Journal of Geological Resource and Engineering 3 (2015) 127-133
doi:10.17265/2328-2193/2015.03.002

Land Use and Land Cover Mapping Based on Band Ratioing with Subpixel Classification by Support Vector Machine Techniques (A Case Study on Ngamoeyeik Dam Area, Yangon Region)

Khin Mar Yee[1], Khine Phoo Wai[1], Bak Jinhyung[2] and Choi Chul Uong[1]

1. Department of Spatial Information Engineering, Pukyong National University, Busan 608737, Republic of Korea

2. Planning and Budget Division, Ministry of Environment, Government Complex, Sejong-Si 30103, Republic of Korea

Abstract: Making map of LULC (Land Use and Land Cover) is crucially important for the sustainable development of the environment. However, the exactly data on how environmental factors influence the LULC situation at the various scales because the nature of the natural environment is naturally composed of non-homogeneous surface features, so the features in the satellite data also have the mixed pixels. Band ratioing with subpixel classification is an enhancement process in which the digital number value of one band is divided by that of any other band in the sensor array. The main objective of this study is to increase classification accuracy of LULC mapping based on band ratioing with subpixel classification by Support Vector Machines (SVMs). This process was applied with a soft approach at allocation as well as at a testing stage and to minimize the shadow and the topographic effects. The result shows the overall accuracy is increased from 61.18% of without band ratioing to 90.35% of band ratioing. The error matrix and confidence limits led to the validation of the result for LULC mapping.

Key words: Band ratioing, Support Vector Machines, change detection.

1. Introduction

Human activities are one the most widespread causes of the decrease forest land and habitat destruction and decline the natural environment. GIS (Geographic Information Systems) and RS (remote sensing) are great helpful and cost-effective tools for assessing the distribution of spatial and temporal dynamics of LULC [1-4]. Research on land cover and land use change is increasingly becoming important due to the fact that it is a primary factor for global change, reason being its interactions with biodiversity, biogeochemical cycles, climate, human activities and ecosystem processes [5].

Remote sensing data provide valuable multi-temporal data on the processes and patterns of LULC change, and GIS is useful for mapping and analyzing these patterns [6]. Image processing and making map of LULC is a key application of remote sensing data. Very updated land use and land cover information is needed at local, regional and national administrative levels for land use planning and managements.

2. Study Area

The researcher was interested in the effects of class based ratioing on land use land cover mapping. As shown in Fig. 1, the study area is Ngamoeyeik dam and its surrounding, northeastern part of the Yangon area. This area was chosen as the study area because it has many obvious land use and land cover features. It lies between latitude 17˚ 18′ 49″ and 17˚ 27′ 16″ North and longitude 96˚ 4′ 24″ and 96˚ 12′ 56″ East and area is

Corresponding author: Khin Mar Yee, Ph.D. student, research fields: remote sensing and geographic information system (RS & GIS). E-mail: kmyee2012@gmail.com.

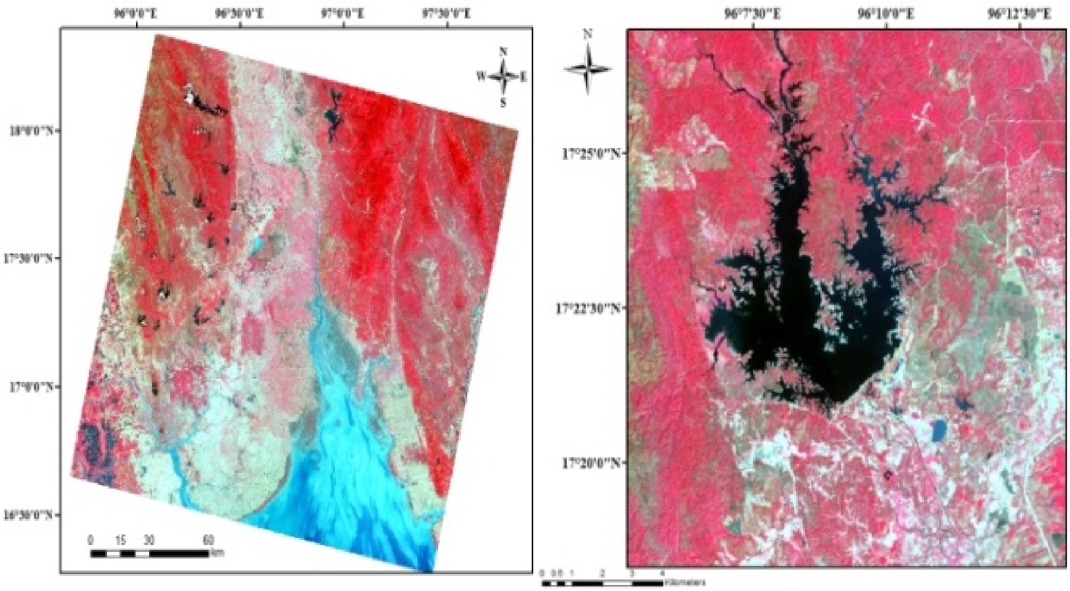


Fig. 1 Location of study area.

255.7 square kilometer. Northern and western part of the study area cover forest area, the center is Ngamoeyeik dam (water area), the eastern and southern parts occupy dry forest area, shallow water area, cultivated land and open land area.

3. Data and Methodology

Landsat 8 (OLI) data (path 132, row 48) was used to evaluate LULC changes and band ratioing with subpixel classification with cloudless area. Spatial resolution is 30 m. The image was processed by USGS (U.S. Geological Survey), UTM zone is 47 and Datum is WGS 84. This paper includes band ratioing with subpixel and SVM of supervised classification techniques. The integrative methods can be found the new findings for testing research.

Band ratioing is the very simple and powerful technique in remote sensing. Ratioing is an enhancement process in which the DN value of one band is divided by that of any other band in the sensor array. Ratioing is considered to be a relatively rapid means of identifying LULC (Land Use Land Cover) features [7-11]. Sometimes differences in brightness values from identical surface materials are caused by topographic slope and aspect, shadows or seasonal changes in sunlight illumination angle and intensity [12]. Ratio images also reduce or eliminate the effects of shadow. To minimize the effects of environmental factors, ratio may also provide unique information not available in any single band that is useful for discriminating soils and vegetation [13]. Moreover, these conditions may hamper the ability of an interpreter or classification algorithm to identify correctly surface materials or LULC in a remotely sensed multi-spectral image.

SVM (Support Vector Machine), machine learning algorithm has been proposed that may overcome the limitations of non-parametric algorithms. SVMs, first introduced by Boser et al., 1992 and discussed in more detail by Vapnik, 1995, 1998, have their roots in statistical learning theory [16] whose goal is to create a mathematical framework for learning from input training samples with known identify and predict the outcome of data points with unknown identity.

An SVM is basically a linear learning machine based on the principle of optimal separation of classes. The aim is to find a linear separating hyperplane that separates classes of interest. The hyperplane is a plane in a multidimensional space and is also called a decision surface or an optimal separating hyperplane or an optimal margin hyperplane. The linear separating hyperplane is placed between classes in such a way that it satisfies two conditions. First, all the data vectors that belong to the same class are placed on the same side of the hyperplane. Second, the distance or margin between the closest data vectors in both the classes is maximized [15]. In other words, the optimum hyperplane is one that provides the maximum margin between the two classes.

SVM is a supervised non-parametric statistical learning technique, therefore there is no assumption made on the underlying data distribution. In its original formulation [14] the method is presented with a set of labeled data instances and the SVM training algorithm aims to find a hyperplane that separates the dataset into a discrete predefined number of classes in a fashion consistent with the training examples.

SVMs have often been found to provide higher classification accuracies than other widely used pattern recognition techniques, such as the maximum likelihood and the multilayer perceptron neural network classifiers. Furthermore, SVMs appear to be especially advantageous in the presence of heterogeneous classes for which only few training samples are available. In the context of hyperspectral image classification, some pioneering experimental investigations preliminarily pointed out the effectiveness of SVMs to analyze hyperspectral data directly in the hyperdimensional feature space, without the need of any feature-reduction procedure [16, 17].

Change detection has various useful applications associated with land cover/use changes such as coastal change and urban sprawl [18]. During the study, data analyses were carried out using ENVI 4.8 and ArcMap 10.2.1 software.

The procedures of this paper consist of three phases. The first phase is a pre-field work including collection of training samples. The second phase is band ratioing and classification algorithms. The third phase is the data analysis of change detection and accuracy assessment based on post field work data collection and ground truth.

Based on the collected training sample data, the spectral profile of the image was plotted on screen to specify the maximum and minimum reflectance for different features, such as cultivated land, forest, dry forest, open land, water (dam) and shallow water. They are noted the maximum and minimum band for this class and sample band ratioing was completed. The result of ratioing images are saved separately. The resulted ratioing image layers were stacked to use the input image data for sub-pixel classification algorithm. The reconnaissance survey had been carried out at the beginning of field in order to familiarize with the study area. Total selected pixels were divided into two parts, one for training and another for testing or allocating the classifiers, so as to remove any possible bias resulting from the use of the same set of pixels for both testing and training phases. The collected training samples had been located on the images and the location of training sites are shown in Fig. 2 and Table 1.

It is the major part of analysis, and the soft classification algorithm, Support Vector Machines with radial basic algorithm has been used for both images (subset Landsat 8 OLI layer stacked image as well as subset images). There were two methods for classification, first method was the classification of image which was without ratioing image and the other was the classification of ratioing and layer stacked image. Change detection is used to correlate and compare two sets of imagery to identify changes. Using change detection statistics is to compile a detail tabulation of changes between two classification images.

The results of change detection indicated the difference between band ratioing and without band ratioing. The resulted rationing image layers were stacked to use the input image data for subpixel classification algorithm.

There are two phases for classification, first phase was the classification image which was without band ratioing and other was classification of band ratioing.

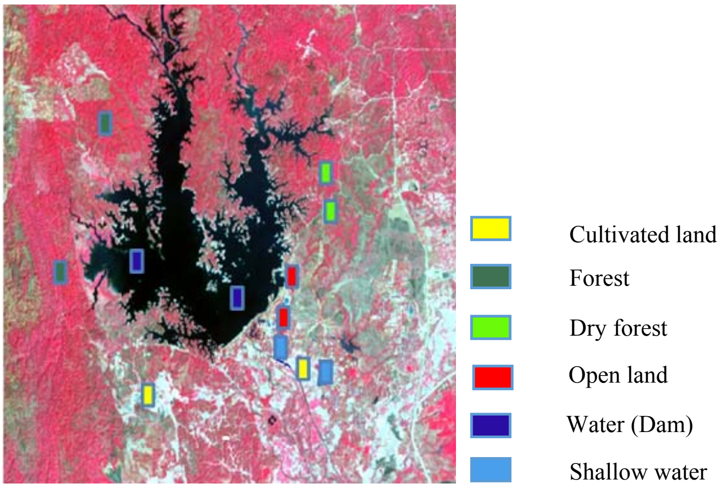


Fig. 2 Location of training samples.

Table 1 Location of training sample and collected pixel.

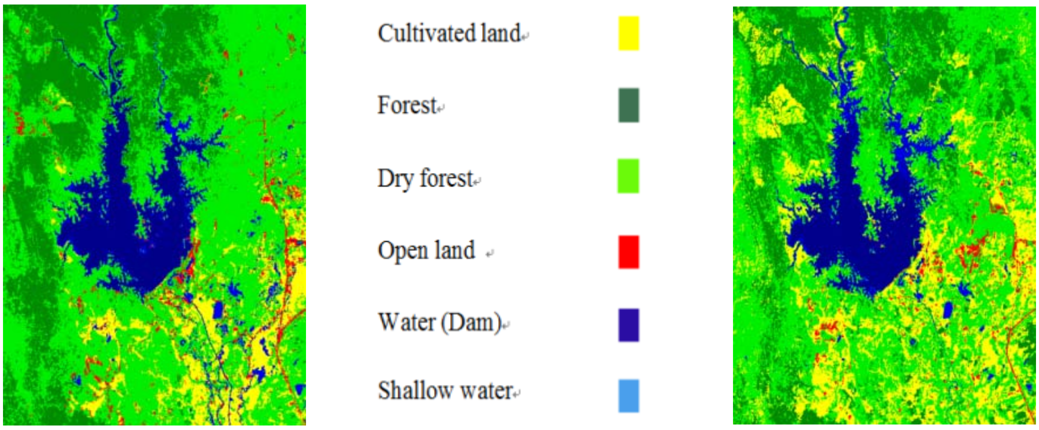
LULC class	Sample location	Sample location	pixel
Cultivated land	1,730	3,876	50
	1,566	3,901	50
Forest	1,470	3,786	50
	1,514	3,639	50
Dry forest	1,753	3,692	50
	1,767	3,729	50
Open land	1,709	3,825	50
	1,721	3,791	50
Water (Dam)	1,548	3,777	50
	1,661	3,811	50
Shallow water	1,705	3,862	50
	1,754	3,877	50

In addition to accuracy assessment is required to minimize the common sources of error in remotely sensed data. The classified images were then assessed for accuracy based on a random selection of 50 reference pixels for each time period. Overall accuracy is computed by dividing the total corrected pixels by the total number of pixels in the error matrix. This statistics indicates the probability of a reference pixel is being correctly classified and is the measurement of omission error.

3. Results and Discussion

Fig. 3 shows the spatial distribution and LULC changes of band ratioing and without band ratioing considered in this research. According to Table 2, the cultivated land is highly increased from 23.59 square kilometer to 50.26 square kilometer after band ratioing and forest area changes from 55.24 square kilometer to 48.16 square kilometer after band ratioing. The dry forest is just a little change from 134.94 square kilometer and decreased to 121.05 square kilometer. Open land was 7.77 square kilometer before band ratioing and decreased to 4.58 square kilometer after band ratioing. Water in Dam was 22.78 square kilometer before band ratioing and decreased to 20.34 square kilometer after band rationing respectively. Shallow water is 11.38 square kilometer and decreased to 11.31 square kilometer after band rationing. Using change detection statistic is to compile a detailed tabulation of changes between two classification images.

The change detected using this routine differs significantly from a simple differencing of two images.



Without band ratioing With band ratioing

Fig. 3 Change area of LULC without band ratioing and band ratioing.

Table 2 Changed area of LULC without band ratioing and band ratioing.

LULC	Without band ratioing (Sq km) (%)	With band ratioing (Sq km) (%)	Change area (Sq km) (%)
Cultivated land	23.59 (9.2)	50.26 (19.7)	26.67 (10.5)
Forest	55.24 (21.6)	48.16 (18.8)	-7.08 (2.8)
Dry forest	134.94 (52.8)	121.05 (47.3)	-13.89 (5.5)
Open land	7.77 (3)	4.58 (1.8)	-3.19 (1.2)
Water (Dam)	22.78 (8.9)	20.34 (7.9)	-2.44 (1)
Shallow water	11.38 (4.5)	11.31 (4.4)	-0.07 (0.1)

Table 3 Accuracy assessment of LULC.

LULC Class	Cultivated land	Forest	Dry forest	Open land	Water (Dam)	Shallow water	Selected sample	Omission error	Commission Error	Accuracy (%)
Cultivated land	36	13	14	13	0	0	76	52.63	61.84	47.37
Forest	13	43	20	0	0	0	76	35.53	43.42	56.58
Dry forest	20	14	34	5	0	3	76	55.26	73.68	44.74
Open land	14	0	5	51	0	6	76	32.89	23.68	67.11
Water (Dam)	0	0	0	0	65	11	76	14.48	11.84	85.53
Shallow water	0	0	17	0	9	50	76	34.21	26.32	65.79
Total	83	70	90	69	74	70	456			
Without band ratioing		→	Overall accuracy = 279/456 = 61.18%							
LULC Class	Cultivated land	Forest	Dry forest	Open land	Water (Dam)	Shallow water	Selected sample	Omission error	Commission error	Accuracy (%)
Cultivated land	71	2	3	0	0	0	76	6.58	17.11	93.42
Forest	3	67	6	0	0	0	76	11.84	11.84	88.16
Dry forest	4	8	63	1	0	0	76	17.11	15.79	82.89
Open land	6	0	3	67	0	0	76	11.84	1.32	88.16
Water (Dam)	0	0	0	0	74	2	76	2.63	7.89	97.37
Shallow water	0	0	0	0	6	70	76	7.89	2.63	92.11
Total	84	77	75	68	80	72	456			
With band ratioing		→	Overall accuracy = 412/456 = 90.35%							

So, the sample size for the image is at least 76 points for each class for the ground truth points. Error matrix is an appropriate beginning for many analytical statistical techniques. Accuracy assessment is very important to understand the output results and making good decisions. All LULC classes of without band ratioing, expect water (dam) is less than the specified accuracy of 85% at the lower confidence limit.

If the total number of correct pixels in a category is divided by the total number of pixels that were actually classified in that category, the result is the measurement of commission error. Table 3 illustrates the accuracy assessment, the error matrix, derivation of omission and commission errors.

The result of this table shows the overall accuracy is 61.18 percent for without band ratioing and 90.35 percent for with band ratioing. The classification with band ratioing, the evaluation of the errors of commission, all LULC classes exceeded (cultivated land and dry forest are nearly 85%) the specified accuracy of 85%. Dry forest did not meet the 85% criterion when errors of omission is evaluated when classification after band ratioing.

4. Conclusion

The results showed that by properly accounting for mixed pixels in all stages, higher level of accuracy could be achieved band ratioing with subpixel classification. This issues considered in this paper are extended to which effect of dimensionality of the feature space, effect of training sample size have an influence on land use and land cover classification accuracy using different classification algorithms. Remote sensing and the digital image processing were helpful in many fields and the classification technique was also advanced in current time.

By doing the class based ratioing technique for digital image classifications, the shadow and the topographic effects which can produce errors in land use and land cover classification of multispectral satellite images were minimized.

Acknowledgment

Firstly, the author gratefully acknowledges the guidance of Professor Dr. Choi Chul Uong, PE & RS lab, Spatial Information Engineering Department, Pukyong National University. Secondly, special thanks for Mr. Ahn Hoyong, Ph. D. candidate, PE & RS lab, Spatial Information Engineering Department, Pukyong National University for his valuable suggestions during the preparation of this paper. Thirdly, the author would like to thank the reviewers for their many helpful comments.

References

[1] Hathout, S. 2002. "The Use of GIS for Monitoring and Predicting Urban Growth in East and West St Paul, Winnipeg, Manitoba, Canada." *J. Environmental Management* 66 (3): 229-38.

[2] Herold, M., Goldstein, N. C., and Clarke, K. C. 2003. "The Spatiotemporal Form of Urban Growth: Measurement, Analysis and Modeling." *J. Remote Sensing of Environ* 86 (3): 286-302.

[3] Lambin, E. E., Geist, H. J., and Lepers, E. 2003. "Dynamics of Land Use and Cover Change in Tropical Regions. Annual Rev." *Environ and Resources* 28 (1): 205-41.

[4] Serra, P., Pons, X., and Sauri, D. 2008. "Land-Cover and Land-Use Change in a Mediterranean Landscape: A Spatial Analysis of Driving Forces Integrating Biophysical and Human Factors." *Applied Geograph* 28 (3): 189-209.

[5] Xiao, J., Shen, Y., Ge, J., Tateishi, R., Tang, C., Liang, Y., and Huang, Z. 2006. "Evaluating Urban Expansion and Land Use Change in Shijiazhuang, China, by Using GIS and Remote Sensing." *Landscape and Urban Planning* 75 (1-2): 69-80.

[6] Zhang, Q., Wang, J., Peng, X., and Shi, P. 2002. "Urban Built-Up Change Detection with Road Density and Spectral Information from Multi-temporal Landsat TM Data." *Int. J. Remote Sensing* 23 (15): 3057-78.

[7] Howarth, P. J., and Boasson, E. 1983. "Landsat Digital Enhancement for Change Detection in Urban Environment." *Remote Sensing of Environment* 13 (2): 149-60.

[8] Howrath, P. J., and Wickmore, G. M. 1981. "Change Detection in the Peace-Atha-Basca Delta Using Digital Landsat Data." *Remote Sensing of Environment* 11 (1): 9-25.

[9] Nelson, R. F. 1983. "Detecting Forest Canopy Changes Due to Insect Activity Using Landsat MSS." *Photogrammetric Engineering & Remote Sensing* 49 (9): 1303-14.

[10] Todd, W. J. 1977. "Urban and Regional Land Use Change Detected by Using Landsat Data." *Journal Research U.S. Geological Survey* 5 (5): 529-34.

[11] Wilson, J. R., Blackman, C., and Spann, G. W. 1976. "Land Use Change Detection Using Landsat Data." In *Proceedings of the Fifth Annual Remote Sensing of Earth Resources Conference*, 91.

[12] Jensen, J. R. 1996. *Introductory Digital Image Processing: A Remote Sensing Perspective,* 2nd edition. Saddle River, NJ: Prentice-Hall, Inc., 316.

[13] Satterwhite, M. B. 1984. "Discriminating Vegetation and Soils Using Landsat MSS and Thenatic Mapper Bands and Bad Ratios." In *Proceedings of the 50th Annual Meeting of American Society of Photogrammetry,* 479-85.

[14] Vapnik, V. 1979. *Estimation of Dependences Based on Empirical Data Nauka*. Moscow: Nauka. (in Russian) (English translation: Springer Verlag, New York, 1982)

[15] Vapnik, V., and Chervonenkis, A. 1974. *Theory of Pattern Recognition*. Moscow: Nauka.

[16] Gualtieri, J. A., and Cromp, R. F. 1998. "Support Vector Machines for Hyperspectral Remote Sensing Classification." In *Proceedings of the SPIE, 27th AIPR Workshop: Advances in Computer Assisted Recognition, Washington,* 221-32.

[17] Melgani, F., and Bruzzone, L. 2004. "Classification of Hyperspectral Remote Sensing Images with Support Vector Machines." *IEEE Transactions on Geoscience and Remote Sensing* 42 (8): 1778-90.

[18] Shalaby, A., and Tateishi, R. 2007. "Remote Sensing and GIS for Mapping and Monitoring Land Cover and Land-Use Changes in the Northwestern Coastal Zone of Egypt." *Applied Geography* 27 (1): 28-41.