LAND USE AND LAND COVER CLASSIFICATION MAPPING BASED ON BAND RATIOING WITH SUBPIXEL OF SUPPORT VECTOR MACHINE TECHNIQUES (A CASE STUDY ON NGAMOEYEIK DAM AREA, YANGON REGION)

Khin Mar Yee, Khine Phoo Wai, Choi Chou Uong

Spatial Information Engineering Department, Pukyong National University

(Kmyee2012@gmail.com)

ABSTRACT

The earth is naturally composed of non-homogeneous surface features, so the features in the satellite data also have the mixed pixels. The study area is Ngamoeyeik Dam and its surrounding, Yangon region. Ratioing 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 perform Land Use and Land Cover (LULC) mapping based on band ratioing with subpixel of Support Vector Machine (SVM) techniques. 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. This paper work processed using the data of Landsat 8 (OLI) image (Multispectral bands) and ground truth data. To get the better classification results for LULC mapping, some image processing techniques such as band ratioing was used. The ground truths were held on 76 points of each class to check with the image. They were compared with change detection and accuracy assessment. The level of accuracy assessment of this study shows that the accuracy of band ratioing is higher than without band ratioing. The error matrix and confidence limits led to the validation of the result for LULC mapping.

KEY WORDS: Band ratioing, Support Vector Machine, Accuracy assessment

Introduction

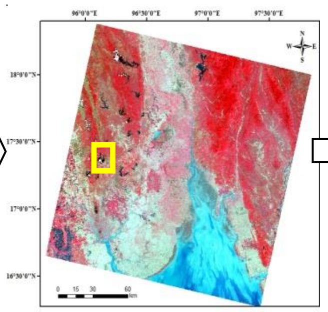
Map making of LULC is a key application of remote sensing data. Updated land use and land cover information is needed at local, regional and national administrative levels for land use planning and managements. All of the remote sensing data can be examined by the various methods. This paper includes band ratioing, SVM of supervised classification and accuracy assessment techniques. The integrative methods can be found the new findings for testing research.

Study Area

The study area is Ngamoeyeik dam and its surrounding, northeastern part of the Yangon area. It lies between latitude 17˚ 18′ 49″ and 17˚ 27′ 16″ North and longitude 96˚ 4′ 24″ and 96˚ 12′ 56″ East (Figure 1). The area is 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. Therefore, this paper includes six types of land use and land cover classes to test band ratioing with subpixel of Support Vector Machine (SVM) techniques.

Table 1. Image Information

ORIGIN	Image courtesy of the U.S. Geological Survey
REQUEST_ID	0501402065168_00027
LANDSAT_SCENE	LC81320482014038LG
FILE_DATE	2014-02-07T06:24:41Z
SPACECRAFT_ID	LANDSAT_8
SENSOR_ID	OLI_TIRS
WRS_PATH	132
WRS_ROW	48
CLOUD_COVER	1.15
IMAGE_QUALITY	9
IMAGE_QUALITY	9
MAP_PROJECTIO	UTM
DATUM	WGS84
ELLIPSOID	WGS84
UTM_ZONE	47
GRID_CELL_SIZE	30



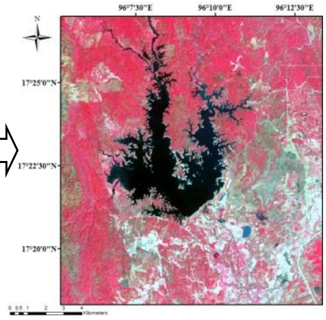


Figure 1. Location of Study Area

Aim and Objectives

The main aim of this paper is to examine to get higher accuracy using band ratioing with subpixel of Support Vector Machine Classifier. This paper studies the minimization the shadows and topographic effects which can produce errors in land use and land cover classifications of multispectral satellite images by first applying band ratioing and them performing soft classification. The next analysis is on change detection, error matrix, accuracy assessment and 95 percent confidence limits using by ground truth points.

Data and Methodology

The image satellite is Landsat 8 (OLI) data (path 132, row 48) 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 (Table 1). Among these six reflectance bands from the multispectral bands selected to use of land use and land cover classification (band 2 to band 7). During the study, data analyses were carried out using ENVI 4.8 software.

Band ratioing is the very simple and powerful technique in remote sensing. To get the better classification results for LULC mapping, Support Vector Machines (SVM) is used for a relatively new generation of techniques for classification and regression problems. An SVM is basically a linear learning machine based on the principle of optimal separation of classes. First, all the vectors data that belong to the same class are placed on the same side of the hyperplane. Second, the distance or margin between the closest vectors data in both the classes are maximized (Vapnik and Chervonenkis 1974; Vapnik 1982).

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 (soft classifiers). The third phase is the data analysis of change detection and accuracy assessment based on post field work data collection and ground truth. The following diagram illustrates the procedure of the study plan.

Image Processing

The reconnaissance survey had been carried out at the beginning of field in order to familiarize with the study area. The purpose of the field work was to collect training and test pixels for Landsat 8 (OLI) data. 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. 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. The samples and their respected spectral characteristics were shown in Table 2. The collected training samples had been located on the images and the location of training sites are shown in Figure 2 and Table 3. They are noted the maximum and minimum band for this class and sample band ratioing was completed.

Table 2. Maximum and Minimum Reflected Bands for LULC Classes

No	LULC class	Sample Location		Max Reflected Band No	Min Reflected Band No
		Latitude	Longitude		
1	Cultivated land	17.36267	96.17617	6	3
2	Forest	17.34563	96.10592	5	4
3	Dry forest	17.39489	96.19595	5	4
4	Open land	17.35793	96.17228	6	2
5	Water (Dam)	17.35399	96.16219	2	7
6	Shallow water	17.34937	96.17118	2	7

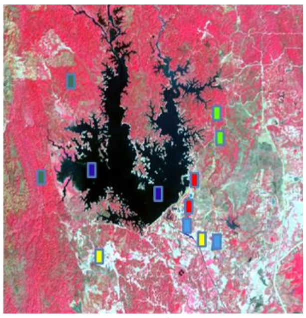
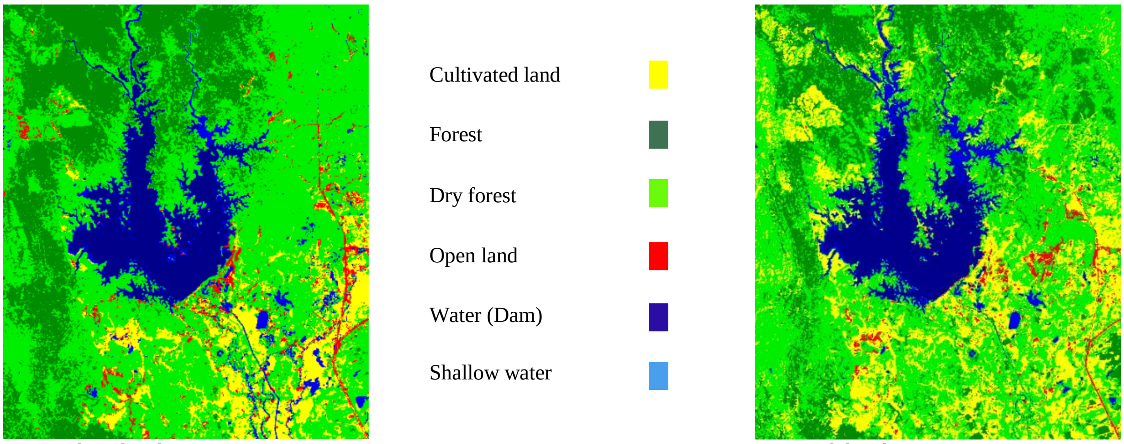


Figure 2. Location of Training Sample

Table 3. location of Training sample and collected pixel

LULC class		Sample location		pixel
Cultivated land		1730	3876	50
		1566	3901	50
Forest		1470	1786	50
		1514	3639	50
Dry forest		1753	3692	50
		1767	3729	50
Open land		1709	3825	50
		1721	3791	50
Water (Dam)		1548	3777	50
		1661	3811	50
Shallow water		1705	3862	50
		1754	3877	50

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. 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. There are two phases for classification, first phase was the classification image which was without band ratioing and other was classification of band ratioing. While the statistics report does not include a class for class image difference, the analysis focuses primarily on the without band ratioing state classification changes. The results of change detection were shown on the following table. The resulted classified images were shown in Figure 3.



Without band ratioing With band ratioing

Figure 3. Change Detection Result of without band ratioing and with band ratioing of Ngamoeyeik Dam area

Results and Discussion

Table 4 and Figure 4 show the results of the change area of the LULC of Ngamoeyeik Dam and its surrounding area. The class of total of cultivated land was 23.59 square kilometer before band ratioing stage and increased to 50.26 square kilometers after band ratioing. Forest area was 55.24 square kilometer before band ratioing and decreased to 48.16 square kilometer after band rationing respectively.

Table 4. Change area of LULC without band ratioing and band ratioing

LULC	Without band ratioing (Sq Km)	With band ratioing (Sq Km)	Change area (Sq km)	Image Difference (Percent)
Cultivated land	23.59	50.26	26.67	10.43
Forest	55.24	48.16	-7.08	2.77
Dry forest	134.94	121.05	-13.89	5.43
Open land	7.77	4.58	-3.19	1.25
Water (Dam)	22.78	20.34	-2.44	0.19
Shallow water	11.38	11.31	-0.07	0.03

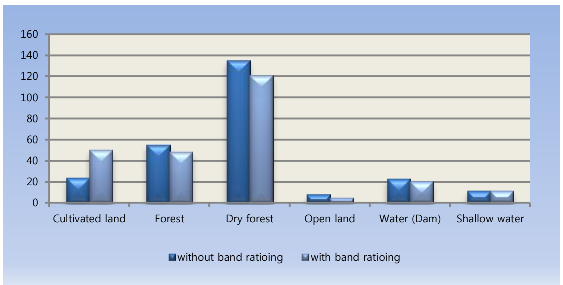


Figure 4. Change area of LULC without band ratioing and band ratioing

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 (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.

This paper is expected accuracy of 95 percent and an acceptable error of 5 percent. 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. Accuracy assessment is required to minimize the common sources of error in remotely sensed data. 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. 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 5 is the error matrix, derivation of omission and commission errors. According to the error matrix, the overall accuracy is 61.18 percent for without band ratioing and 90.35 percent for with band ratioing. The results of the six LULC classes can be examined for error of commission and omission that is shown in Table 6 based on the evaluation of the errors of commission.

Table 5. Error matrix, Derivation of Omission and Commission Errors

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 **%**										

Table 6. Individual Category Error Evaluation

without band ratioing							
LULC		Commission			Omission		
	Point Correct	n	% correct	95% confidence limits	n	% correct	95 % confidence limits
Cultivated land	36	83	43.37	32-55	76	47.37	36-59
Forest	43	70	61.43	49-74	76	56.58	45-68
Dry forest	34	90	37.78	27-48	76	44.74	33-57
Open land	51	69	73.91	63-85	76	67.11	56-78
Water (Dam)	65	74	87.84	80-96	76	85.53	80-94
Shallow water	50	70	71.43	60-83	76	65.79	55-77
with band ratioing							
LULC		Commission			Omission		
	Point Correct	n	% correct	95% confidence limits	n	% correct	95 % confidence limits
Cultivated land	71	84	84.53	76-93	76	93.42	89-98
Forest	67	77	87.01	93-95	76	88.16	80-96
Dry forest	63	75	84	75-93	76	82.89	74-92
Open land	67	68	98.53	95-100	76	88.16	80-96
Water (Dam)	74	80	92.5	86-99	76	97.37	93-100
Shallow water	70	72	97.22	93-100	76	92.11	85-99

All LULC classes of without band ratioing, expect water (dam) is less than the specified accuracy of 85 % at the lower confidence limit. 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.

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.

Acknowledgement

The author gratefully acknowledges the guidance of Professor Dr. Chou Uong Choi, PE&RS lab, Spatial Information Engineering Department, Pukyong National University, and 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.

References

Alqurashi A.F., Kumarl L., 2013. Investigation the use of Remote Sensing and GIS Techniques to Detect LULC Change, A River, School of Environmental and Rural Science University, Makkah, Saudi Arabia, (198-199), pp 193-204

Brown M., Lewis H.G., and Gunn S.R., 2000. Linear Spectral Mixture Models and Support Vector Machines for Remote Sensing, Speech and Intelligent System research, University of Southonpton, pp 1-15

Burtridge R., Trotter M., Burton B., Holdon S., 2001. Drag design by machine learning: SVM for pharmaceutical data analysis, University College, London, pp 5-14.

San B.T., Sumer E.O., Gurcay B., 2003. Comparison of band ratioing and spectral indices methods for detection alunite and kalonite minerals using aster data in Biga region, Turkey

Congalton R.G., 1991. A review of assessing the accuracy of classification of remotely sensed data. Remote Sensing Environment, pp 36-37, 35-47

Duan K.B., and Keerthi S.S., 2005. Which is the Best Multiclass SVM Method An Empirical Study, Bioinformatic Research Center, Nanyang Technological University

Jensen J.R., 1996. Introductory Digital Image Processing, A Remote Sensing Perspective, United States of America

Vapnik V.N., 1999. An Overview of Statistical Learning Theory, IEEE Transactions on Neural Networks, Vol. 10. No.5, pp 988-999.

Yang L., Xian G., Klaver J.M., and Deal B., 2003. Urban Land-Cover Change Detection through Sub-Pixel Imperviousness Mapping, Using Remotely Sensed Data