THE COMPARSION STUDY OF PADDY RICE THEMATIC MAPS BASED ON PARAMETER CLASSIFIER (MLC) AND REGIONAL OBJECT OF KNOWLEDGE CLASSIFIER (RG+ROSE)

Tsu Chiang Lei[*1] Shiuan Wan[2] Cheng-Yi Huang[3] and Jhe-Yuan Li[4]

[1] Associate Professor of Department of Urban Planning and Spatial Information, Feng Chia University, Taichung 407, Taiwan; Tel: +886-4-24511248
E-mail:tclei@fcu.edu.tw

[2] Associate Professor of Department of Information Management, Ling Tung University, Taichung, Taiwan ;Tel: +886-4-24629863;
Email:shiuan123@mail.ltu.edu.tw

[3] Graduate Student of Department of Urban Planning and Spatial Information, Feng Chia University, Taichung, Taiwan; Email:cd80031@gmail.com

[4] Graduate Student of Department of Environmental and Spatial Information Science and Technology, Feng Chia University. Taichung, Taiwan; Email:m9616939@fcu.edu.tw

KEY WORDS: Regional Object Classifier, Paddy Rice Thematic Map, Rough Set theory

ABSTRACT: In Taiwan, Rice has been a crop of global importance, which has drawn a great attraction of using remote sensing techniques for evaluating on its production. Traditionally, the paddy rice area is determined by pixel based image data to create Thematic Map. The Paddy Rice Thematic Map can be used to effectively estimate the crop production. In usual, the classification approaches in Remote sensing data is considered to apply a supervised approach (such as Maximum Likelihood Classifier, MLC). However, the pixel based procedure disobeys the rules of human image classification processes. Accordingly, a Regional Object with an ideal classifier is presented in this study. To observe the difference of our developed model and classifier, this study proposes a Region-based approach for Regional Object Classification (ROC). This new concept can effectively reduce the salt and pepper effect of classification results from very high resolution images by using a conventional pixel-based classifier. More specifically, it employs an information technique of a Rough Set Data Explorer (ROSE) to cope with the Regional Object Classification problem. Through this process, the Regional Object with ROSE classifier has overall accuracy and Kappa accuracy of region-based concepts which is better than those of the pixel-based concepts (MLC) classifier in the evaluation of paddy rice from ADS-40 images.

1. INTRODUCTION

Theoretically, the salt and pepper effect may occur while the determination on very high resolution images through the pixel based classification (such as: Maximum Likelihood Classification, MLC) process. In the past, there are two major concepts to handle the above problem. First, ancillary information of image process can provide a better help on improving the performance of classification (Haralick and Shaunmugam, 1973; Clarke, 1986; Woodcock and Strahler, 1983; Chica-Olmo and Abarca-Hernandez, 2000). On the other hand, it is widely accepted that image segmentation process can also enhance image recognition (Soille, 1999). This concept applies useful image data to transform them into Object Oriented Information (OOI). OOI can facilitate and accelerate the image classification. Moreover, OOI can handle images and then divide them into different regions which are given a reasonable representation and explanation (Blaschke and Strobl, 2001). This is so called *Region Representation*. As a result, it preferable for the process to represent the true spatial pattern of target objects rather than uniform pixels. Technically, the object-based methods contain two steps: *segmentation* and *classification* (Lu and Weng, 2007).

On the other hand, Data Mining had become a brand new approach in analyzing classification and geosciences. Due to the assessments and nonlinear character of classification problems, utilization of the Rough Set classification approach can be considered as a potentially superior approach (Goh and Law, 2003; Sinha and Laplante, 2004; Lei et al., 2008; Wan et al., 2008; Liu, 2010, Chen et al., 2011). Specifically, in this study, we use the ROSE (Rough Set Data Explorer)

(Prędki et al., 1998) classification technique to tackle the uncertainties arising from the materials and parameters involved in an observed paddy rice area. However, the ROSE model is a software system that implements basic elements of rough set theory and rule based discovery techniques (Pawlak, 1982; Ziarko, 1993). The classification model was developed at the Laboratory of Intelligent Decision Support Systems of the Institute of Computing Science in Poznan (Prędki et al., 1998). In this study, the entire paper is presented as three parts. The first part introduces study area, material and procedure. The second part presents the research method: MLC and Regional Object with ROSE. The third part discusses the outcomes and conclusions are drawn.

2. STUDY AREA, MATERIAL AND PROCEDURE

2.1 Study Area and Material

The study area is located in Yunlin County, Taichung, Taiwan (see Figure 1). The component targets include paddy rice, grass, bare land, buildings, asphalt roads, bodies of water, shadow areas and other areas which are shown on Figure 1. Our study focuses on extracting paddy rice map data (thematic map) to provide a better expression of knowledge classification. In the next section, the material is introduced. The material used in this study was obtained from a Leica Airborne Digital Sensor ADS-40 image on 05/15/2008. The image is 2402×2702, representing an area of 162.26 ha (see Figure 1).

2.2 The Plan of Study

In this study, an original multispectral image with several pieces of ancillary information is used to enhance the quality of image accuracy. Two of the classifiers, including the traditional MLC (pixel-based concept) and ROC (region-based concept) method, are used as parallel techniques for comparison. In essence, ancillary information is a feasible solution to increase the feature space by means of texture information and indicators. MLC is a statistical approach which seems quite difficult in that it eliminates the salt and pepper effect in high resolution images. Solutions are expected to be applied to obtain better classification performance. The best advantage of using rough set theory is its great attribute reduction ability (Goh and Law, 2003; Lei et al., 2008; Wan et al., 2008; Wan et al., 2010[a]). Through this process, the procedure not only effectively and objectively constructs the knowledge rules among image classification results, but it also reduces the uncertainty effective of spectral and ancillary information on remote sensing data systems.

3. INTENT ATTRIBUTES AND EXTENT ATTRIBUTES FOR OOI

Generally, the OOI data is includes two parts: 1. Intent attributes of an object 2. Extent attributes of an object.

3.1 The OOI intent for spectral and ancillary information

3.1.1 Spectral Information: In this study, the original material of this research was based on B, G, R, NIR bands of an ADS-40 image. The satisfactory level is reached by extracting the paddy rice area through the original bands or a traditional pixel-based classification process. Consequently, in addition to the original bands, some ancillary data (indicator and texture information) are also considered as classification evidence to classify categories of the image.

3.1.2 NDVI (Indicators): To determine the density of vegetation on a patch of land, researchers must observe the distinct colors (wavelengths) of visible and near-infrared sunlight reflected by the plants. Nearly all satellite Vegetation Indices are employed to quantify the density of plant growth on the earth: near-infrared radiation (NIR) minus red radiation (R) divided by near-infrared radiation plus red radiation (Bannari et al., 1995). The result of this formula is called the Normalized Difference Vegetation Index (NDVI).

3.1.3 GLCM (Texture Information): Texture features have long been used in remote sensing applications for representing and retrieving regions similar to a query region (Haralick et al., 1973; Marceau, et al., 1990). In this study, the Gray Level Co-occurrence Matrix (GLCM) and associated texture feature calculations are used as the image analysis techniques. However, in this study, NDVI, R-Homogeneity, G- Homogeneity and IR-Homogeneity were selected as indicator methods for our texture information. We also used remote-sensing data for image classification. The window size of texture methods is 3 by 3, and the preventive data of format is floating single type for easy extraction of knowledge information.

3.2 The Image Fusion through OOI Extent

In this process, the concept of "ancillary information to gray level transforms" is used to reduce the multi-dimension change to one dimension of image (Lei et al., 2009). Some of the same gray level values of given bands in the multispectral image multiply the probability value of the same gray level for the whole image. Then, the calculation function is applied to the accumulation of digital values with regard to the different bands at same location for the whole image. Through this process, the multi information of the bands is transformed into a one-dimensional image to keep the majority of information for each category of every band. The probability values of each gray level are counted by the number of outcomes of a given image and then they are divided into the total number of pixels in the image for each band. After all the bands on the image are transformed, it is used to count the total value of the synthetic image at the same location of every band. Finally, the multi-bands are successfully transformed into one-dimensional information by using this process. However, in this study, the R, G, B, NIR, NDVI, R-homogeneity, G-homogeneity and NIR-homogeneity bands are used and they are transformed into one-dimensional gray level images.

3.3 Region-Growing (Shape)

Basically, the region growing (RG) method is among the most commonly used segmentation methods for the data extraction images (Pal and Pal, 1993). Specifically, in the first stage, it adopts *similarity* to merge the nearby pixels close to a given pixel. In the second stage, it applies *area* to group nearby regions into a larger region. The terms of *similarity* and *area* are predefined variables based on the observed objects. In this study, 167,405 sub-image patches are calculated through the RG method according to the similarity variable and area variable which are found as 25 and 25, respectively. These two values are found iteratively by trial and error.

4. BASIC PRINCIPLE OF PIXEL-BASED AND REGION-BASED CLASSIFICATION METHOD

In this study, parallel analysis is used to extract the paddy rice in an ADS-40 image by using two classification processes: (1) Maximum Likelihood Classification (2) Regional Object with ROSE Classification.

4.1 Maximum Likelihood Classification

Maximum Likelihood Classification (MLC) is a popular method for classifying pixel based images. The maximum likelihood decision rule is determined by the probability density function, which considers a pixel for a particular class (Dean and Smith, 2003). Equal probability is assumed for all classes and the input bands have normal distributions (Foody et al., 1992).

4.2 Rough Set Theory

A conventional rough set can only resolve data that is pre-classified into certain levels of groups. As a matter of fact, actual environmental data is distributed uniformly (particularly with the remote sensing data format). Hence, a Discrete Rough Set is employed as an appropriate tool to evaluate them. Particularly, the integration of Regional Object with ROSE contains three strategies which are established simultaneously.

(1) Segment process: this process relies on the region-growing method.

(2) Object formalized process: our target object was integrated with the new image (inter information) and geometric shape data by the RG (extent information) method.

(3) Object interpreted process: the paddy rice objects were expressed by various attributes from the ROSE classifier.

The category knowledge database is expected to generate some given decisions of categories for series of analysis. As aforementioned, in this study, the efficiency of pixel-based classification and region-based classification is compared.

5. RESULTS

5.1 Results of MLC

In a traditional pixel-based classifier, all the pixels embedded in an image are basically independent units without any descriptive data of spatial association rules. Consequently, the pixel-by-pixel classification results usually suffer from the salt and pepper effect. As shown in Figure 2, a substantial amount of noise distribution is present on the patches in the interior and exterior areas. Due to these noise signals, some of the street trees and highway grasslands are misclassified.

5.2 Results of Regional Object with ROSE

Two main steps are designed for the analysis. First, available samples are selected. Second, those samples are used as materials to analyze through ROSE classification. In this study, Regional Object with ROSE is based on Boolean operation, and the best fictitious cutting point is selected as the threshold for this attribute. Thus, the data in each attribute is discretized into two levels. Then, the program builds the information table and the core attributes are found mathematically. In this study, the attribute reduction processes are conducted to reduce the raw data into four *core* attributes: (1) IR, (2) NDVI, (3) G-Homo, and (4) IR-Homo. Then, the rules can be formulated as:

$$\begin{cases} \textbf{\textit{IF}} \quad \text{IR} \leq 59.09 \text{ and IR-Homo} \geq 88.02 \text{ and NDVI} \geq 165.96 \\ \textbf{\textit{Then}} \quad \text{Decision= 1 (Paddy Rice)} \\ \textbf{\textit{IF}} \quad \text{IR>59.09} \qquad \text{or} \\ \text{IR-Homo<88.02} \qquad \text{or} \\ \text{NDVI<165.96} \\ \textbf{\textit{Then}} \quad \text{Decision= 2 (Non Paddy Rice)} \end{cases} \tag{1}$$

As the formulated rules are established, the training samples (81 paddy rice versus 70 non-paddy rice) are excluded from the entire image. To be more specific, 167,405 patches are embedded into the formulated rules. The outcomes are shown in Figure 3. The Regional Object with ROSE employs the characteristics of “shape” + “geometry” + “object”. Hence, here is less salt and pepper effect than there is in MLC (See Figure 2). To achieve a comprehensive rice paddy field structural result, region-based classification is a much better approach, when compared to traditional pixel-based classification.

5.3 Results of Accuracy Evaluation

Table 1 shows the classification results for paddy rice and non-paddy rice when using the MLC and Regional Object with ROSE method. There are 400 sample points selected randomly, including 150 sample points of paddy rice and 250 sample points of non-paddy rice. The overall accuracy and Kappa value of Regional Object with ROSE (96.50% and 0.9223; see Table 1 (b)) is better than the conventional MLC (92.50% and 0.8358; see Table 1 (a)) in the evaluation of paddy rice and non-paddy rice for an ADS-40 image. Furthermore, omission errors (7.35%) and commission errors (13.70%) were obtained with the MLC of paddy rice. With the Regional Object with ROSE method, the omission errors and commission errors were 4.41% and 5.8%, respectively. The above phenomenon is further explained by the fact that the paddy rice patch may include some uncertainty information. Comparing the divergence of the two research methods, the Regional Object with ROSE successfully eliminates the uncertainty information. The Regional Object with ROSE method greatly decreases the commission errors from 13.70% to 5.8%. The outcome of accuracy verification is much higher than that of the pixel-based classification (MLC) result.

6. SUMMARY AND CONCLUSIONS

In this study, Regional Object Classification (ROC) is an innovative solution to analyze paddy rice patch extraction which is generally analyzed by pixel-based classification. The results indicate that the overall accuracy with the Regional Object with ROSE method increases from 92.50% to 96.50% and the “salt and pepper” effect is dramatically reduced. Three observations can be drawn as follows:

1. The Regional Object with ROSE method performs better for the paddy extraction of overlapped land cover in a high resolution image.
2. Regional Object with ROSE and MLC have very different outcomes. Regional Object with ROSE uses a prominent classifier with core factors and thresholds. It can reinforce the extent (such as shape, area and so on) of the study objects. Although pixel-based methods and Regional Object with ROSE both can use ancillary information in their classification, Regional Object with ROSE can record the integrity of regional data. This allows for more reliable performance than the pixel-based methods are capable of. Further, Regional Object with ROSE can improve the relations among different pixels; therefore, higher accuracy of classification outcomes can be attained. In essence, Regional Object with ROSE can avoid the fragile parts of an image which are most likely to induce salt and pepper effects.
3. Regional Object with ROSE is capable of providing important ancillary information for paddy rice. More specifically, the record of each object includes clearly extracted information (spectral, texture, indicator, and so on), offering better material to determine the classification outcomes. Conversely, pixel-based classification does not retain the records of that information which may be profoundly applied for further analysis. For instance, the ancillary information may depict the change of the target objects (such as paddy rice) in different scenarios. If this process can be retained, it can help in the detection (objects+ ancillary information) of the variations of change of the paddy rice.

Hence, the major difference between region object classification and pixel classification is that Regional Object with ROSE can provide comprehensive internal and external OOI of paddy field images. In other words, the classification rules can be formulated to provide information with regard to the single objects of an intricate image. More specifically, a prominent classifier with Regional Object with ROSE can tackle the objects not only by looking at the extent attributes of an object but also by analyzing the intent attributes of an object. As a result, compared to the dispersed results from the pixel-based classification algorithm, Regional Object with ROSE offers great value due to the abundant information it provides. It can provide useful factors to construct “object rules” which are better than pixel-based + MLC.

REFERENCES

Bannari, A., Morin, D., Bonn, F., and Huete, A. R., 1995, A review of vegetation indices, *Remote Sensing Reviews*, 13, pp. 95-120.

Blaschke, T., and Strobl, J., 2001, What’s wrong with pixels? Some recent developments interfacing remote sensing and GIS. GeoBIT/GIS, Vol. 6, pp. 12-17.

Chen, Y. M., Miao, D. Q., Wang, R. Z., and Wu, K. H., 2011, A rough set approach to feature selection based on power set tree, *Knowledge-Based Systems*, Vol. 24 (2), pp. 275-281.

Chica-Olmo, M. and Abarca-Hernández, F., 2000, Computing Geostatistical Image Texture for Remotely Sensed Data Classification, *Computers & Geosciences*, 26(4), 373-383.

Clarke, K. C., 1986, Computation of the Fractal Dimension of Topographic Surfaces using the Triangular Prism Surface Area, *Computer & Geosciences*, 12(5), 713-722.

Dean, A. M., and Smith, G. M., 2003, An evaluation of perparcel land cover mapping using maximum likelihood class probabilities, *International Journal of Remote Sensing* 24, (14): 2905-2920.

Foody, G. M., Campbell, N. A., Trood, N. M., and Wood, T. F., 1992, Derivation and application of probabilistic measures of class membership from the maximum likelihood classification, *Photogrammetric Engineering & Remote Sensing*, Vol. 58(9), pp. 1335-1341.

Goh, C. and Law, R., 2003, Incorporating the Rough Sets Theory into Travel Demand Analysis, *Tourism Management*, 24, pp. 511-517.

Haralick, R.M., Shaunmmugam, K. and Dinstein, I., 1973, Textural Features for Image Classification, *IEEE Transaction System, Man, and Cybernetics*, 67, 786-804.

Hojjatoleslami, S. A. and Kittler, J., 1998, Region growing: A new approach, *IEEE, Image Processing*, vol. 7, no. 7, pp. 1079-1084.

Lei, T. C., Wan, S., and Chou, T. Y., 2008, The Comparison of PCA and Discrete Rough Set for Feature Extraction of Remote Sensing Image Classification – A Case Study on Rice Classification, Taiwan, *Computational Geosciences*, vol. 12, pp. 1-14.

Lei, T. C., Wan, S., Li, J. Y., Yeh, H. C., 2009, *Using Regional Object Classification Model to Extract Paddy Rice Patch from ADS- 40 Image*, 30th Asian Conference on Remote Sensing (ACRS2009), Beijing, China.

Liu, G. L., 2010, Rough set theory based on two universal sets and its applications, *Knowledge-Based Systems*, Vol. 23(2), pp. 110-115 .

Lu, D., and Weng, Q., 2007, A survey of image classification methods and techniques for improving classification performance, *International Journal of Remote Sensing*, Vol. 28(5), pp. 823-870.

Marceau, D. J., Howarth, P. J., Dubois, J. M. and Gratton, D. J., 1990, Evaluation of the Grey-lvel Co-occurrence Matrix Method for Land-cover Classification using SPOT Image, *IEEE Transactions on Geoscience and Remote sensing*, 28(4), 513-519.

Pal, S. K., and Mitra, P., 2002, Multispectral image segmentation using the rough-set-initialized EM Algorithm, *IEEE Transactions on Geoscience and Remote Sensing*, 40(11), pp. 2495–2501.

Pal, N. R., and Pal., S. K., 1993, A review on image segmentation techniques, *Pattern Recognition*, Vol. 26(9), pp. 1277-1294.

Pawlak, Z., 1982, Rough sets, *Int. J. Computer and Information Sci.*, 11, 341-356.

Prędki, B., Słowiński, R., Stefanowski, J., Susmaga, R., and Wilk, S., 1998, ROSE - Software Implementation of the Rough Set Theory, L. Polkowski and A. Skowron (Eds.): RSCTC'98, LNAI 1424, pp. 605-608, Berlin.

Sinha, D. and Laplante, P., 2004, A Rough Set-based Approach to Handling Spatial Uncertainty in Binary Images, *Engineering Applications of Artificial Intelligence*, 17, 97-110.

Soille, P., 1999, *Morphological image analysis: Principles and applications*, Springer-Verlag Telos, ISBN-3540656715.

Wan, S., Lei, T. C., Huang, P. C., Chou, T. Y., 2008, The Knowledge Rules of Debris Flow Event: A Case Study for Investigation Chen Yu Lan River, Taiwan, *Engineering Geology*, (98):102-114.

Wan, S., Lei, T. C., 2009, A Knowledge-based Decision Support System to Analyze the Debris-Flow Problems at Chen Yu-Lan River, Taiwan, *Knowledge-Based Systems*, (22): 580-588.

Wan, S., Lei, T. C., Chou, T. Y., 2010a, A novel data mining technique of analysis and classification for landslide problems, *Natural Hazards*, doi:10.1007/s11069-009-9366-3.

Woodcock, C. E., Strahler, A. H. and Jupp, D. L. B., 1988, The Use of Variogram in Remote Sensing, *Remote Sensing of Environment*, 25, 323-348.

Ziarko, W., 1993, Analysis of Uncertain Information in the Framework of Variable Precision Rough Sets, *Foundations of Computing and Decision Sciences*, Vol. 18, No. 3-4, 381-396.

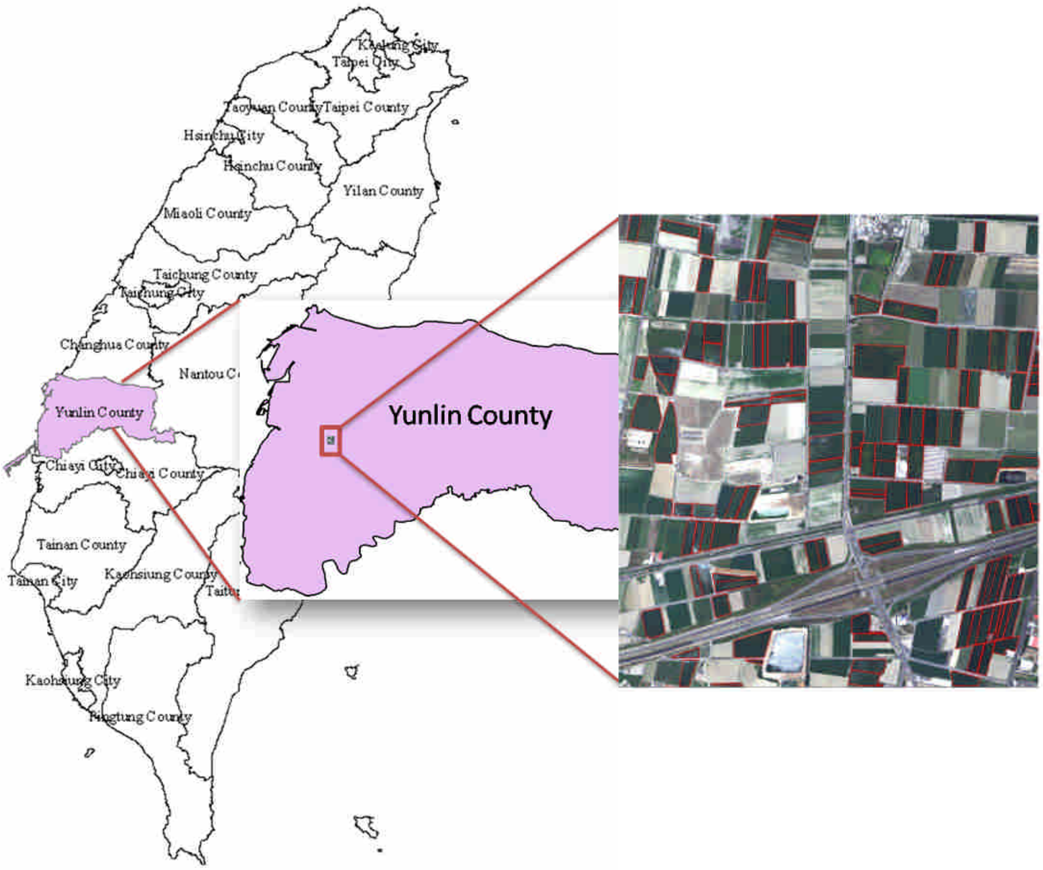

Fig1. The study area of Yunlin County, Taiwan

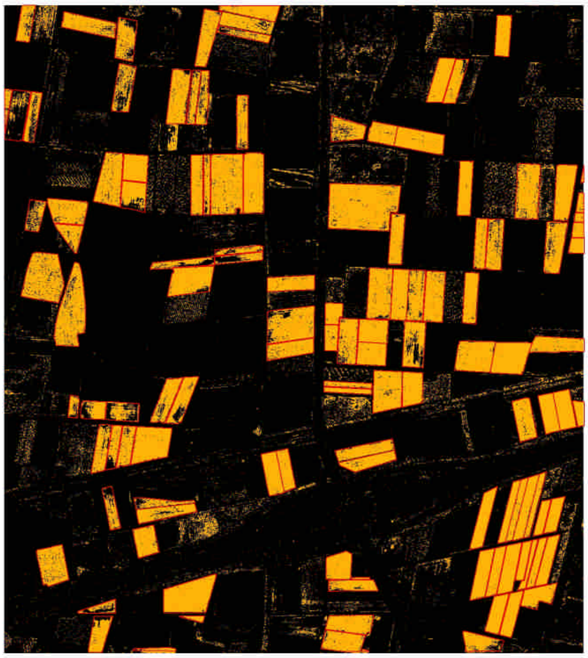

Fig. 2 The pixel based classification result
by MLC method
(Red line is paddy rice of GIS Map)

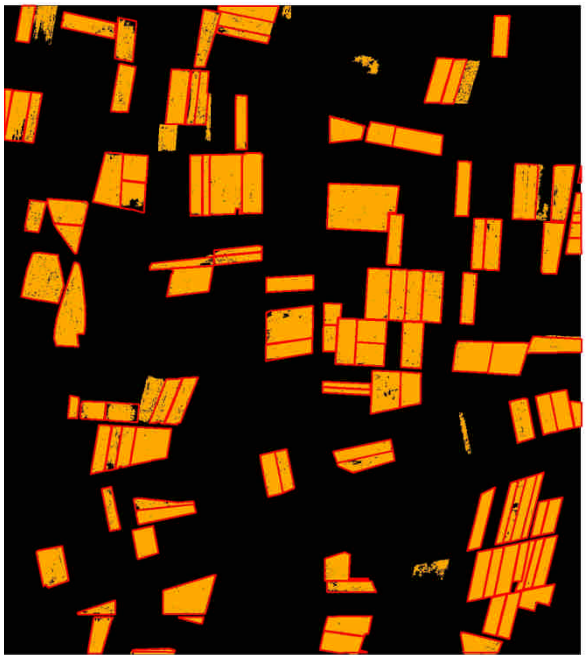

Fig. 3 The region based classification result
by Regional Object with ROSE method
(Red line is paddy rice of GIS Map)

Table 1. The Error Matrix result of pixel-based concepts (MLC) and Regional Object with ROSE.

(a) MLC Result

Map Class / Ground Reference	Paddy Rice	Non-paddy Rice	Total	User's Accuracy
Paddy Rice	126	20	146	86.30%
Non-paddy Rice	10	244	254	96.06%
Total	136	264	400	
Producer's Accuracy	92.65%	92.42%		92.50%
Overall Kappa Statistics = 0.8358				

(b) Regional Object with ROSE Result

Map Class / Ground Reference	Paddy Rice	Non-paddy Rice	Total	User's Accuracy
Paddy Rice	130	8	138	94.20%
Non-paddy Rice	6	256	262	97.71%
Total	136	264	400	
Producer's Accuracy	95.59%	96.97%		96.50%
Overall Kappa Statistics = 0.9223				