TEXTURE-BASED LAND USE CLASSIFICATION OF REMOTE SENSING DATA USING UN-SUPERVISED METHODS WITH MAKOV RANDOM FIELDS TECHNIQUES

Teerapat Butkhot[1] and Pipat Reungsang[1,2,*]

[1]Department of Computer Science, Faculty of Science, Khon Kean University, A. Muang, Khon Kaen, 40002, Thailand; Tel: +66-43-362188
[2]Reginal Centre for Geo-Informatics and Space Technology, Northeast Thailand, Khon Kean University, A. Muang, Khon Kaen, 40002, Thailand; Tel: +66-43-348268; E-mail: reungsang@kku.ac.th

KEY WORDS: Land use and land cover classification, Expectation and Maximization, Unsupervised classification, Markov Random Fields, Remote Sensing

Abstract: Land use and land cover (LULC) classification is a fundamental task in high-level spatial analysis on Remote Sensing data. Several methods and algorithms have been developed to categorize the land into its various uses including supervised learning and un-supervised learning methods. While the use of supervised learning method can significantly improve the accuracy of classification processes compares to un-supervised learning method. The issues arise including a number of training samples required to build models and the need of pre-determined classes of Remote Sensing data which time consuming and domain specific. To overcome the issues, our study proposes a modified technique for classifying Remote Sensing data based on unsupervised-learning methods. The proposed technique works in two stages: (i) pixel-based classification and (ii) classification refinement using Markov Random Fields (MRFs) technique. For the first stage, an expectation and maximization (EM) algorithm is implemented cooperating with pixel-based features generated based-on color intensities and texture descriptors. While the later stage, a graphical model is constructed on image lattices by representing pixels as nodes in the graph, which are linked by edges representing neighborhood interactions of the pixels. To obtain classification refinement, the MRF technique is applied to the graph to solve labeling problems. To demonstrate our technique, 2002 Landsat-TM5 images covering areas in Khon Kean Province, Thailand were classified. Experimental results show that classification accuracy is improved after applying MRFs to refine the classification results. Therefore, overall with our approach, classification accuracy can be improved with less processing time and not domain specific compare to supervised methods.

INTRODUCTION

Land use classification is one of the geographical analysis procedures that play a crucial role in land planning and utilizing. Classification (in context of segmentation) is a process of partitioning images into contiguous regions based intrinsic properties of regions (Zabih 2001). Land use classification; therefore, is process that segment/identify regions in map images based on appearances of regions relating to their utilization. The appearances of regions in image maps can be exposed in a number of ways, for instance, color intensities and textures. The manual classification of land use is usually time-consuming, subjective and is prone to inter and intra-reproducibility (Szuster, Chen et al. 2011). Therefore, automated classification of land use has been proposed. Fundamental technique for identifying regions in map images is to perform pixel-based classification that partitions pixels of images into different classes (Howarth 1992). Land use classification techniques can be divided into two main categories: (i) supervised learning classification; and (ii) unsupervised learning classification (Hames 2009). Supervised learning classification method takes into an account of training data (labeled data) to generate predictive models, which are used to perform final decisions. A number of techniques has been reported in the literature. Maximum likelihood classifier (MLC) is one of many early methods that has been used for classifying land use in map images (Howarth 1992; Hames 2009; Ramita Manandhar 2009; and 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 *et. al* (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. Howarth *et. al.* used a number of different features (i.e. color intensities and texture) to perform land use classification using maximum likelihood classifiers (Howarth 1992). They reported that textural features provide promising results. Supported Vector Machines (SVMs) (Cortes and Vapnik 1995) are one of commonly used techniques that have been applied to land use classification applications. SVMs classify data by generating predictive models maximizing the margin between data classes in labeled data. Learned model is then used to make decisions in map images. Szuseter *et. al.* studied performances of a number of classifiers for classifying lands in map images (Szuster, Chen et al. 2011). They showed that there are only marginal different between the 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.

Unsupervised learning approach is a family of classification techniques that are computed without requiring labeled training data. The technique is also known as data clustering method where data items are partitioned based on cohesive properties of data item without requiring labeled data. One of the most widely used unsupervised techniques for land use classification is *K*-means algorithms (ISODATA) (Rozenstein and Karnieli 2011; Szuster, Chen et al. 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.

This paper is aimed at developing an automated method for classifying Remote Sensing data based on unsupervised learning method. The proposed technique works in two stages: (i) pixel-based classification and (ii) classification refinement using Markov Random Fields (MRFs) technique. For the first stage, an expectation and maximization (EM) algorithm is implemented cooperating with pixel-based features generated based on color intensities and texture descriptors. While the later stage, a graphical model is constructed on image lattices by representing pixels as nodes in the graph, which are linked by edges representing neighborhood interactions of the pixels. To obtain classification refinement, the MRF technique is applied to the graph to solve labeling problems.

METERIALS AND METHODS

The objective of this paper is to perform land use classification using textural appearance of relating to pixels image to classify them into different land types. This pixel-based classification scheme is performed using an unsupervised method, i.e. Expectation and Maximization (EM) algorithm, to produce probability results, which are used in a post-processing method to improve classification performance. Therefore, This section justify the techniques for classifying lands in map images, which is organized as follows: Section (A) provides the details of image data used in this work. Section (B) explains the technique to generate color-intensity-based and texture-based features used to classify pixels into different land types. Section (C) describes the classification method, i.e. EM algorithm that is used in this work before the post-processing methods for improving classification is presented in Section (D).

A. Image Data

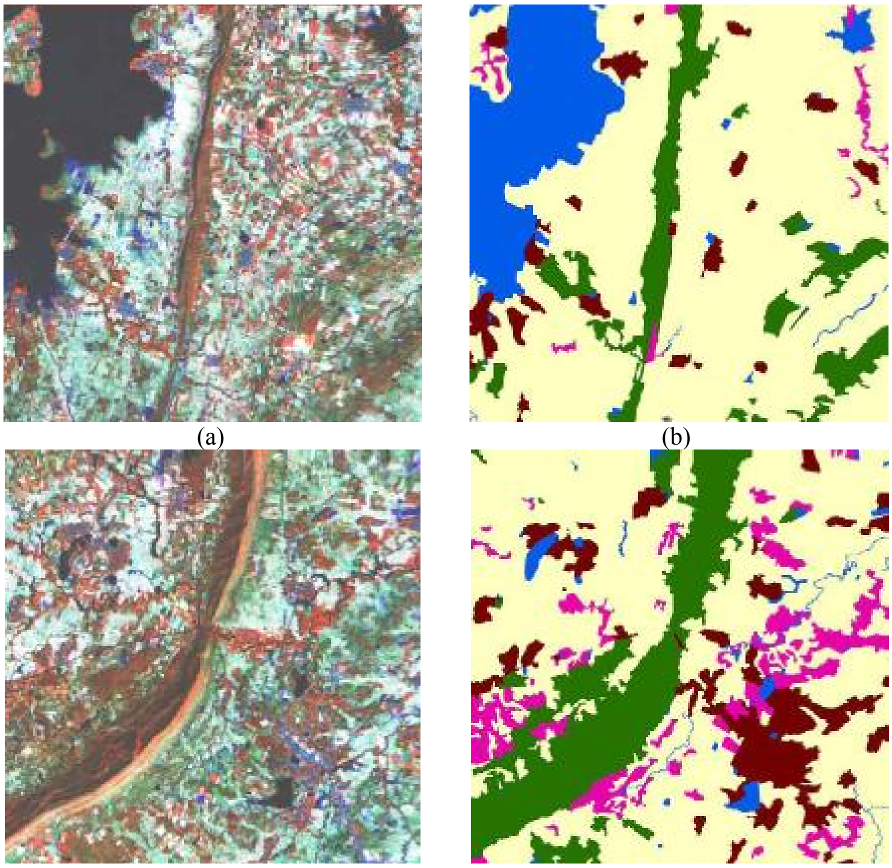
The study area is in Khon Kean Province, Thailand. Remote sensing data of map images are obtained from Landsat TM^5 at 30m-resolution. Map images are divided into sub-images (500x500 pixels) to simplify computation. Examples of image data are illustrated in Figure 1.

B. Feature Generation

To derive pixels descriptor for classification, we exploit object-based features drawn from color and textural information, as follows:

1) Color Features

- color spaces: image are transformed into 3 different color spaces to generate pixel-based descriptor, i.e. grayscales, RGB and LaB*.
- the first principle component of RGB: The principal component analysis (PCA) (Jolliffe 2002) is one of the widely used linear dimensionality reduction techniques. It is an unsupervised method, which provides a compact representation by projecting the data onto a new space, such that the first principal component is associated with the largest Eigen values of the covariance matrix. The largest Eigen value corresponds to the direction where data have the largest variance while the smallest one corresponds to the smallest variance. Hence, we can select the projection matrix composed of the first few Eigen vectors associated with the largest Eigen values. By doing so, we can reduce the dimensionality of our feature space to preserve most of the variance in the data *Y*. The projection vectors presenting the maximum variance directions are computed by solving the Eigen value problem:



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

Figure 1: Examples of image data: (a) and (c) are sub-images obtained from Landsat TM[5], (b) and (d) are corresponding manually digitized images, which are used as the ground truth.

$$CV = \lambda V, \tag{1}$$

where C is the covariance matrix of data samples Y in the feature space R^D. If we assume that the resulting Eigen values λ_i are in the descending order, we can use the Eigen vectors associated with the first d largest Eigen values to project Y to Y' 0 as follows:

$$Y' = (Y - \mu_Y)V_d, \tag{2}$$

where μ_Y is the mean of Y , and V_d are the first d eigenvectors associated with the d largest Eigen values. To derive this color feature we project I_{rgb} into a compact space and set $d = 1$ to generate the PCA feature.

2) Texture Features

Images are divided into overlapping sub-images $(m \times m)$. For each sub-image, we construct the first order statistics properties, i.e. average gray-scales (μ) and its standard derivation (σ) and defined as one of texture features. In addition, a textural feature based on the relationships of pixels in sub-images is also utilized using gray scale co-occurrence matrices (GLCMs). GLCMs are generated based on co-occurring grey values of pixels with particular spatial relationships with respect to the displacement (d) and direction (θ). We set $\theta = \{0^o;\ 45^o;\ 90^o;\ 135^o\}, d = 1$. Thus, 4 co-occurrence matrices are constructed. After normalization, texture features are extracted by calculating statistical measurements from the matrices (i.e. entropy, correlation, contrast, inverse difference moment and angular second moment) -- presented in Table 1. (Haralick 1979). The calculation of GLCMs is shown in Figure 2.

Table 1: list of statistical features generated from co-occurrence matrices.

Feature	Computation
Entropy	$-\sum_i \sum_j G_d(i,j) \log G_d(i,j)$
Contrast	$\sum_i \sum_j (i-j)^2 G_d(i,j)$
Correlation	$\sum_i \sum_j \frac{(i-p_x)(j-p_y)G_d(i,j)}{\sigma^2}$
Inverse different moment	$\sum_i \sum_j \frac{G_d(i,j)}{1+(i-j)^2}$
Angular second moment	$\sum_i \sum_j G_d(i,j)^2$

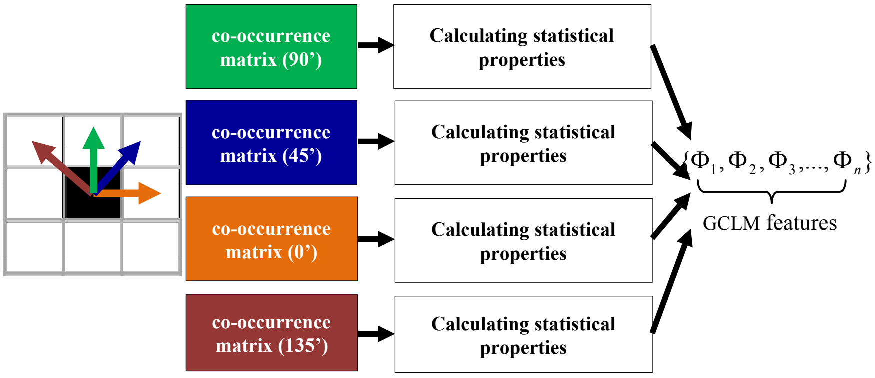


Figure 2: Examples of calculating GLCM features: directions ($\theta = \{0^o, 45^o, 90^o, 135^o\}$) and the displacement ($d = 1$ pixel) used to generate GLCMs

In addition to textural features derived from spatial relationships of gray values of pixels in image tiles, texture descriptors by taking into account macro patterns of pixels in image tiles are utilised, i.e. Local Binary Patterns (LBPs) (Ojala, Pietikainen et al. 1994). LBPs are feature extraction methods that capture the proportion of the micro patterns. The conventional LBP operator is applied to every pixel p by examining the eight neighbours and generating binary representation of pixels by thresholding gray values of the neighbouring pixels with the value of p. Therefore, this allows us to construct an eight-digit binary number, $b_1, b_2, \ldots, b_8$ where $b_i = 0$ if the intensity of the i^{th}neighbour is less than, or equal to, that of p, otherwise $b_i = 1$, and this denotes binary patterns or representation. The new macro value is assigned to p by converting binary representation to decimal number. For a given tile, therefore, we can generate a histogram of the converted decimal numbers. LBPs can generate 256 features to represent a texture descriptor of areas. The computation of the LBP features for a given pixel p is illustrated in Figure 3.

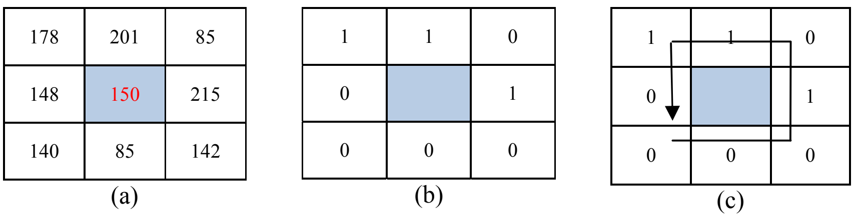


Figure 3 : Examples of computation of the LBP (a) original gray intensities of a given pixels and its neighbours; (b) binary representations of eight-neighbours and (c) a value generated from the binary representation: $(00001011)_2 = 11$

The final textural descriptors used in this work are derived from grayscale statistics. Average grayscales (μ_g) and standard deviation (σ_g) are calculated from image windows and used as a texture feature for classification.

C. Classification

In the previous section, we present the techniques of extracting discriminate descriptor of pixels in order to perform classification. To achieve the task, given an image I comprising of pixel $\mathrm{P} = \mathrm{p}_1, \ldots, \mathrm{p}_\mathrm{N}$ (where N is a number of pixels), we wish to evaluate $\Phi: P \times C \rightarrow K = \{1,2, \ldots, k\}$ $k =$ number of classes and Φ is called a classifier. We define the set of the feature vectors of pixels $x_1, x_2, .., x_N: N$ 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:

$$p(x; a_k, S_k, \pi_k) = \sum_{k=1}^{m} \pi_k p_k(x), \;\; \pi_k \geq 0, \;\; \sum_{k=1}^{m} \pi_k = 1, \tag{3}$$

$$p_k(x) = \varphi(x; a_k, s_k) = \frac{1}{(2\pi)^{\frac{d}{2}} |S_k| 1/2} \exp\left\{-\frac{1}{2}(x - a_k)^T S_k^{-1}(x - a_k)\right\}, \tag{4}$$

where m is the number of mixtures, p_k is the normal distribution density with the mean a_k and covariance matrix S_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:

$$L(x, \theta) = logp(x, \theta) = \sum_{i=1}^{N} log\left(\sum_{k-1}^{m} \pi_k p_k(x)\right) \rightarrow \underset{\theta \in \ominus}{max,} \tag{5}$$

$$\ominus = \{(a_k, S_k, \pi_k) : \} \tag{6}$$

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:

$$\alpha_{ki} = \frac{\pi_{k\varphi}(x; a_k, S_k)}{\sum_{j=1}^{m} \pi_{j\varphi}(x; a_j, S_j} \tag{7}$$

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

$$\pi_k = \frac{1}{N} \sum_{i=1}^{N} \alpha_{ki}, \;\; a_k = \frac{\sum_{i=1}^{N} \alpha_{ki} x_i}{\sum_{i=1}^{N} \alpha_{ki}}, \;\; S_k \frac{\sum_{i=1}^{N} \alpha_{ki} \Gamma}{\sum_{i=1}^{N} \alpha_{ki}}, \tag{8}$$

Where

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

D. Refinement Process

From the classification (using the methods explained in Section C), a Markov Random Fields (MRF) is applied to refine classification results by considering local consistency between pixels (Kindermann and Snell 1980). The MRF applied in this work is probabilistic models that consider relationships between pixels in a neighborhood to decide pixels classes. Given X and C; X is an image composing $\{p_1, \ldots, p_N\}$ as a set of pixels and $C = \{c|1, \ldots, c_N\}$ (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:

$$\mathrm{H}(X) = H_1(X) + H_2(X) \tag{9}$$

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

$$H_1(X) = \sum_{p\epsilon X} \sum_{t\epsilon\Omega p} \phi(p,t) \tag{10}$$

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 :

$$\phi(p,t) = \begin{cases} 1 & \text{if } p \text{ and } t \text{ are in the same class} \\ 3 & \text{otherwise;} \end{cases} \tag{11}$$

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, Veksler et al. 2001).

EXPERIMENT AND RESULTS

The previous sections explain the techniques to generate feature sets from image windows (i.e. color intensity and texture-based) and a technique for performing an unsupervised learning method to classifying image data before a post-processing can be conducted to improve classification results. This section provides the details of experiments and results of land use classification using the proposed technique. To evaluate the proposed technique, 5 sub-images are collected from the map images. Images are divided into overlapping windows (15x15 pixels) at the center of the pixel in images. Each image window is used to generate features as discussed in Section B. Pixels in images are classified into 5 classes (according to the land types defined in this work) i.e. (i) water, (ii) forest, (iii) undefined areas, (iv) urban areas and (v) agricultural areas. To achieve this task, the EM algorithm is implemented. A separate set of sub-images (from the map image) are collected and used to initialize the EM parameters (a_k and s_k). Classification is performed and results are used to carry out refining process using a MRF (explained in Section D). To evaluate the performance of the proposed technique, classified images are compared to the ground images provided by manual digitization (see Section Image data for details). Accuracies are defined as corrected classified pixels compared to the ground truth images. Experiments are conducted and results are shown on Table 2.

Table 2: Results of land use classification using the proposed method.

Features	EM		EM and MRF	
	Acc.(%)	Kappa.	Acc.(%)	Kapp.
RGB	0.35	0.48	0.39	0.41
Gray Level	0.33	0.49	0.37	0.43
Lab*	0.36	0.47	0.40	0.48
Projected RGB using PCA	0.39	0.46	0.41	0.47
Grayscale co- occurrence matrices	0.42	0.36	0.42	0.36
Local Binary Patterns	0.44	0.42	0.47	0.45
Grayscale co- occurrence matrices + Gray Level	0.33	0.52	0.33	0.51
Grayscale Statistics	**0.47**	**0.53**	**0.52**	**0.61**

Table 2 presents the results of classification performed using 8 different feature subsets using EM algorithm and the refinement method using a MRF. The results presented in Table 2. show that grayscale statistics provides promising results in classifying lands in map images using the EM algorithm and the MRF (accuracy of 52% ±11). Color intensity generally result lower accuracy than textural features. The best performance of color intensity is obtained by using the projected RGB (accuracy of 41%±9). Compared to the textural feature using grayscale statistics, the textural features is significantly statistical better than the projected RGB ($p = 0.003$). In addition, the performance of classification is improved the refinement process using the MRF. With grayscale statistics, the classification accuracy is significantly statistical improved using the refinement process from 47%±9 to 52%±11 (p=0.01) The overall comparison of classification accuracy is depicted in Figure 4.

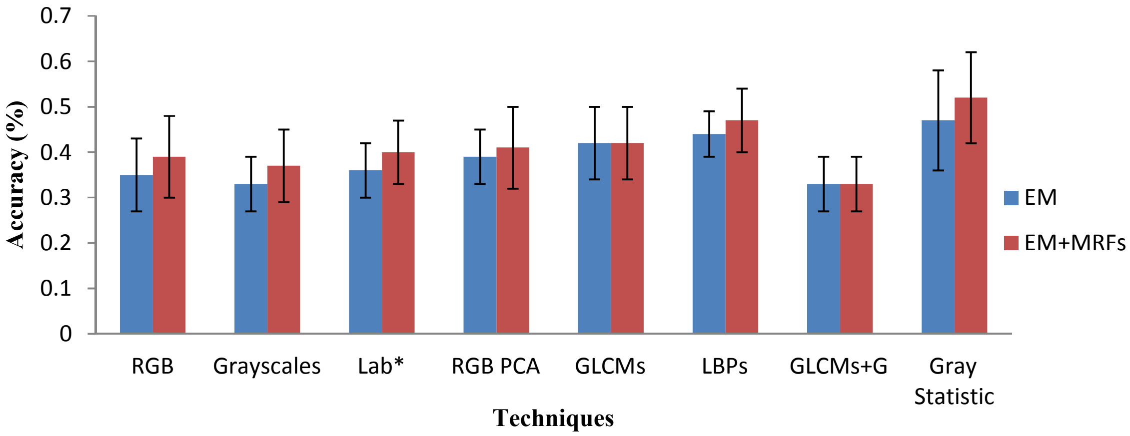


Figure 4 : Comparison of different feature subset using the unsupervised learning method.

Figure 5 show the performance of classification of each land types using the proposed techniques with the textural features derived from greyscale statistics. The results show that water is the proposed technique is able to classify water areas well (high precision and recall). The textual appearance of water is smooth and homogenous - see Figure 6- therefore, the textural features are able to differentiate water from other land types well. However, poor results are drown to undefined area and urban area class. These 2 land types are varied and uniform in textural appearance resulting in classified into agricultural areas, which is the majority class. As a consequence, classifying agricultural areas give high recall and low precision.

CONCLUSION

This paper present a technique for classifying lands in map images into 4 different type - i.e. (i) water, (ii) forest, (iii) undefined areas, (iv) urban areas and (v) agricultural areas - using an unsupervised learning method and texture-based features. The word is divided into 3 steps. Textural features is first generated from images before the classification is carried out based using the EM algorithm. Classification results are then used in a post-processing (refinement) step aiming at improving classification performance. The post-processing step is performed based on a graphical model method (Markov random field: MRF) that takes into an account the local interaction between pixels in images.

The experiment results show that textural features provides better Classification results compared to color intensity-based features. Grayscales statistic is the most discriminative features between feature subsets to classify lands in map images. In addition, the post-processing method improve the performance of classification.

To improve classification results, we could improve the performance of classification by incorporate information relating to local context of lands in images. This will take into account the spatial relationships between land types and use this information to improve classification results. In addition, extending to larger volume of data is also one of our future works.

Classified \ Actual	A	B	C	D	E	Precision	Recall
A	47330	4959	679	588	5307	0.80	0.80
B	6573	125625	19350	36312	100791	0.44	0.60
C	704	1924	5637	2866	154021	0.03	0.11
D	838	38088	11149	17027	103882	0.10	0.22
E	3580	37450	13898	21312	404944	0.84	0.53

(a)

Classified \ Actual	A	B	C	D	E	Precision	Recall
A	46754	1863	426	341	3341	0.89	0.79
B	7383	134159	20438	39329	100601	0.44	0.64
C	642	1217	5282	2233	128111	0.04	0.10
D	302	32476	9524	13798	80555	0.10	0.18
E	3944	38331	15043	22404	456337	0.85	0.59

(b)

A	Water
B	Forest
C	Undefined areas
D	Urban areas
E	Agricultural areas

Figure 5 : Confusion matrices presenting the performances of land use classification using the proposed technique with the textural features derived from greyscale statistics; (a) the results of classification performed by the EM algorithm and (b) the results obtained from the refinement process.

REFERENCES

Boykov, Y., O. Veksler, et al. (2001). "Fast Approximate Energy Minimization via Graph Cuts." IEEE Trans. Pattern Anal. Mach. Intell. **23**(11): 1222-1239.

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

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.

Haralick, R. M. (1979). "Statistical and structural approaches to texture." Proceedings of the IEEE **67**(5): 786-804.

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.

Jolliffe, I. T. (2002). Principal Component Analysis, Springer.

Kindermann, R. and J. L. Snell (1980). Markov Random Fields and Their Applications, AMS.

Ohkubo, A. and K. Niijima (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.

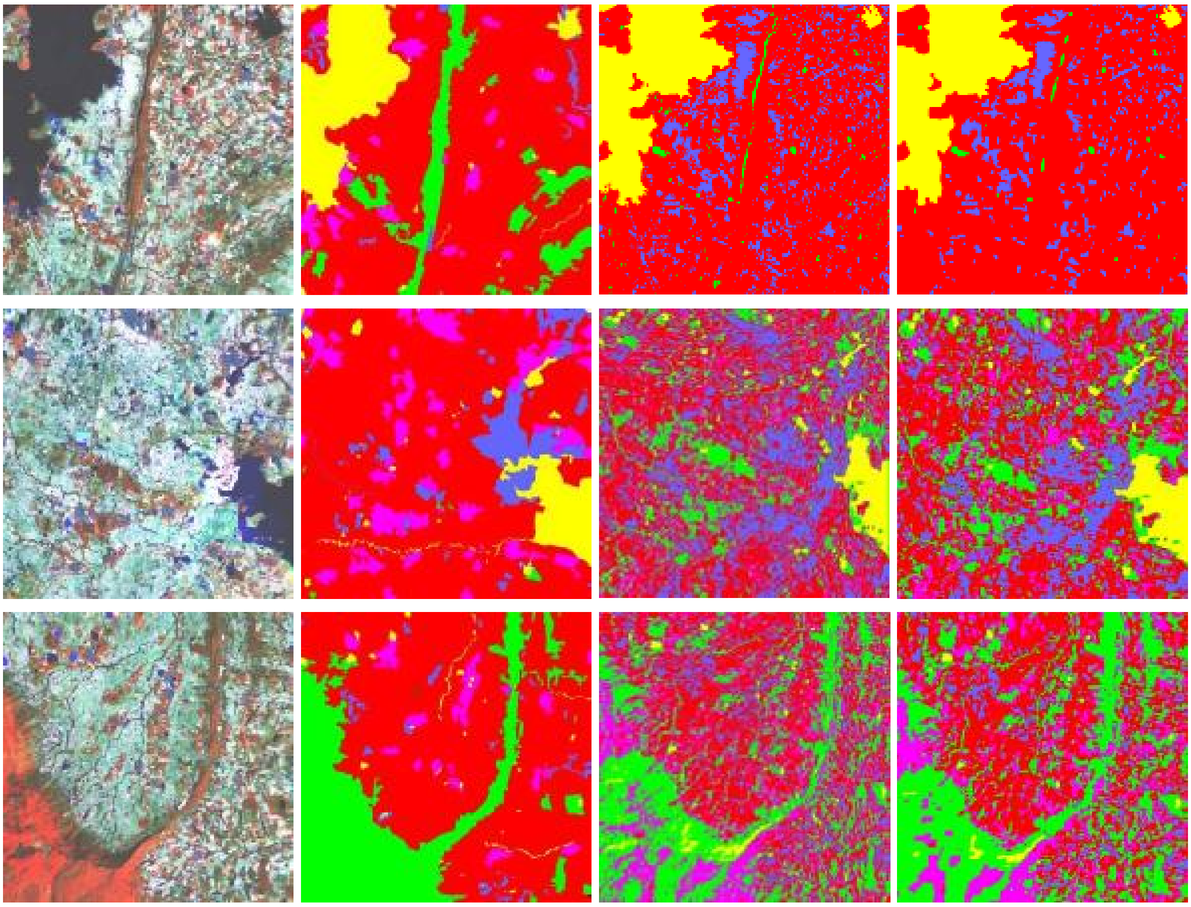
Ojala, T., M. Pietikainen, et al. (1994). Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. Pattern Recognition, 1994. Vol. 1 - Conference A: Computer Vision & Image Processing., Proceedings of the 12th IAPR International Conference on.

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.

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

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

Zabih, Y. B. a. O. V. a. R. (2001). Fast Approximate Energy Minimization via Graph Cuts. IEEE Transactions on Pattern Analysis and Machine Intelligence**:** 1222-1239.



Map Images **Manual Digitization** **EM** **EM and MRFs**

Figure 6 : Examples of classification results using the proposed classification technique with the textural features derived from greyscale statistics.