SPATIAL OBJECT BASED REMOTELY SENSED IMAGE CLASSIFICATION TECHNIQUE FOR TEA PLANTATION MAPPING IN SRI LANKA

M D Samanthi Ranjanie[1], Damitha D Karunarathna[2]
[1]Postgraduate Institute of Archaeology of University of Kelaniya (PGIAr),
407, BauddhalokaMawatha, Colombo 07, Sri Lanka
Email: msamanthi@gmail.com
[2]E-Learning Centre, University of Colombo School of Computing (UCSC),
35, Reid Avenue, Colombo 07, Sri Lanka
Email:ddk@ucsc.lk

KEY WORDS: Land-use Mapping, SORSICT

ABSTRACT: This study presents a spatial object based remotely sensed image classification technique (SORSICT) to identify the tea plantation land-use from satellite imagery. In Geographic Information Systems, supervised learning is used to classify remotely sensed imagery data to develop land-use maps. The traditional supervised classifiers used for this process tend to generate irrelevant classes that subjected to use the post-classification methods for correcting the errors of the land-use classes.

This research found a solution for the problem of how do the current image classification algorithms be refined to generate accurate land use thematic maps for the tea plantation in Sri Lanka. The new approach is a refined supervised classification which uses the conventional minimum distance decision rule. It incorporates a spatial object based threshold scheme to limit the insertion of the irrelevant land-use types in a particular region. That spatial threshold scheme is a vector-polygon layer generated by a Geographic Information System using the prior knowledge about the land-use pattern of the particular region. It consists of the major divisions of the region with the attribute of possible land-use types in each division. Those major divisions are the spatial objects that provide the boundaries to select only the relevant classes of land-use types for the classification.

SORSICT is tested for upcountry tea plantation of Sri Lanka. Quick Bird-2008 Satellite imagery of a selected district of Sri Lanka was used to evaluate the proposed approach. Comparison of the results of SORSICT with the conventional methods reveals higher accuracy in SORSICT at identifying tea plantation from remotely sensed data. SORSICT provides a solution for the problem of generating irrelevant classes in remotely sensed imagery classification by applying a spatial threshold scheme.

1. INTRODUCTION

In per pixel crop classification two kinds of effects namely spectral variability and mixed pixel event frequently occur and strongly corrupt the classification result (Smith & Fuller, 2001). To improve the classification accuracy eliminating these effects advanced classification approaches and techniques were developed. In the new generation a departure from the classical pixel-spectra based image analysis remote sensing has been the adoption of object-based and knowledge-driven methods that absorb the advantage of the spatial, spectral and temporal characteristics (Hay et al.,2008; Deren et al.,2000;Arbiol et al., 2006; Richards et al., 2005; Duong, 2003).

Tea plantation occupies 5% of cultivated land area and absorbs 13 % of labour force in the country as a prominent foreign exchange earner of Sri Lanka(Annual Report, 2011; Tea Market Update, 2012). It is obvious that the GIS and remote sensing techniques could be used in many different ways to support the management of tea plantation in Sri Lanka.

Only a few research publications can be found on the application of remote sensing technology for tea plantation management in local and international level. Dutta (2006) presents an application of remote sensing technology for tea plantation to demark the extent and the type of tea planted in deferent land areas. In order to map tea type and extent remote sensing offers efficient and reliable factors. Some previous studies on remote sensing data analysis have revealed that spectral characteristics of tea leaves and canopy of Sri Lankan tea plantations (R.M.S.S. Rajapakse, et al, 2002; Pushpakumari, H.M.M.P., et al, 2008). However these studies do not concern about land-use mapping of tea plantation in Sri Lanka.

This paper proposes a new method to refine a current image classification algorithm to generate accurate land use thematic maps for the tea plantation in Sri Lanka.

2. STUDY AREA AND DATA

To test the proposed classification method two sample areas are selected from Kandy District of Central province of Sri Lanka in extent of 5.73ha and 32.6ha. The major landuse type of these sample areas are tea crop.

The sample images are subset from multispectral satellite imagery dataset of Quick-Bird (USA) of 2008 with 2.4m resolution and four bands (Red, Green, Blue and Near Infrared). To enhance the spatial resolution the multispectral dataset was pan-sharpened using panchromatic data with 0.6m resolution of equal source and year of multispectral dataset.

The ground reference data collected by field surveys in 2008 is used for the training stage in supervised classifications and tocreate the ground referenced GIS vector layer. Existing maps and survey reports of the estate were used to gather ancillary data.

3. METHODOLOGY

The solution proposed to improve the accuracy of the classification refines the supervised classification approach using minimum distance decision rule. It uses a data format from a Geographic Information System to define the spatial boundaries which are used to determine the appropriate land-use types. The proposed classification process above explained comprises the following six phases.

3.1. Pre-processing the image.

The typical process used at this stage is pan-sharpening, which merges the high-resolution panchromatic image with the corresponding lower resolution multispectral imagery to create a single high-resolution color image.

3.2. Selecting ROIs.

The pan-sharpened, high resolution colour image is used to identify regions of interest (ROIs) to generate class signatures representing different types of land coverage. The identification of ROIs is a manual process which required human expert knowledge. Two text files are created in this step namely, class definition file and sample definition file.

The class definition text file is composed, in each of its lines, four numbers for the class number and RGB values and one string for the class name. For example, the line that defines the ‘tea’ is written as 2 0 255 0 Tea. The first number that represents the class identification number is unique. Second column has the Red, Green and Blue values of the representative colour for the classification output image. The class name should be meaningful to the represented land-use type of the real ground.

The training samples are manually selected for every class. Most of the classes have more than one sample training sites, because of the spectral heterogeneity of data within one category. Therefore at-least one pixel should be given for one variation of the category.

3.3. Signature formation.

After the classes and samples are defined, the signatures are generated through an automated process. The input image file, class definition file and sample definition file are applied to generate signatures that used as the electronic keys of the land use categories in the classification.

For a Minimum Distance Classifier the signature is an n dimensional mean vector corresponding to the average of all pixels in all samples for a particular class, where n is the number of features (or image bands). The mean vector μ for class *i* is calculated for each feature as:

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} x_j \tag{1}$$

Where N is the number of samples and x_j is the j^{th} sample. For example the signature of class tea for the image sample 1 is generated as:

2 137.58 216.65 213.69

3.4. Creating polygon layer of land-use prior knowledge.

From the fourth step the modification of existing classification method is initialized. The spatial data boundaries that used to select the appropriate signatures are created using land-use prior knowledge. So that a Geographic Information System vector data structure - polygon layer (feature class / shape file) is built up consisting meaningful polygon objects that represent the attributes of the major land-use or land-cover categories of the area.

In other words that polygon layer consists of the boundaries of major land-use types of study area.

Figure 3.1 delineates the features of GIS layer corresponding to the image sample 1 and 2 that created using the software package, ArcGIS 9.2.

(a) (b)

Figure 3.1. (a) GIS Layer for sample No.1 (b) GIS Layer for sample No.2

3.5. Polygon detection and signature selection.

It is the most important stage that formulates the new modification. The central coordinate values of the candidate image pixel are compared with the GIS layer and recognize the corresponding polygon area that the pixel belongs. Polygon detection is done using the solution presented by Philippe Reverdy (Chowdhury & Talukder, 2005) the Signed Angle Algorithm to determine whether a point lies on the interior of a polygon. After recognizing the corresponding polygon area that the pixel belongs, the attributes of that polygon which consist of possible land-use types of the area are written to new data structure. That object is used to select the signatures in the classification process. In other words, the signatures which represent those possible land-use classes are only considered as the appropriate signatures which can be assigned to the candidate pixel.

3.6. Classification.

The automated process of classification by analysing the discriminant function for only the sorted signatures corresponding to the pixels is performed in this step. The modified classification algorithm is illustrated through the following pseudo code.

```
For all pixel,
For all the polygon bounding boxes,
        If the pixel’s central coordinates are fallen in a polygon,
                Then read the attributes of landuse types of that polygon,
                        Select the signatures of those landuse types,
        Else check the pixel coordinate with the next polygon,
For all Signatures,
Analyze the Discriminant Function (Calculate minimum distance between the pixel value and the signature),
Decide the most suitable class from the selected signatures for the candidate pixel,
Assign the class to the candidate pixel,
Write for a new output image with predefined color of the class.
```

The proposed classification process above explained is further demonstrated by figure 3.2. The boxed section of the flow chart shows the modification of the algorithm that used to select the appropriate signatures for a given pixel. If the accuracy is insufficient the classification is repeated from the step of feature extraction.

4. Results and Evaluation

The application of the proposed algorithm of image classification for tea plantation in Sri Lanka generated the following results (Figure 3.3) for a one image sample. The results of conventional method of minimum distance classification tool of ENVI 4.3 software package is used to compare the results of proposed method in evaluation.

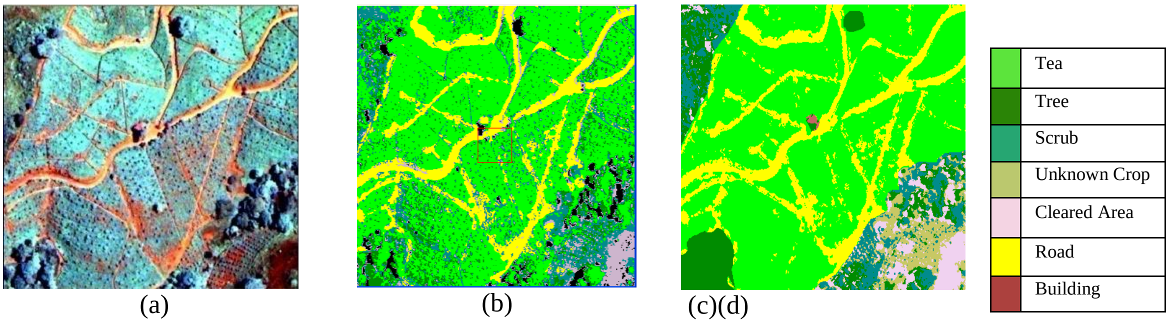

Figure 3.3: (a) Input image (b) Classification output of image sample 1 using conventional method (c) Classification output of image sample 1 using proposed method (d) legend

The evaluation process is delineated in figure 3.4. The data generated by intersections of two layers (reference layer and classified output layer) represent the correctly classified data while the relative complement produces the incorrectly classified data.

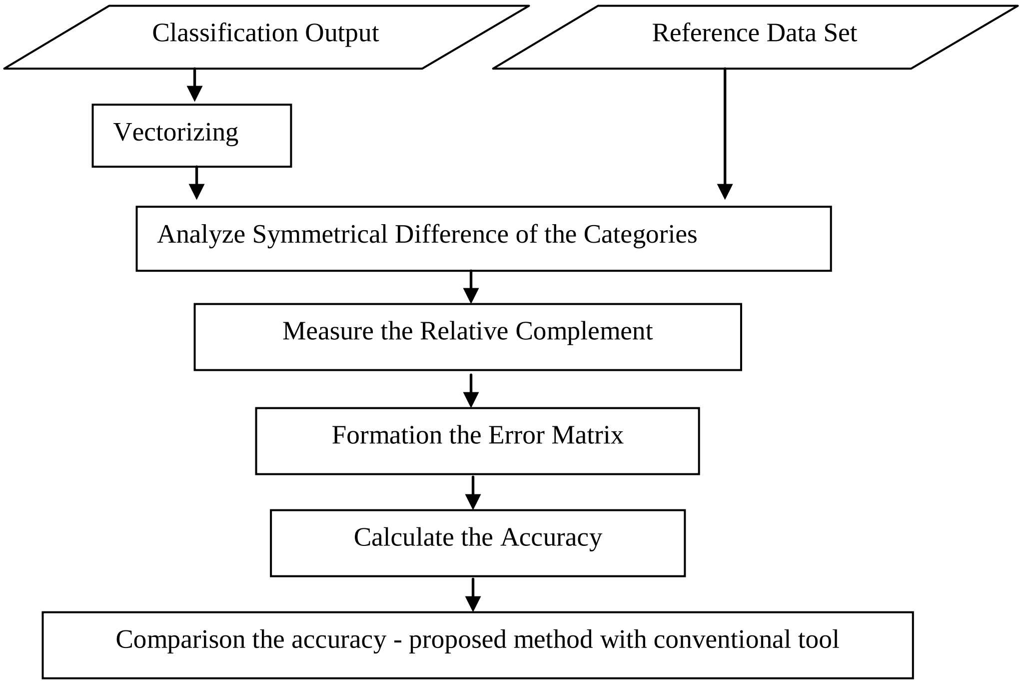

Figure 3.4: The Evaluation Process of Classified Data

The reference data set for both samples are created through manual interpretation that presented in figure 3.5.

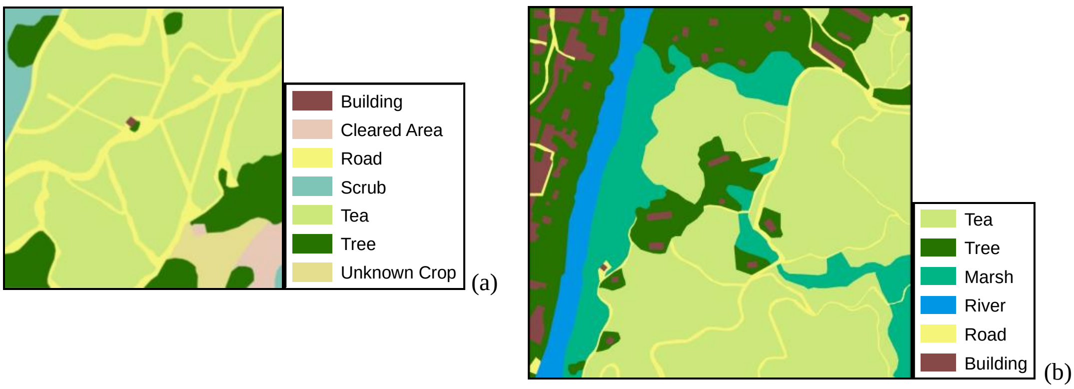

Figure 3.5: Reference Data Sets for Accuracy Assessment; (a) Sample 1 (b) Sample 2

Accuracy Cell Arrays (Error Matrix) are created to calculate the relative accuracy of the reference data to the classified data and vice versa. The relative accuracy of reference data and classified data in proposed and conventional classification methods for both samples are tabulated. The producer's accuracy (2) and consumer's accuracy (3) are calculated.

For a given class,

$$\text{Producer's Accuracy} = \frac{\text{Reference Value}}{\text{Reference Total}} \times 100 \quad (2)$$

$$\text{Consumer's Accuracy} = \frac{\text{Classified Value}}{\text{Classified Total}} \times 100 \quad (3)$$

According to the Accuracy Cell Arrays the relative accuracy of reference data and classified data in proposed and conventional classification methods for both samples are tabulated (Table 1).

Table 1. Accuracies of Classification Results

Sample 1		
	Proposed Method	Conventional Method
Overall Accuracy	81.41 %	68.43 %
Average Class Accuracy	75.14 %	42.64 %
Sample 2		
Overall Accuracy	75.46 %	52.95 %
Average Class Accuracy	56.36 %	51.18 %

Figure 4.1 shows the producer accuracy of sample 1.

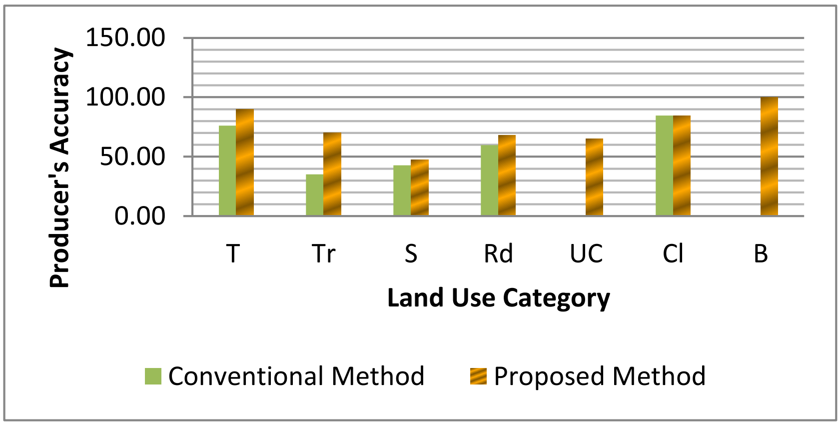

Figure 4.1 : Producer's Accuracy of Classification Results for Sample 1

The measurements of accuracies reveal that the proposed method achieved higher accuracy than the conventional method.

5. Conclusion

The proposed classification method uses GIS polygon layer to select the appropriate signatures to avoid the misclassifications. Minimum Distance classification algorithm is modified to test the proposed classification method using remotely sensed satellite imagery of up country tea plantation area in Sri Lanka.

The evaluation of conventional Minimum Distance classifier and modified Minimum Distance classifier is performed quantitatively and qualitatively. The measurements of accuracy assessment that produced by the quantitative evaluation shows higher reliability of the proposed classification results than the conventional classification results for both study samples. Some spatial errors of the classification results identified by human eye interpretation and prior knowledge are higher in conventional classification results. But the proposed classifier generates the landuse of tea plantation obviously in the real location and in an approximate extent. The proposed method can be accepted to generate land use map according to the evaluation of higher validity and reliability achieved by it through the comparison of the same classifier of conventional method.

References

Arbiol, R., Zhang, Y., Palá, V. (2006). Advanced classification techniques: a review. ISPRS Commission VII Mid-term Symposium “From Pixel to Processes”. Enschede, NL

(2011). *Annual Report.* Colombo: Central Bank of Sri Lanka.

(2012). *Tea Market Update.* Colombo: Sri Lanka Tea Board.

Chowdhury, F. M. &Talukder, K. H. (2005). The Unidirectional Edge Method: A New Approach for Solving Point Enclosure Problem for Arbitrary Polygon. *Asian Journal of Information Technology 4 (5)*, 537-540.

Deren LI , Kaichang DI & Deyi LI. (2000). Land Use Classification of Remote Sensing Image with GIS Data Based on Spatial Data Mining Techniques. *International Archives of Photogrammetry and Remote Sensing.* Amsterdam: International Congress for Photogrammetry and Remote Sensing.Vol. XXXIII, Part B3, 238-245

Duong, N. D. (2003). Land use/ cover mapping of mountainous area using knowledge based classification approach – Case study of Vietnam. *Geoinformatics*, vol.14, no.1.23-28

Dutta, R. (2006). Assessment of Tea Bush Health and Yield Using Geospatial Techniques. *IIRS*, (pp. 6-33). Dehradun.

Hay, G. J. & Castilla, G. (2008). Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline . In T. L. Blaschke, *Object-based Image Analysis: Spatial Concepts for Knowledge-driven Remote Sensing Applications,*Berlin: Springer. (pp. 75-89).

H.M.M.P. Pushpakumari, N.D.K. Dayawansa and Ranjith Premalal De Silva. (2008). Relationships between satellite derived vegetation indices and leaf biomass for mid and up country tea plantations in Sri Lanka. *GEO OBSERVATEUR* (No. 16).

Smith, G.M. & Fuller, R.M. (2001). An integrated approach to land cover classification: an example in the Island of Jersey. *International Journal of Remote Sensing* , 3123-3142.

Richards, J. A. & Jia, X. (2005). *Remote Sensing Digital Image Analysis* (4 ed.). New York: Springer Berlin Heidelberg.

R. M. S. S. Rajapakse, N. K. Tripathi & K. Honda. (2002). Spectral characterization and LAI modelling for the tea (Camellia sinensis (L.) O. Kuntze) canopy. *International Journal of Remote Sensing , 23* (18), 3569-3577.

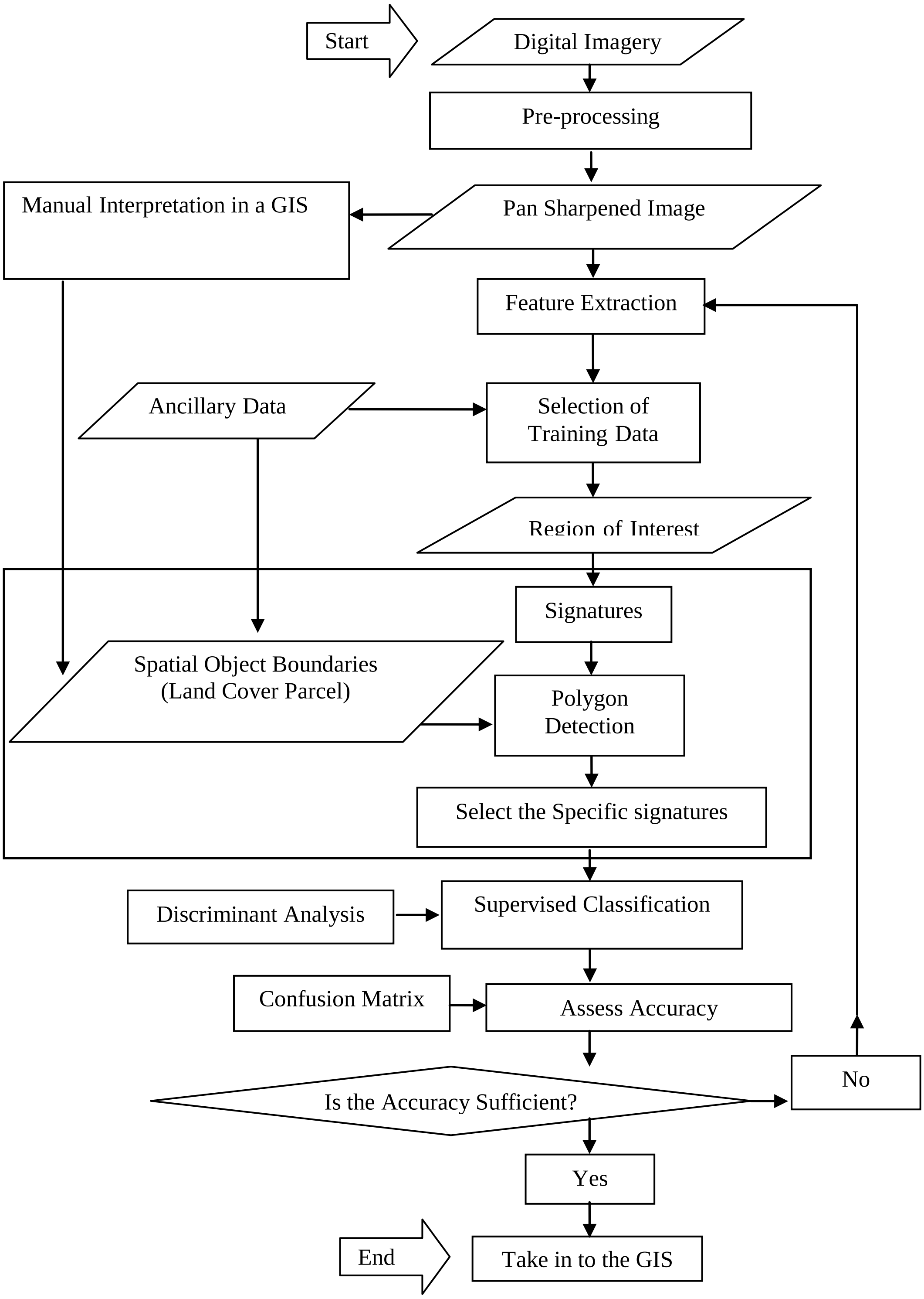

Figure 3.2: Flow of the proposed classification approach