapos, R.A. Filho, D.A. Roberts, M.O. Smith, and A.R. Filledspie (1995). Classification of multispectral images based on fractions of endmembers: application to land-cover change in the Brazilian Amazon. *Remote Sensing of Environment*, 52: 137-154.
- Asner, G.P., A.J. Elmore, R.F. Hughes, A.S. Warner, and P.M. Vitousek (2005). Ecosystem structure along bioclimatic gradients in Hawaii from imaging spectroscopy. *Remote Sensing of Environment*, 96: 497-508.
- Giambelluca, T.W. and T.A. Schroeder (1998). Climate. In: S.P. Juvik and J.O. Juvik (Eds.) *Atlas of Hawaii*, pp.49-59. University of Hawaii Press, Honolulu, Hawaii.
- Goetz, A.F.H., G. Vane, J.E. Solomon, and B.N. Rock (1985). Imaging spectrometry for earth remote sensing. *Science*, 228: 1147-1153.
- Gong, P., R. Pu, and B. Yu (1997). Conifer species recognition: an exploratory analysis of in situ hyperspectral data. *Remote Sensing of Environment*, 62: 189-200.
- Green, R.O., M. Berman, P. Switzer, and M.D. Craig (1998). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geosciences and Remote Sensing*, 26: 65-74.
- Green, R.O., M.L. Eastwood, C.M. Sarture, T.G. Chrien, M. Aronsson, B.J. Chippendale, J.A. Faust, B.E. Pavri, C.J. Chovit, M. Solis, M.R. Olah, and O. Williams (1998). Imaging spectroscopy and the airborne Visible/Infrared Imaging Spectrometer (AVIRIS). *Remote Sensing of Environment*, 65: 227-248.
- Hirano, A., M. Madden, and R. Welch (2003). Hyperspectral image data for mapping wetland vegetation. *Wetlands*, 23: 436-448.
- Jensen, J.R. (2005). *Introductory digital image processing: a remote sensing perspective*. Pearson Prentice Hall, Upper Saddle River, NJ, pp.444.
- Li, L., S.L. Ustin, and M. Lay (2005). Application of AVIRIS data in detection of oil-induced vegetation stress and cover change at Jornada, New Mexico. *Remote Sensing of Environment*, 94: 1-16.
- Okin, G.S., D.A. Roberts, B. Murray, and W.J. Okin (2001). Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments. *Remote Sensing of Environment*, 77: 212-225.
- Roerts, D.A., M.O. Smith, and J.B. Adams (1993). Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data. *Remote Sensing of Environment*, 44: 255-269.
- Schmidt, K.S. and A.K. Skidmore (2001). Exploring spectral discrimination of grass species in African rangelands. *International Journal of Remote Sensing*, 22: 3421-3434.
- Schmidt, K.S. and A.K. Skidmore (2003). Spectral discrimination of vegetation types in a coastal wetland. *Remote Sensing of Environment*, 85: 92-108. *Remote Sensing of Environment*, 85: 92-108.
- Schmidtlein, S and J. Sassin (2004). Mapping of continuous floristic gradient in grasslands using hyperspectral imagery. *Remote Sensing of Environment*, 92: 126-138.
- Smith, M.O., S.L. Ustin, J.B. Adams, and A.R. Gillespie (1990). Vegetation in deserts: I. A regional measure of abundance from multispectral images. *Remote Sensing of Environment*, 31: 1-26.
- Underwood, E., U. Susan, and D. DiPietro (2003). Mapping nonnative plants using hyperspectral imagery. *Remote Sensing of Environment*, 86: 150-161.
- Ustin, S.L. and Q.F. Xiao (2001). Mapping successional boreal forests in interior central Alaska. *International Journal of Remote Sensing*, 22: 1779-1797.
- Vitousek, P.M., G. Aplet, and D. Turner (1992). The Mauna Loa environmental matrix: foliar and soil nutrients. *Oecologia*, 89: 372-382.
- Wagner, W.L., D.R. Herbst, and S.H. Sohmer (1990). *Manual of the flowering plants of Hawaii*. Vol.1. University of Hawaii Press, Honolulu, HI.