The 34th
ACRS
Bridging Sustainable Asia
Bali - Indonesia, 20-24 October 2013

Proceedings

THE 34TH ASIAN CONFERENCE ON REMOTE SENSING

the festive of science, education, nation and culture

Indonesian Remote Sensing Society and
Asian Association on Remote Sensing

ISBN: 978-602-9439-33-5

IMPROVING LAND USE CLASSIFICATION USING CONTEXT-BASED TECHNIQUES

Teerapat Butkhot[1] Pipat Reungsang[1,2,*] and Rasamee Suwanwerakamtorn[1,3]

[1]Department of Computer Science, Faculty of Science, Khon Kaen University,
A. Muang, Khon Kaen 40002 Thailand, teerapat.bn@angmail.com
[2]Groundwater Research Center, Khon Kaen University,
A. Muang, Khon Kaen 40002 Thailand, reungsang@kku.ac.th
[3]Geo-Informatics Centre for Development of Northeast Thailand, Khon Kaen University,
A. Muang, Khon Kaen 40002 Thailand, rasamee@kku.ac.th

*Corresponding author: reungsang@kku.ac.th

ABSTRACT

Our research aims to automatically classify images from Thailand Earth Observation Satellite (THEOS) into different land types using context-based techniques. To achieve this task, three major steps were performed as follows: (i) images were first pre-processed to produce ground truth without unrelated objects (unrelated-artifacts); (ii) features extraction from image pixels were then carried out by extracting information from each pixel using Bag of Word (BOW) technique to represent the feature vector; (iii) and the last, image classification was performed using Expect and Maximization (EM) algorithm before contextual information was incorporated to improve the classification results by using Markov Random Fields (MRFs) technique. To evaluate the proposed technique, four THEOS images within Khon Kaen Province were collected and classified. Analysis results showed that using our technique for classifying land was superior compare to the conventional technique.

Keywords: Land use classification, expectation and maximization, context-based classification, Markov Random Fields, VZ-texture technique

INTRODUCTION

Land use classification is one of the geographical analysis procedures that play a vital role in land planning and utilizing. In context of segmentation, classification is a process of dividing data into different groups or classes based on their coherent features. Land use classification of map images (remote sensing data); therefore, is a process that segments/identifies regions in map images based on appearances of regions relating to their utilization. Conventional schemes for classifying lands from satellite images into different types based on their usages (called Land Use Classification: LUC) using manual-based approaches are time consuming and prone to reproducibility (Szuster et al., 2011). The development of computer technology cooperated with emerging of machine leaning paradigm offers a new perspective of performing land use classification with automate-based approaches. The challenge of automated land use classification is the high degree of variation in map images such as color and texture of land types. A fundamental technique for segmenting map images is to perform pixel-based classification based on features of the pixels (pixels-based features) such as color, intensity, and texture. Different classification techniques, in addition, can produce different classification. As a consequence, a number of classification techniques have been proposed to classify land in map images recently. Maximum likelihood classifier (MLC) is one of many early methods that have been used for classifying land use in map images (Howarth, 1992; Hames, 2009; Ramita, 2009; Rozenstein and Karnieli, 2011). The techniques estimate model parameters by determining the likelihood of labeled data of given data classes (land types). The model parameters are then used to derive the prediction of image data. In addition, Artificial Neural Network (ANN) has been applied to the classification problem. For example, Ohkubo and Niijima (1999) applied Neural Network to perform pixel-based classification using color properties. The use of color features may produces poor results due to the fact that there is color variation in data images. A number of different features (i.e. color intensities and texture) were used to perform land use classification using maximum likelihood classifiers

(Howarth, 1992). They reported that textural features provided promising results. Supported Vector Machines (SVMs) are one of commonly used techniques that have been applied to land use classification applications (Cortes and Vapnik, 1995). SVMs classify data by generating predictive models that maximizes the margin between data classes in dataset. Learned model is then used to make decisions in map images. Szuseter and Chen (2011) studied performances of a number of classifiers for classifying lands in map images. They showed that there are only marginal different results of classification using different classifiers (i.e. SVMs, ML and ANN). In general, supervised learning methods for land use classification, require labeled data (training data) relating to a number of data classes to generate predictive models. K-means algorithm (ISODATA) is one of unsupervised techniques that have been widely use to classify map images (Rozenstein and Karnieli, 2011; Szuster and Chen, 2011). K-means algorithms assign class labels (K classes) to data items by determining the distance between data items and the mean of data classes. Thus, K-means algorithms derive hard-assignment and exclude information relating to uncertainty of data items, which is useful for post-processing methods to improve classification performance. Expectation and maximization (EM) algorithms, therefore, are proposed to perform unsupervised classification that produce probability outcomes, which are used in this work.

One issue of applying pixel-based-feature techniques (such as color and texture) to classify image maps is that it does not consider the relationships between pixels in images, which may not produce promising results in classification. This paper, therefore, proposes a technique that uses contextual information for classifying lands in map images. The contextual information used in this work takes into an account the spatial relationships of pixels. The proposed technique works in three stages: (i) feature generation using Bag of Word technique, (Leung and Malik, 2001) (ii) classification and (iii) refinement using Markov Random Fields (MRFs) techniques. This paper is organized as follows: material and method will be first presented in the first section. This section will explain the details of data used in this work, and the technique for generating feature and the classification technique. Experiments and results will be demonstrated in the second section before discussion and conclusion are provided in the last section.

METERIALS AND METHODS

This section describes the details of the proposed method for classifying lands in map images based on their utilization. The overall process of the method is illustrated in Figure 1. In addition, the data used in this work for evaluating the proposed method is presented. This section is, therefore, divided into 2 sub-sections including (i) data and materials and (ii) methodology.

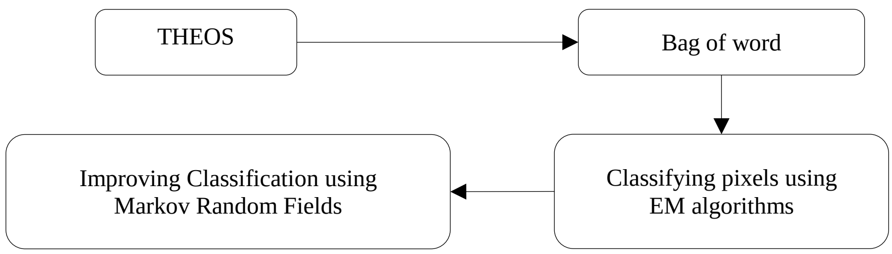
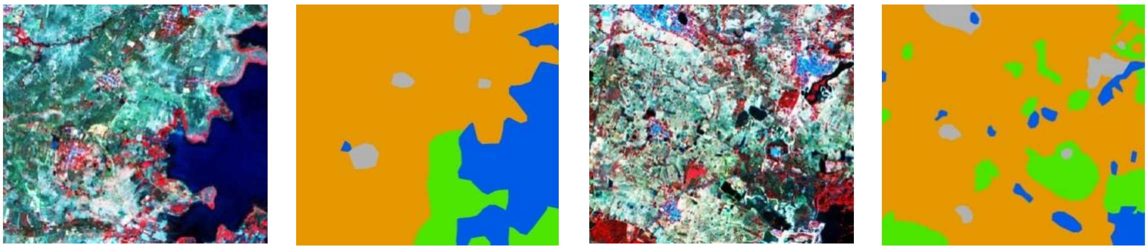


Figure 1. Overall process of the method for classifying lands in map images based-on their utilization using a context-based approach.

A. Data and Materials

The objective of this work is to develop a method for classifying lands into different types. In order to evaluate the performance of the proposed method, remote sensing data (map images) was first specified. In this work, we collected 2007 map images from THEOS in Khon Kaen province areas, Thailand. Image data were divided into sub-images with the resolution of 1000x1000 pixels. Later, 20 sub-images were randomly selected as used as the data in this work. In addition to this data, we collected 5 sub-images with the same resolution and used as a prior data for initializing models in classification step (explained in classification section). Each image data was manually digitized by a specialist, and used as a ground truth data. An example of data is shown in Figure 2.



(a) (b) (c) (d)

Figure 2. Example of image data: (a) and (c) were images collected from THEOS while (b) and (d) were digitized images corresponding to (a) and (c).

B. Methodology

The previous section describes the image data that utilized in this work. This section explains insight into the proposed method for land use classification. The section details the underline technique (Bag of Word) for generating features that were used in classification stage. At this stage, object-based approach, Expectation and Maximization algorithm was implemented as the primary classifier. In order to cooperate contextual information into the process, a post-processing was carried out by taking the output from object-based classification to refine the results using a Markov Random Field technique. Overall processes of the methodology were depicted in Figure 3.

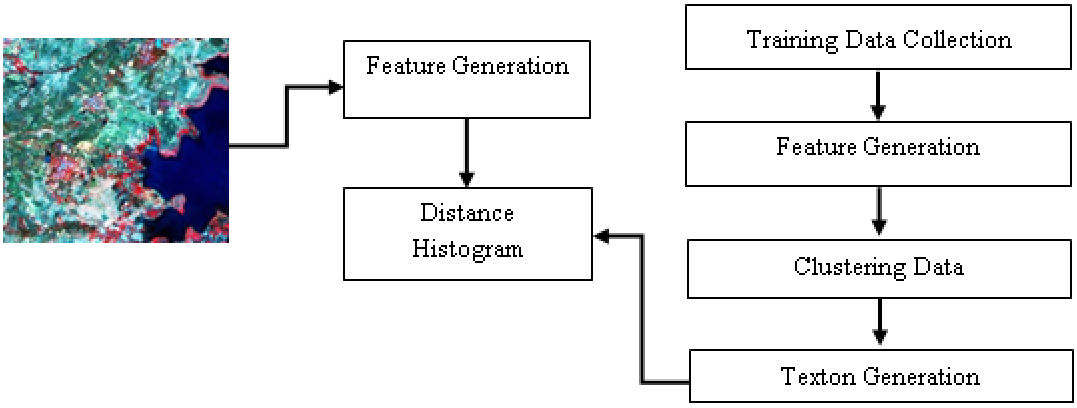


Figure 3. Overall process for feature generation used in classification process**.**

Feature generation using Bag of Words

To generate an object feature, an input image composed of a given pixels, features can be generated based on their appearance. In this work, we had exploited color properties (RGB color model) of the pixels as the base features for classification. To generate the feature, we implemented Bag of Word technique which separated in 2 steps:

1. texton (code book) generation: this step was aimed at defining a template that can be used to generate object feature. To define textons, an additional data was used (called template data). Each pixel in the image is represented by *RGB*-value. The template data is then clustered into

T groups (where T is the number of textons) using K-means algorithm. The mean (m) of each group is finally is used as a texton $\boldsymbol{T} = \{m_1, m_2, m_3, \dots, m_T\}$.

2. feature generation : this step generates object features for classification based-on the texton that was generated in the previous step. Each pixel is proceeded to generate *RGB*-value, which is defined as the feature (y). Using the texton ($\boldsymbol{T}$), the object feature ($\boldsymbol{f}$) is calculated for generating a distance histogram of y and **T** as $\boldsymbol{f} = \{d_1, d_2, d_3, \dots, d_T\}$ where $d_i = ||y - m_\mathrm{i}||$.

Classification

In the previous section, we presented the techniques of extracting discriminate descriptor of pixels in order to perform classification. To achieve the task, given an image I comprising of pixel $\mathrm{P} = \mathrm{p}_1, \dots, \mathrm{p}_\mathrm{N}$ (where N is a number of pixels), we wish to evaluate Φ: $P \times C \rightarrow K = \{1,2, \dots, k\}$ $k =$ number of classes and Φ is called a classifier. We define the set of the feature vectors of pixels $d_1, d_2, .., d_t: N$ as vectors from a d-dimensional Euclidean space. The EM (Expectation-Maximization) algorithm estimates the parameters of the multivariate probability density function in the form of a Gaussian mixture distribution with a specified number of mixtures. Considering the set of the feature vectors $x_1, x_2, .., x_N: N$, vectors from a d-dimensional Euclidean space can be drawn from a Gaussian mixture:

$$p(x; a_k, S_k, \pi_k) = \sum_{k=1}^{m} \pi_k p_k(x) \tag{1}$$

where

$$\pi_k \geq 0, \quad \sum_{k=1}^{m} \pi_k = 1$$

$$p_k(x) = \varphi(x; a_k, s_k) = \frac{1}{(2\pi)^{\frac{d}{2}} |S_k| 1/2} \exp\{Z\} \tag{2}$$

where

$$z = -\frac{1}{2}(x - a_k)^T S_k^{-1}(x - a_k)$$

Where m is the number of mixtures, p_k is the normal distribution density with the mean a_k and covariance matrixS_k, π_k is the weight of the k-th mixture. Given the number of mixtures m and the samples $x_i, i = 1..N$ the algorithm finds the Maximum-Likelihood Estimates (MLE) of the all the mixture parameters, i.e. a_k ,S_k and π_k:

$$L(x, \theta) = logp(x, \theta) = \sum_{i=1}^{N} log\left(\sum_{k-1}^{m} \pi_k p_k\,(x)\right) \rightarrow \frac{max,}{\theta \in \ominus} \tag{3}$$

$$\ominus = \{(a_k, S_k, \pi_k) \colon \} \tag{4}$$

where

$$a_k \in R^d, S_k = S_k^T > 0, S_k \in \ R^{d \times d}, \pi_k \geq 0, \sum_{k=1}^{m} \pi_{k=1.}$$

EM algorithm is an iterative procedure. Each iteration includes two steps, in the first step (Expectation-step, or Estep), we find a probability p_i, k (denoted α_i, k in the formula below) of sample i belonging to mixture k using the currently available mixture parameter estimates:

$$\alpha_{ki} = \frac{\pi_{k\varphi}(x; a_k, S_k)}{\sum_{j=1}^{m} \pi_{j\varphi}(x; a_j, S_j} \tag{5}$$

While the second step (Maximization-step, or M-step), the estimated mixture parameter are refined using the computed probabilities:

$$\pi_k = \frac{1}{N}\sum_{i=1}^{N}\alpha_{ki}, \quad a_k = \frac{\sum_{i=1}^{N}\alpha_{ki}x_i}{\sum_{i=1}^{N}\alpha_{ki}}, \quad S_k \frac{\sum_{i=1}^{N}\alpha_{ki}\Gamma}{\sum_{i=1}^{N}\alpha_{ki}} \tag{6}$$

where

$$\Gamma = (x_i - a_k)(x_i - a_k)^T$$

Refinement Process

From the classification (using the methods explained in previous section), a Markov Random Fields (MRF) will be applied to refine classification results by considering local consistency between pixels (Kindermann and Snell, 1980). The MRF applied in this work was probabilistic models that consider relationships between pixels in a neighborhood to decide pixels classes. Given X and C; X is an image composing $\{p_1, \dots, p_N\}$ as a set of pixels and $C = \{c|1, \dots, c_N\}$ (where c_N is a number of defined regions) as the pixel classes obtained from the output of classification (single-stage or two stage classification), we can define an energy function for a MRF as:

$$\mathrm{H}(X) = H_1(X) + H_2(X) \tag{7}$$

where H_1 is a binary term measuring the similarity of pixelsand their neighborhoods and yet determines global and localhomogeneity of the image. The binary term can be defined as:

$$H_1(X) = \sum_{p \epsilon X}\sum_{t \epsilon \Omega p}\phi(p,t) \tag{8}$$

where Ω_p is a set of neighborhood pixels of p, p and t are neighboring pixels; $\phi(p,t)$ is a function of the similarity of p and t. The function is set to small positive numbers if s and t are in the same class (C) and set to bigger positive numbers otherwise. In this work, we define $\phi(p,t)$as:

$$\phi(p,t) = \begin{cases} 1 & \text{If } p \text{ and } t \text{ are in the same class} \\ 3 & \text{otherwise;} \end{cases} \tag{9}$$

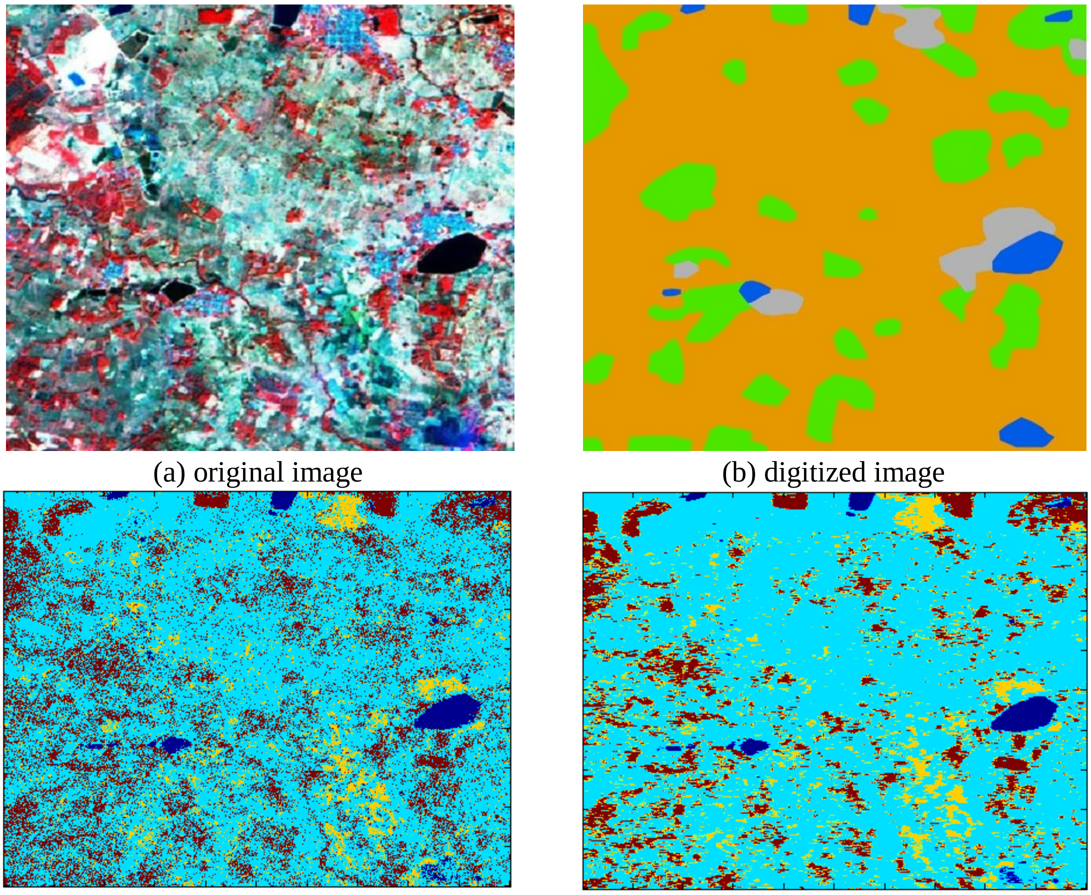
H_2 is an unary term and can be defined by a function of class memberships. Using the results from pixel-based classification, we define H_2 as a function of likelihood function: $H_2 = -log(p(C|P))$, where $\mathrm{p}(C|P)$ is obtained from the EM algorithm. Having defined an energy function, energy minimization is performed using min-cut techniques (Boykov and Veksler, 2001).

EXPERIMENTS AND RESULTS

The previous section explained the proposed method for classifying lands from map images into different types. The proposed method works by generating features from image pixels based on color intensity (object-based features) before classification was performed using EM algorithms. After the classification, a refinement process was carried out to improve the classification using a MRF. This section, therefore, presents experiments that we had conducted to evaluate the proposed technique. We collected 20 sub-images (as our dataset) from the map image within Khon Kaen Province, Thailand. To evaluate the proposed technique, a 5-fold cross validation was performed where each fold was used as prior information (used to initialize the EM parameters). The experimental results showed that the average accuracy of the object-based features is 72.35±5.36. After apply the contextual information, the average accuracy was achieved at 76.34+7.45. Therefore, applying the contextual information can improve the performance of object-based feature classification. In addition, this improvement is statistically significant (p=0.004). The results are shown in Table 1 and example of classification outputs are shown in Figure 4 and Figure 5, respectively.

Table 1 Classification result for each land types.

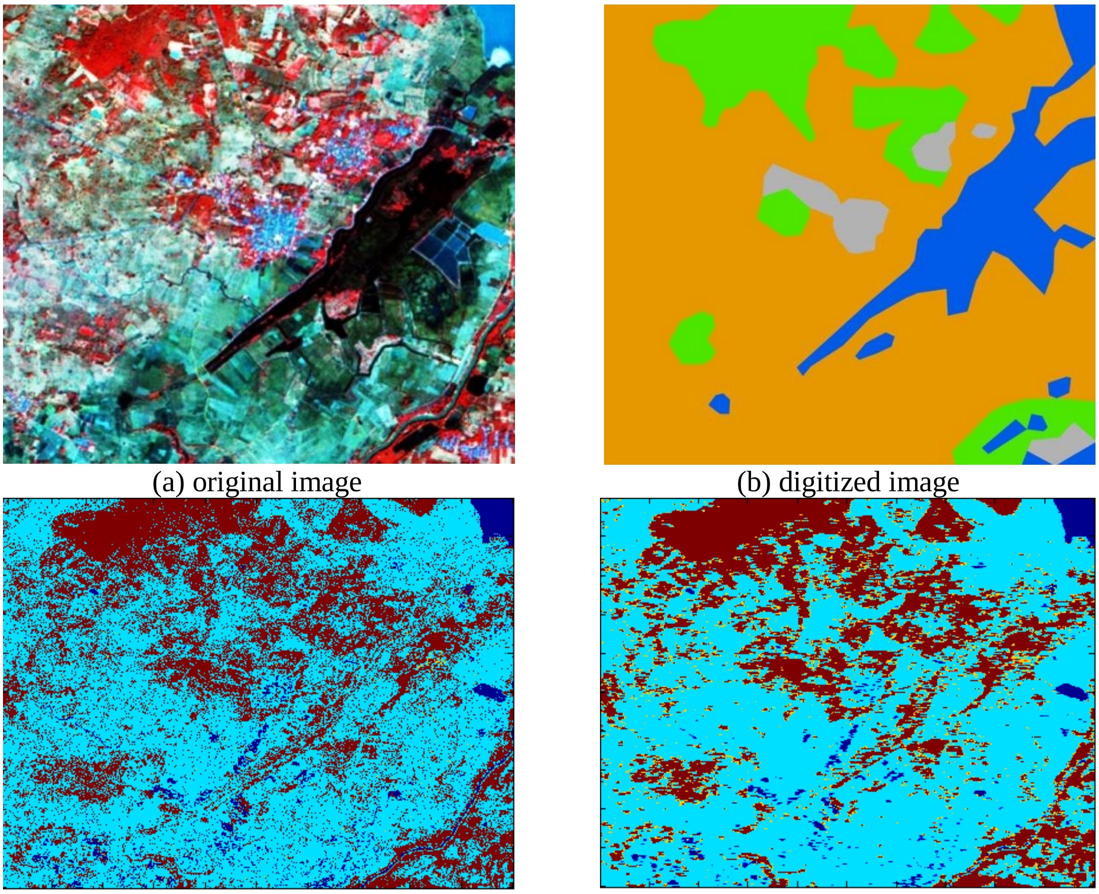
Land Types	Object-based feature		Context-based method	
	True Positive Rate	**False Positive Rate**	**True Positive Rate**	**False Positive Rate**
Water	0.78	0.24	0.82	0.18
Forest	0.70	0.44	0.72	0.34
Urban Areas	0.42	0.14	0.55	0.17
Agricultural Areas	0.54	0.24	0.66	0.18



(a) original image (b) digitized image

(c) object-based feature image (d) context-based technique image

Figure 4 Examples of classification results



(a) original image (b) digitized image

(c) object-based feature image (d) context-based technique image

Figure 5 Examples of classification results.

Table 1 shows the performance of classifying our dataset into different land types. Water and forest areas were well classified for both techniques with high true positive rate (TPR). However, forest areas had high false positive rate (FPR = 0.34) as there were some miss-classification between forest and agricultural areas. This miss-classification was found here as there was marginal different in color properties (as we use color as a based feature for classification in this work). Therefore, a number of agricultural areas were classified into forest, resulting low TPR (0.66) in agriculture and high FPR (034) in forest. In addition, urban area had low positive rate (0.55), which represents poor classification. The issue with this land type was the urban area where was minority class in our dataset. This leads to the problem of unbalanced data in multi-class classification. One possible solution to the problem – as to improve the classification results – is to perform hierarchical classification with the fusion of different features such as color texture. Easy land types (such as water and forest) can be first classified before the remaining land types may be preceded by other methods and features.

CONCLUSIONS

This paper addresses an issue of land-use classification using object-based features techniques. The object-based techniques for classifying land types can produce poor results as they do not consider the relationships between pixels in images. Therefore, this work proposed a technique to classify land into different types by cooperating contextual information into the classification process. This contextual information takes into account of the relationships between pixels in the images. Experimental results showed that context-based classification was superior to the classification using object-based feature alone.

In future work, hierarchical-based classification techniques would be explored. Easy land types can be first classified using a simple classification technique. The remaining land types (difficult classes) can be classified with different classification techniques and features.

REFERENCES

[1] Boykov Y., Veksler O., Zabih R., 2001, Fast Approximate Energy Minimization via Graph Cuts, *IEEE Trans Pattern Anal Mach Intell*, 23(11): 1222-1239.

[2] Cortes C., Vapnik V., 1995, Support-Vector Networks, *Mach. Learn*, 20(3): 273-297.

[3] Hames F. S. A.-A. a. A. S., 2009, Comparison of Four Classification Methods to Extract Land Use and Land Cover from Raw Satellite Images for Some Remote Arid Areas, Kingdom of Saudi Arabia, *JKAU; Earth Sci*, 20(1): 167-191.

[4] Howarth D. J. M. a. P. J., 1992, A Comparison of Spatial Feature Extraction Algorithms for Land-Use Classification with SPOT HRV Data, *REMOTE SENS. ENVIRON*, 40: 137-151.

[5] Jolliffe I.T., 2002, Principal Component Analysis, Springer.

[6] Kindermann R., Snell J. L., 1980, Markov Random Fields and Their Applications, AMS.

[7] Leung T., Malik J., 2001, Representing and Recognizing the Visual Appearance of Materials using Three-dimensional Textons, *Int. J. Comput. Vision*, 43(1): 29-44.

[8] Ohkubo A., Niijima K., 1999, A new supervised learning method of neural networks and its application to the land cover classification, Geoscience and Remote Sensing Symposium, 1999, IGARSS '99 Proceedings, *IEEE 1999 International*.

[9] Ramita Manandhar I. O. A. O. a. T. A., 2009, Improving the Accuracy of Land Use and Land Cover Classification of Landsat Data Using Post-Classification Enhancement, *Remote Sens*, 1: 330-344.

[10] Rozenstein O., Karnieli A., 2011, Comparison of methods for land-use classification incorporating remote sensing and GIS inputs, *Applied Geography*, 31(2): 533-544.

[11] Szuster B. W., Chen Q., Borger M., 2011, A comparison of classification techniques to support land cover and land use analysis in tropical coastal zones, *Applied Geography*, 31(2): 525-532.