APPLICATION OF COMBINED PIXEL-BASED AND SPATIAL-BASED APPROACHES FOR IMPROVED MIXED VEGETATION CLASSIFICATION USING IKONOS

Atiek Widayati*
Bruno Verbist*
Allard Meijerink**

*International Centre for Research in Agroforestry - Southeast Asia Regional Office (ICRAF SEA), PO Box 161, Bogor, Indonesia , Tel.: 62-251-625 415; Fax.625 416,
Email: a.widayati@cgiar.org, b.verbist@cgiar.org

**International Institute for Geo-Information Science and Earth Observation (ITC), P.O. Box 6, 7500 AA Enschede, The Netherlands, Tel: +31 (0)53 4874 444; Fax: +31 (0)53 4874 400;
E-mail : meijerink@itc.nl

KEY WORDS: *pixel-based, spatial-based, mixed vegetation classification, coffee gardens, Lampung, Indonesia*

ABSTRACT

Classifying a mosaic of coffee systems, each in a different stage of structural complexity is not obvious when that ranges from monoculture to a complex agro-ecosystem, with various shade and fruit trees mixed in different degrees of density. Distinction into different sub-classes incorporating tree complexity and tree cover, is important as tree density and the generally related amount of litter are important from a soil erosion perspective. In this study, the objective was to classify different coffee garden systems plus several other minor vegetation classes existing in the area using IKONOS in Sumberjaya district, Lampung Province, Indonesia. Pixel-based classification approach was integrated with spatial-based approach to reach an improved classification result. In the supervised pixel-based approach training samples are collected to generate statistical parameters for the classifier to classify the whole image. The spatial-based approach refers to segmentation procedure, known also as object-based classification. Two methods of integration were explored and pure pixel-based-approach was as well conducted for comparison purpose. Results were then tested using ground check data. The methods tested are: pure spectral approach of (a) supervised classification using maximum likelihood classifier, integration with segmentation which was done in two ways, by (b) classifying the segments and by (c) combining the pixel-based classified image with segment image using majority rule. Of all the three methods the combination using majority rule showed the highest overall accuracy. Several points were discussed as feedback to the methods tried as well as to improve the classification result

INTRODUCTION

Sumberjaya district located in the northwestern part of Lampung Province, Indonesia is dominated by coffee gardens, which cover 70 % of the area (Dinata, 2002). The district, which also coincides with the catchment of Way Besai, the major river in the area, was a conflict area between guardians of state forest land and farmers, triggered by deforestation and a perceived loss of watershed function related to erosion and river sedimentation during the last decades. For that reason it was considered important to obtain detailed information on land use, with a special focus on the various coffee systems as they are considered an important factor in the erosion process.

Low resolution imagery like LANDSAT (E)TM or SPOT allows to discriminate to a certain extent between monoculture coffee on one hand and multistrata systems on the other. From imageries from 1973, 1986 and 2000 (Dinata, 2002). It was concluded that besides deforestation there was a parallel evolution from monoculture systems towards various multistrata systems. To what extent can high resolution imagery like IKONOS be useful to differentiate in more detail between those various mixed coffee systems.

SITE DESCRIPTION

The study area is located in the district of Sumberjaya, Lampung Province, Indonesia. Figure 1 shows the location of Sumberjaya District and the study site within this district. Coffee gardens is the major land use in this area, which occupy most of its undulating terrain.

Coffee gardens exist in a range of complexities as well as canopy covers , and specifically in this site, they can be classified as Sun coffee, Simple shade coffee, Shade polyculture coffee and Jungle coffee (Verbist, *et al*., 2002). Figure 2 shows the illustration of those types.

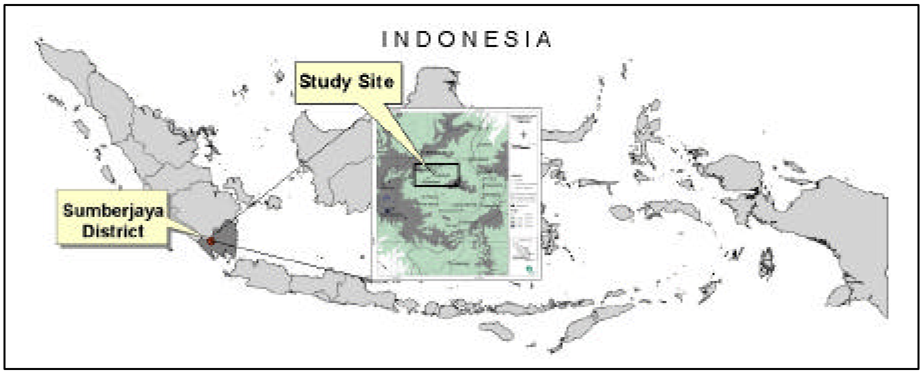

Figure 1. Study site as a subset area of Sumberjaya District

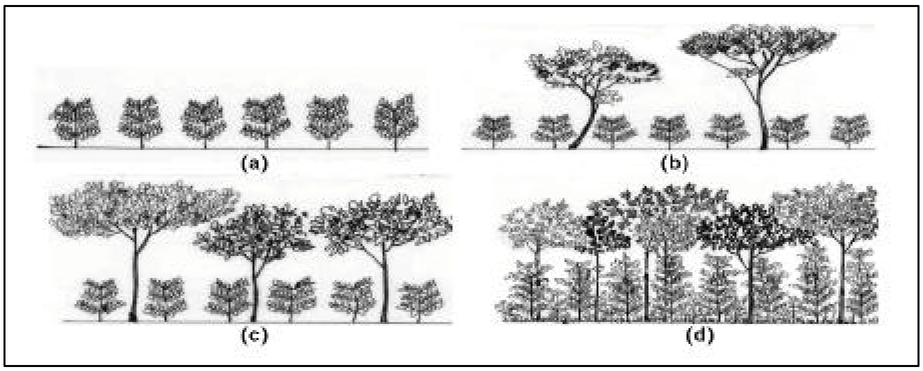

Figure 2 . (a) sun (monoculture) coffee, (b) simple shade coffee, (c) shade polyculture coffee, (d) jungle coffee *(drawings by Wiyono, ICRAF-SEA)*

The types of coffee fields are described briefly below (Verbist, *et al*., 2002) :

- **Sun coffee** (Figure 2 (a)) field is monoculture field, which is generally intensively managed.
- **Simple shade coffee** (Figure 2 (b)) is a simple shade coffee field which uses a single, pruned canopy species to provide shade. The most common species are *Erythrina, Gliricidia sepium, Paraserianthes* sp. The vertical structural diversity of this type is rather poor.
- **Shade polyculture coffee** (Figure 2 (c))**:** the coffee is planted under shade trees like *Erythrina lithosperma*, *Leucaena glauca* and *Albizzia falcata* and is often mixed with fruit trees or medicinal plants. This type of coffee field also often occurs as homegarden near by settlements.
- **Jungle coffee** (Figure 2 (d)) rarely occurs at present. It is an old system, which involved coffee planting under the forest canopy. The coffee trees would not be pruned and would grow quite tall (<8 m).

Coffee gardens possess different levels of patchiness. Jungle coffee has an almost 100 % closed canopy cover, and therefore the level of patchiness is relatively lower than the others. On a relative short distance a shade polyculture coffee may range from a relatively closed canopy cover to a more open one, and therefore the degree of patchiness of this type is more various than jungle coffee. Simple shade coffee commonly exists with more open canopy compared to jungle coffee or shade polyculture coffee, and with at least two levels of strata (coffee bushes and the shade trees) plus an open ground to various extents, in general, this type has a high degree of patchiness. For the sun coffee, the level of patchiness depends very much on the density of the coffee bushes. However sun coffee with higher cover than 50 % is rarely found.

Aside from coffee gardens, shrubs or secondary growth may also occur in different patchiness, which is more due to the heterogeneous bush height and various extents of open ground.

METHODS

IKONOS Multi spectral images with 4 m spatial resolution and 4 spectral bands (3 visible bands and 1 near infra red) was used in this study, covering an area of approximately 4 * 9 km^2. It is projected into UTM, zone 48 south, based on ellipsoid WGS 84.

The coffee gardens as classified above have various levels of patchiness, which is image resolution dependent. The patchiness is captured in the image as the within-patch spectral variability resulting in a certain texture. As related to the different levels of patchiness, the classification of coffee gardens are determined by the type of coffee field as well as the canopy cover, which is a sum of the canopies of the coffee and of the shade trees. Other land cover types in the area like ricefields, herbs and grass, cleared land and water are relatively homogeneous and therefore have low within-patch spectral variability. The land cover types to be classified in this study are shown in Table 1.

Table 1. Land cover classes to be distinguished in this study

Land cover Class	Class code
Jungle coffee or shade polyculture coffee ; closed canopy cover (>50 %)	C1
Shade polyculture coffee or simple shade coffee ; 25-50 % canopy cover	C2
Sun (monoculture) coffee; 25-50 % canopy cover	C3
Sun coffee or simple shade coffee; 25 cover	C4
Newly planted coffee; very sparse cover	C5
Ricefield, inundated	R1
Ricefield, dry	R2
Herbs and grass	H
Woody shrubs, secondary growth	S
Cleared/ opened land	B
Waterbody	W

Aside from the classes to be classified above, there are two regions in the study site being masked prior to the classification process. They are: small town of Fajarbulan, located in the western part of the study area and forested hill of Bukit Rigis, located in the southeastern corner of the study area.

The flow diagram of the overall methods applied in this study is shown in Figure 3.

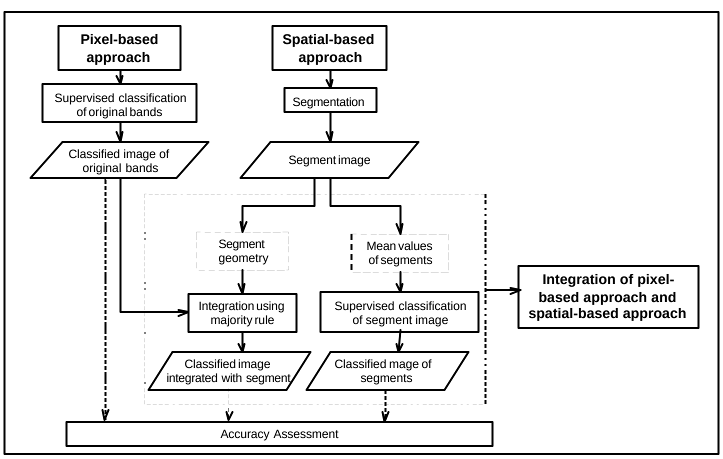

Figure 3. The flow diagram of the approaches applied in the study

Pixel-based approach

The pixel-based approach defined in this study refers to the supervised classification method using Maximum Likelihood (ML) classifier. All image processing was done with ERDAS Imagine software. This pixel-based classification was used as the standard approach to which results of spatial-based approaches were compared and integrated.

Spatial-based approach utilizing segmentation algorithm

The high resolution image of 4 m used in this study showed much variability in neighboring values and, as explained above, some of the classes (like the various shade coffee systems) inherently appear as textured features. The approach in dealing with the textured image was to classify it into objects on the ground, which in this particular case represent patches of vegetation, and for this study, a segmentation approach was applied. Segmentation is basically an automated procedure representing the way human eyes recognize a group of pixels as individual objects. It divides the image into objects and in recognizing those objects, it follows the hypothesis that neighboring image elements belong to the same class (Schiewe, *et al*., 2001).

Segmentation algorithm

A segmentation process incorporates a set of procedures in splitting and merging the region based on what is defined as 'different regions' (to be split) or the 'same region' (to be merged). Segmentation approach is largely used for images with high speckles such as SAR data. This type of segmentation follows a "cartoon model" which is a decomposition of an image into a collection of regions on each of which the intensity is constant (Cook, *et al.*, 1994). The mean values of the region are then applied in the cartoon of the segments.

The segmentation algorithm that will be used in this study is called Merge Using Moments (MUM) (Cook, *et al.*, 1994). The principle behind MUM algorithm is that at the initial stage, each pixel is a region, which is followed by a merging calculation of the 'similar' adjacent regions. For two adjacent regions, it calculates the likelihood (probability) of whether they belong to the same actual region by assessing the statistical properties. The two similar regions are tagged and those tagged regions are in the end merged. If two regions are significantly different, they are left as they are (no merging is conducted). In determining whether two regions are significantly different, the log probability of merging and splitting are calculated using the following algorithms (Cook, *et al*., 1994).

The a priori means (m) of the intensity (I) with N as the number of pixels in a region:

$$m = {}^{1}\!/_{N} \sum_{i=1}^{N} I_i = \bar{I}$$

(1)

The log likelihood of splitting two regions (A and B) is given below, whereby the parameter L represents the level of texture present in the image:

$$l_{split} = L\left(-N_A \log(\bar{I}_A) - N_B \log(\bar{I}_B) - N_A - N_B\right)$$

(2)

The log likelihood for merging two regions are as follows:

$$l_{merge} = L\left(-(N_A + N_B)\log\left(\bar{I}_A + \bar{I}_B\!/_{2}\right) - N_A - N_B\right)$$

(3)

The log likelihood of the difference, $l_D = l_{split} - l_{merge}$ is as follows:

$$l_D = L\left(-N_A \log(\bar{I}_A) - N_B \log(\bar{I}_B) + (N_A + N_B)\log\left(\bar{I}_A + \bar{I}_B/2\right)\right)$$

(4)

A low value of L indicates a textured image and is typical for radar imagery with its inherent *speckle,* of which a great deal is in fact noise. Filtering then helps to increase the signal-to-noise ratio. High value of L means low texture, and it implies a less severe smoothing treatment. As optical imagery doesn't contain '*speckle',* the texture is a reflection of the condition on the ground like patchiness or internal shadows.

In executing the segmentation procedures, to determine whether two regions are to be merged or to be split, a 'merge threshold' should be set (P). Two adjacent regions will be merged if the probability that they are samples from the same distribution is larger than this threshold. Smaller values of the threshold will result in larger regions, which means less segments in the image.

In this study, a segmentation module in ERDAS Imagine environment, called CAESAR Module, which was originally developed for the processing of radar imagery, was used to conduct the segmentation process with some adjustment in setting parameters to process optical image.

Iterative visual inspection was used to decide what input parameters allowed the algorithm to classify appropriate objects on the ground, using the segments of land cover types with a high within-patch spectral variability, i.e coffee gardens, as the bases. As noted by Schiewe, *et al.*, 2001, segmentation is indeed still an iterative and manual approach.

A larger l_D means a higher probability in splitting the regions, and lower probability of the regions to be merged into one region. Visually this will result in an image with more segments which means higher variability of values.

Integration of pixel-based classification and spatial-based approach

Referring to the "Integration between pixel-based approach and spatial-based approach" (see Figure 3), the first integration was a supervised classification of the image of segments, utilizing the same training samples as the ones used for classifying original bands.

The resulting segments have only the mean pixel value of the merged-regions. In the spectral scatter diagram it was seen as lower density of points yet with the same distribution as that of the original bands. With the same training samples used, in the scatter diagram, most training patches will become single-valued. These single-valued patches have lack of statistical parameters in order to use the ML classifier. Hence the minimum-distance-to-means (MD) classifier will be used to classify the image of segments. In classifying an image, MD classifier only considers the shortest euclidian distance between the arbitrary pixel and the training pixels.

The second method of integration, called "Integration using majority rule" (refer to Figure 3), uses post processing spatial layering method and is principally the pure combination of spatial and spectral properties of the image. Image of segments provides the geometry of the homogeneous areas considered as the objects on the terrain. The labels of the land cover classes come from the land cover class resulted from pixel-based classification. The decision to assign a particular land cover class to a certain segment boundary follows majority function. The diagram below (Figure 4) describes this method

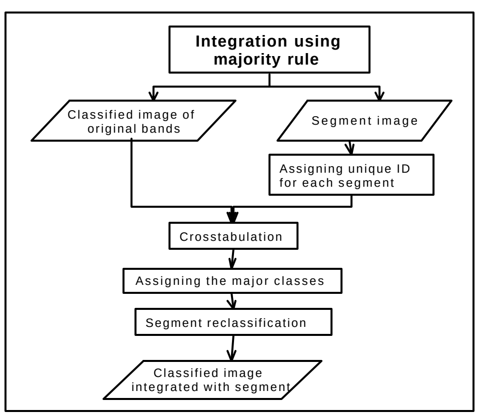

Figure 4. Diagram of integration using majority rule

Accuracy assessment

The accuracy assessment is done with error matrices and Kappa statistics. The per class accuracy will be used as the measure to compare accuracies among methods.

RESULTS AND DISCUSSION

Data collection in the field

The training samples and ground truth samples were collected using GPS (Global Positioning System). To train the classifier approximately 2400 pixels in 52 training patches were collected in the field. For the accuracy assessment, independent samples for groundtruth were also collected during the same period and 502 points clustered in 50 land cover patches were used for this purpose.

Segmentation

In CAESAR Module two parameters had to be set, L and P. Since the effect of changing L is equivalent to the changes in P (Cook, *et al.*, 1994), P was set to a fixed value (10^{-10}) and various L values were tested in order to find the best value in judging whether the image has produced reasonable regions, representing land cover patches on the ground. The values put on trials were 100, 50 and 20, and the one considered by visual inspection to give the best result was L=50. This was mainly so for the coffee class, although the same value had to be applied for the entire image. The resulting image of segments is shown in Figure 5 (b)

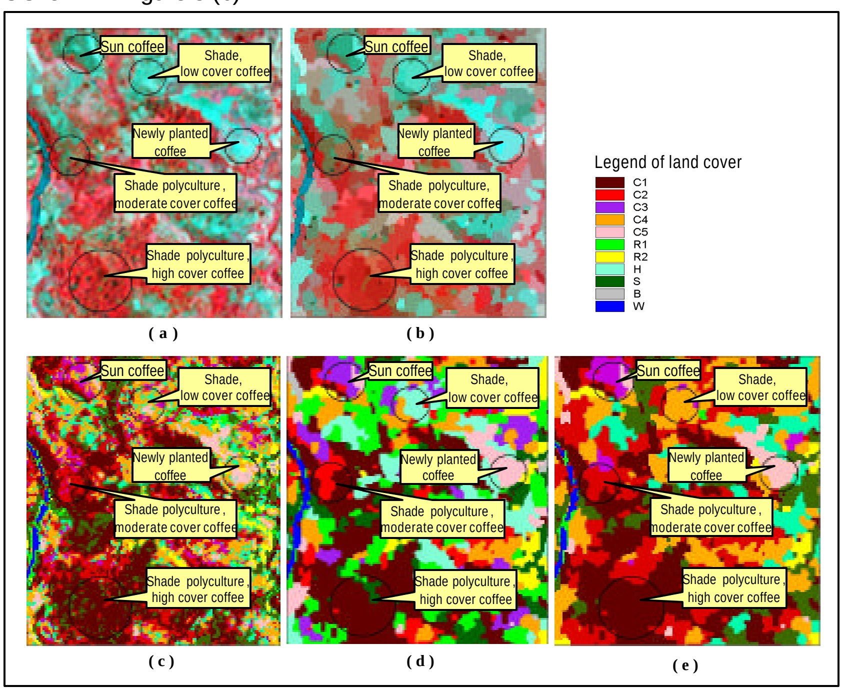

Figure 5. (a) original bands, false color composite , (b) image of segments, (c).classified image of original bands , (d). classified image of segment, (e) classified image integrated with segments

Classified image of original bands

The false color composite image of IKONOS original bands is as in Figure 5 (a) and the classified image of original bands is shown in Figure 5 (c). The classified image of original bands gives 80% overall accuracy. The intermediate classes of coffee gardens (C3 and C4) seem hard to separate and have a low accuracy (65 % and 69 % respectively) while the low accuracy of herbs and grass (52 %), can be attributed to a spectral confusion with dry rice (R2).

Classified image of segments

The Classified image of segments (Figure 5(d)) shows slightly lower overall accuracy (79.6 %) than that of the classified image of original bands.

One thing to note about this result is that when comparing the classified image of segments with the classified image of original bands visually, the land cover with high range of means values were overrepresented e.g. R1 (Inundated rice) and H (Herbs and grass)., because 'herbs and grass' and 'inundated rice' were indeed spectrally mixed with other classes. Since the classification was done to segments, it created an overrepresentation of the larger areas at the cost of smaller features, while in reality they normally exist in small patches. Accuracy improvement occurs in classes of C3, C4, and S (100%, 86% and 100% respectively). So for classes with high within-patch variability accuracy already increased with this method.

Classified image integrated with segments

The integration between classified image and image of segments image using majority rule gives the highest overall accuracy (Figure 5 (e)) of 85.4% of the three methods. Visually, when compared to the classified image of segments alone, it is observed that the overestimation of H and R1 is reduced. This approach improves the accuracy of some coffee classes considered as problematic due to high within-patch variability, i.e. C3 (100%) and C4 (89%). The other coffee gardens classes (C1, C2) show similar levels of accuracy.

The low accuracy of herbs and grass in all methods is due to the spectral problems rather than object classification problem. Herbs and grass (H) and dry ricefield (R2) are indeed spectrally mixed with other classes.

In all, although only limited patches could be sampled for the ground truth test, based on the accuracy assessment done in this study, the integration method using majority rule indicates an improvement for the classification of classes which have high within-patch spectral variability , i.e. sun coffee gardens with moderate cover (C3), sun or simple shade coffee with low cover (C4), and woody shrubs/ secondary growth (S). This takes place because the integration optimizes information from both the original spectral responses and the geometry of the objects on the ground. Based on the t test conducted, at $\alpha = 0.05$ this method indicates significantly higher overall accuracy, although this should be considered carefully due to the small number of sample used for the test.

CONCLUSION AND RECOMMENDATION

Segmentation or commonly also known as object based-classification is superior to a pixel-based multi spectral classification, mainly for classes with a large within-patch spectral variability, as it groups neighboring pixels as objects prior to the assignment of specific land cover type. However, it should be noted that segmentation to this extent still employs much of human intervention in recognizing the objects on the ground. Selection of the value of the segmentation parameter needs local field knowledge.

The whole image was segmented using the same threshold, which was optimized for coffee gardens, causing relatively homogeneous spectral classes like rice or grass to be overrepresented. A so-called multi-level segmentation, whereby a different threshold and relevant parameters can be set for different objects is likely to give better results.

This study indicates that the integration of pixel-based approach and spatial-based approach using majority rule improves land cover classification of a high spatial resolution image covering an area with a local high spectral variability (e.g. mixed-vegetation area, agroforestry area). Further studies with more test samples and more elaborate analysis are needed to test if the higher accuracy is significant.

REFERENCES

CAESAR User Guide, Version 3.0 for ERDAS Imagine, (1998). NASoftware, UK.

Cook, R., *et al.,* 1994. MUM (merge using moments) segmentation for SAR images. In: Europto 1994, Rome.

Dinata, A. E. P., 2002. Deteksi perubahan lahan menggunakan citra satelit multisensor di Sumberjaya, Lampung. Jurusan manajemen hutan, Fakultas kehutanan. Bogor, Institut Pertanian Bogor and ICRAF-SEA, Bogor, Indonesia.

ERDAS Field Guide, 3rd ed., 1995, Atlanta: Georgia, USA.

Feingersh, T., 2000. Synergy of multi temporal SAR and optical imagery for crop mapping, International Institute for Aerospace Survey and Earth Sciences (ITC), Enschede, the Netherlands.

Gorte, B., 1998, Probabilistic segmentation of remotely sensed images, International Institute for Aerospace Survey and Earth Sciences (ITC), Enschede, the Netherlands

Schiewe, J., *et al*., 2001, Potential and problems of multi scale segmentation methods in remote sensing, http://www.definiens-imaging.com/down/

Verbist, B., *et al.,* 2002, Biodiversity within coffee agroforestry systems: Can driving factors of land use change be influenced ?, in press.