# AUTOMATIC LAND-COVER CLASSIFICATION OF LANDSAT IMAGES USING FEATURE DATABASE IN A NETWORK

G. W. Yoon*, S. I. Cho, G. J. Chae, J. H. Park

ETRI Telematics Research Group, Daejeon, Korea – (gwyoon, chosi, cbase, jhp)@etri.re.kr

**KEY WORDS:** Classification, Land cover, Database, Networks, Segmentation, Object, Software

**ABSTRACT:**

In order to utilize remote sensed images effectively, a lot of image classification methods are suggested for many years. But the accuracy of traditional methods based on pixel-based classification is not high in general. And, in case of supervised classification, users should select training data sets within the image that are representative of the land-cover classes of interest. But users feel inconvenience to extract training data sets for image classification. In this paper, object oriented classification of Landsat images using feature database is studied in consideration of user's convenience and classification accuracy. Object oriented image classification, currently a new classification concept, allows the integration of a spectral value, shape and texture and creates image objects. According to classification classes, objects statistics such as mean value, standard deviation, tasselled cap transformation and band ratio component were constructed as feature database. The feature of seven classes (Rural, Forest, Grass, Agriculture, Wetland, Barren, Water) was constructed in this study, it will be served in a network to user for image classification training data sets. Proposed method will be higher classification accuracy than that of traditional pixel-based supervised classification and gives convenient environment to users.

## 1. INTRODUCTION

The remote sensing technology is currently being offered a wide variety of digital imagery that covers most of the Earth's surface. This up-to-date image data is a promising tool for producing accurate land cover maps. To maximize the benefit of such data, automatic and efficient classification methods are needed. To achieve this objective, pixel-based classification has been extensively used for the past years. Currently the prospects of a new classification concept, object-based classification, are being investigated. Recent studies have proven the superiority of the new concept over traditional classifiers (Each, 2003; Darwish, 2003; Mitri, 2002; Niemeyer, 2001; Sande, 2003). The new concept's basic principle is to make use of important information (shape, texture and contextual information) that is present only in meaningful image objects and their mutual relationships.

In order to obtain image objects, classification software is developed by ours. It gives convenient environment to non-specialists, because operated automatically. And, feature database is constructing for automatic land cover classification. Feature database has information of seven class (water, rural, barren, wetland, grass, forest, agriculture) features in Landsat images. Proposed method will be higher classification accuracy than that of traditional pixel-based supervised classification and gives convenient environment to non-specialist users.

## 2. OBJECT ORINTED CLASSIFICATION

The object oriented classification concept is that important semantic information necessary to interpret an image is not represented in single pixels, but in meaningful image objects and their mutual relations. Image analysis is based on contiguous, homogeneous image regions that are generated by initial image segmentation. Connecting all the regions, the image content is represented as a network of image objects. These image objects act as the building blocks for the subsequent image analysis. In comparison to pixels, image objects carry much more useful information. Thus, they can be characterized by far more properties such as form, texture, neighbourhood or context, than pure spectral or spectral derivative information (Baatz, 1999).

### 2.1 Segmentation

Adjacent, similar pixels are aggregated into segments as long as the heterogeneity in the spectral and spatial domains is minimized in this step. Neighbouring segments are fused to a new segment if the resulting heterogeneity is minimized and below a specified level. The definition of heterogeneity is flexible and consists of a trade-off between homogeneity in the spectral domain (e.g. backscatter values in various channels) and form/shape in the spatial domain. Homogeneity in the spectral domain is defined by a weighted standard deviation over the spectral channels. Homogeneity of shape depends on the ratio of an object's border length to the object's total number of pixels (compactness), and the ratio between the lengths of an object's border to the length of the object's bounding box (smoothness). Compactness is minimum for a square; smoothness is minimum if the object borders are not frayed (Benz, 2001).

### 2.2 Classification

Usually classifying means assign a number of objects to a certain class according to the class's description. Thereby, a class description is a description of the typical properties or conditions the desired classes have. The objects then become assigned (classified) according to whether they have or have not met these properties/conditions. In terms of database language

---

* Corresponding author. This is useful to know for communication with the appropriate person in cases with more than one author.

one can say the feature space is segmented into distinct regions which leads to a many-to-one relationship between the objects and the classes. As a result each object belongs to one definite class or to no class. Classic classifiers in remote sensing (e.g., maximum-likelihood, minimum-distance or parallelepiped) thereby assign a membership of 1 or 0 to the objects, expressing whether an object belongs to a certain class or not. Such classifiers are usually also called hard classifiers since they express the objects' membership to a class only in a binary manner. In contrast, soft classifiers (mainly fuzzy systems and/or Bayes classifiers) use a degree of membership/a probability to express an object's assignment to a class. The membership value usually lies between 1.0 and 0.0, where 1.0 expresses full membership/probability (a complete assignment) to a class and 0.0 expresses absolutely nonmembership / improbability. Thereby the degree of membership/probability depends on the degree to which the objects fulfill the class-describing properties/conditions (Baats, 2002).

## 3. EXPERIMENT

### 3.1 Study area and data

The study site is located around Daejeon city in middle part of the Korean peninsular as shown Figure 1. The area measures approximately 575 and comprises rural areas, agriculture areas, forest areas and different areas.

Classification for the area was performed using Landsat TM acquisitions of 13 March 2000 (Figure 1). And, large-scale (1:50,000 scale) land cover map which ministry of environment produced in Korea was used as reference map (Figure 2). It is comprised of seven classes (Rural, Forest, Grass, Agriculture, Wetland, Barren, Water) and manufactured based on 2 September 1998 of Landsat TM. Table 1 show used data in this study.
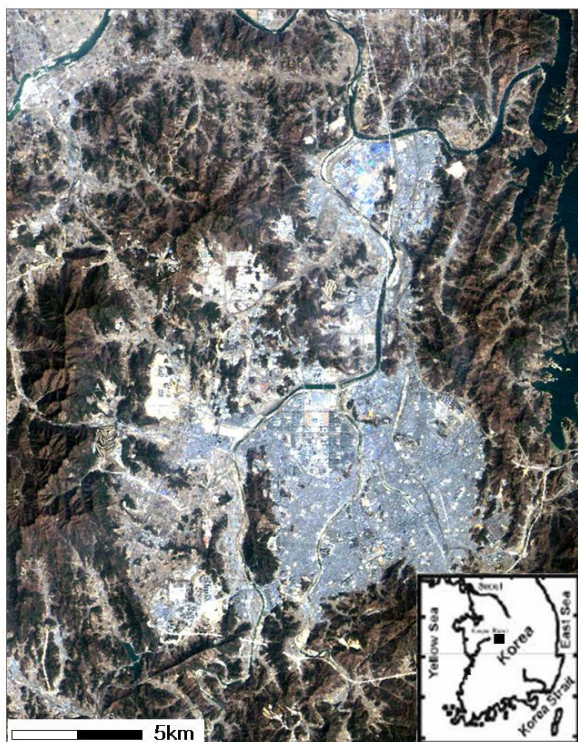


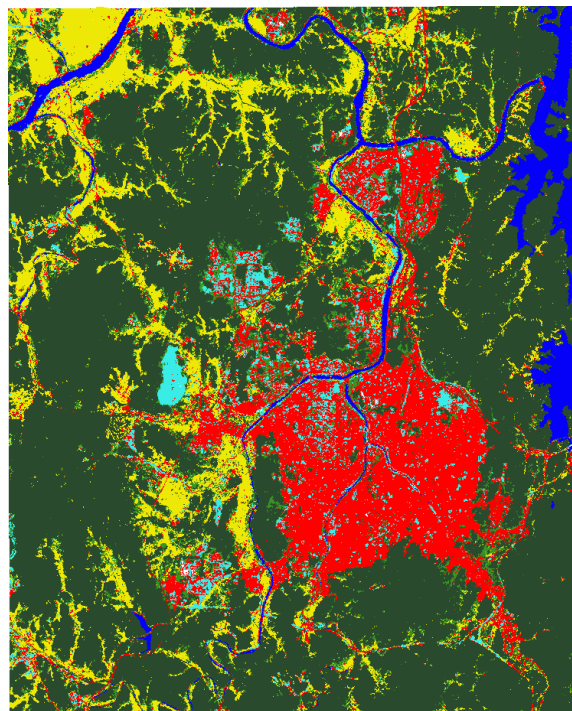Figure 1. Location of the study area and Landsat TM image



Figure 2. Land cover map as reference of study area

| | Path/row | Date | Purpose |
|---|---|---|---|
| Landsat TM | 115/35 | 13 Mar. 2000 | Classification |
| Land cover map | - | 02 Sep. 1998 | Reference |

Table 1.  Used data in this study

### 3.2 Feature database construction

We must select training data in supervised classification. As images are classified based on training data, we select training sites within image that are representative of the land cover classes of interest. The training data should be of value if the environment from which they were obtained is relatively homogeneous. However, if the land cover conditions should change dramatically across the study area, training data of partial in study area would not be representative of spectral conditions. So, we have to select training data carefully and widely in study area. And, the general rule is that if training data are being extracted from $n$ bands then >10n pixels of training data are colleted for each class. This is sufficient to compute the variance-covariance matrices required by some classification algorithms (Jensen, 1996).

It is indispensable step to extract training data in supervised classification, but it requires substantial human operations. Such substantial human operations make training data selection a time-consuming and laborious processing. The specially, as many satellite images are processing, thus operations will be troublesome more and more. In this paper, feature database is introduced to reduce a laborious processing and obtain high classification accuracy. Feature database has statistics calculated training data. As feature database constructed by specialists provide to non-specialists, they can have the advantage of convenience and accuracy.
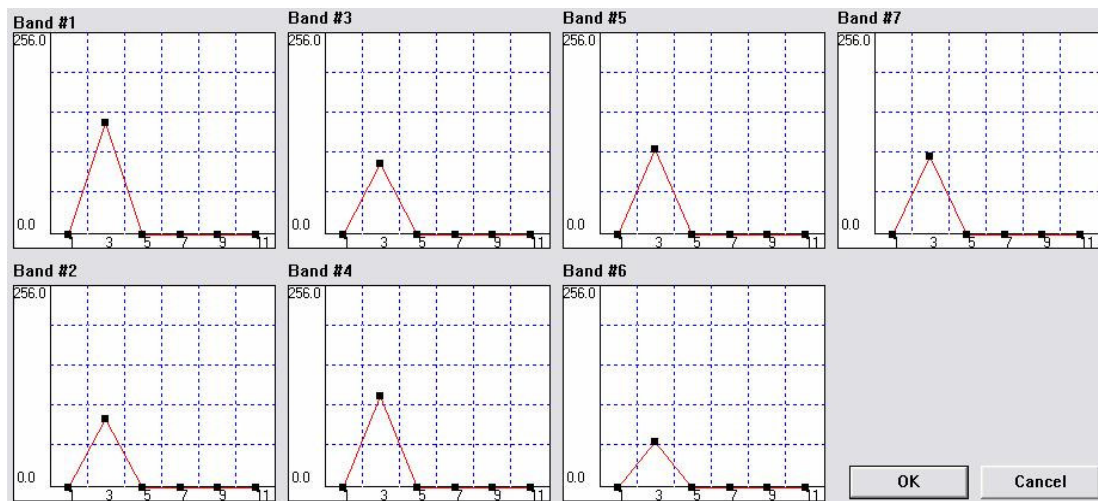
Figure 3. Brightness mean value of rural area

Land cover classification system made by ministry of environment in Korea is used in this study. It has 7 classes including rural area, forest area, grass area, agriculture area, wetland area, barren area, water area. Rural area includes residential area, commercial area, traffic and public facilities. Forest area includes broad-leaved tree area, needle leaf tree area and mixed area. Grass area includes green tract of land, graveyard and hillock. Agriculture area includes rice field, farm and arable land. Wetland area includes swamp, salt field and tidal flat. Water area includes river, lake and sea (Park, 2001).

Brightness value in pixel may be variable according to acquisition time. Difference of brightness value according to time in forest and grass area may be larger than that of other areas. So, feature extracting from images is constructed as database bimonthly (January, March, May, July, September, November). User should use feature information close to acquisition time of database for classification. In feature database, segmentation objects which are generated using two-neighbour centroid linkage region growing method (Hong, 1991) have feature information. Segmentation objects include feature information selectively and constructed as database (Table 2). Figure 3 shows brightness mean value of rural area in March for example among feature information. Feature database is constructing in nowadays.

### 3.3 Processing and result

Satellite imagery information management center (SIMC) of Korea archives past Landsat images and receives Landsat-7 ETM+ images. Using Landsat image database and reference database, we are constructing feature database for land cover classification as mentioned above. When Landsat images are classified, feature database will help users to operate few steps for land cover classification as shown figure 4. First, segmentation should be processed using several parameters (Figure 5). The segmentation process needs for minimum parameters considering users don't have profound knowledge. Level for combing means threshold of region growing. Scale means minimum size of segment. Level for merging means threshold of neighbour segment for merging. Figure 6 shows segment of Landsat image in this study. Images are classified using feature database after image segmentation.

| Brightness | Tasselled cap transformation | Band ratio |
|---|---|---|
| b1 mean<br>b1 std. | brightness mean<br>brightness std. | b2/b1 mean<br>b2/b1 std. |
| b2 mean<br>b2 std. | greenness mean<br>greenness std. | b3/b2 mean<br>b3/b2 std. |
| b3 mean<br>b3 std. | wetness mean<br>wetness std. | b4/b2 mean<br>b4/b2 std. |
| b4 mean<br>b4 std. | haze mean<br>haze std. | b4/b3 mean<br>b4/b3 std. |
| b5 mean<br>b5 std. | | b5/b4 mean<br>b5/b4 std. |
| b6 mean<br>b6 std. | | b7/b5 mean<br>b7/b5 std. |
| b7 mean<br>b7 std. | | |

Table 2. Feature lists for classification training data in database
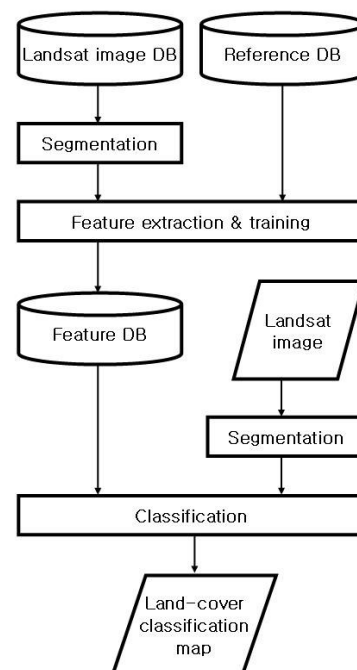


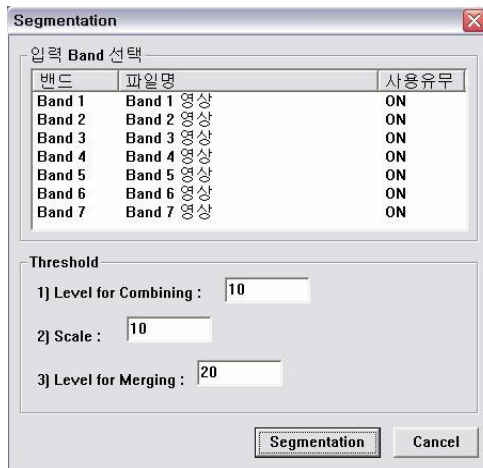Figure 4. Brief flowchart in this study

Figure 5. Segmentation dialog box in the program



Figure 6. Segmentation map of Landsat image in study area
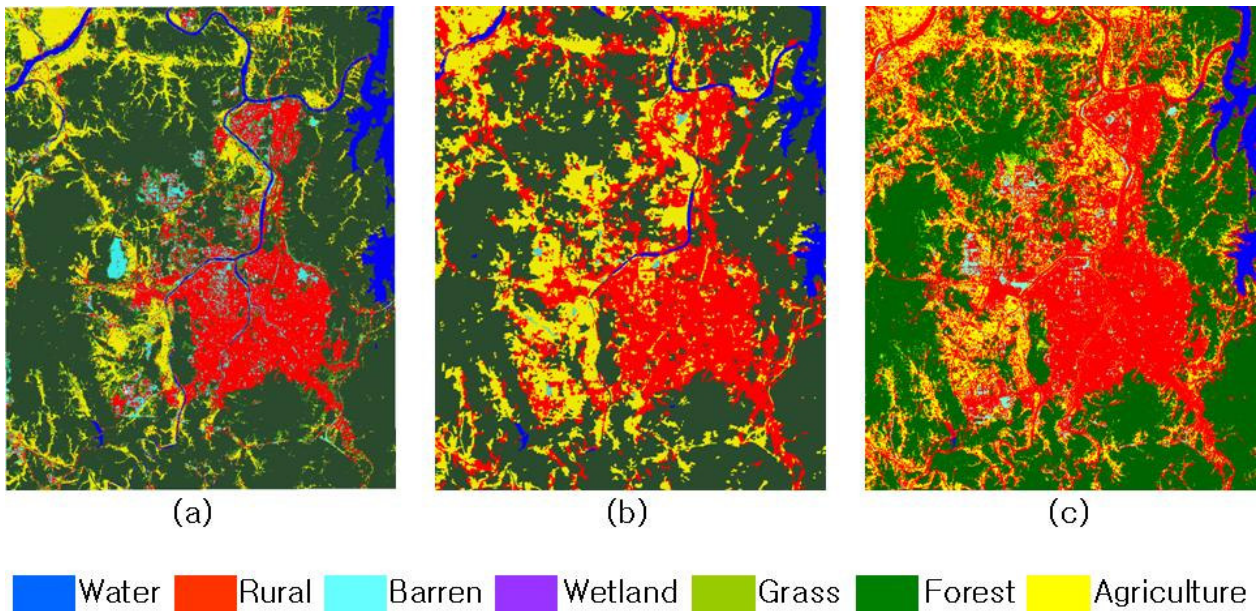


Figure 7. classification map of (a) large scale as reference (b) this study method (c) pixel based classification

That is to say, training data is replaced with feature database. So, users don't feel inconvenience to select training data sets. Figure 7 (b) shows classification result using feature database. Figure 7 (a) is large scale classification map as reference produced by ministry of environment, figure 7 (c) is pixel based classification map using Earth 2.0 software. Although classification result is extracted better using editing classes, we did barely work post-processing in view of non-specialists. Accuracy assessment is planning in the future in consideration of time and area a lot. Examining with the unaided eye, accuracy of method in this study is better than that of tradition method. It will be expected to serve convenient surroundings to users.

## 4. CONCLUSION

We must select training data in supervised classification. As images are classified based on training data, we select training sites within image that are representative of the land cover class of interest. Users don't feel inconvenience to select training data sets sometimes. So, automated classification method using feature database is proposed in this study. Feature database has statistics calculated training data. We construct statistics about brightness, tasselled cap transformation and band ratio in rural area, forest area, grass area, agriculture area, wetland area, barren area and water area in now. As a result of our developed classification software in test area, it is expected that proposed method is higher accuracy than traditional method. It will serve convenient surroundings to non-specialist users.

## 5. REFERENCES

Baatz, M. and Schape, A., 1999. Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation, *Angewandte Geographische Informationsverarbeitung XI. Beitrage zum AGITSymposium*, Salzburg.

Baatz, M., Benz, U., Dehghani, S., Heynen, M., Holtje, A., Hofmann, P., Lingenfelder, I., Mimler, M., Sohlbach, M., Weber, M. and Willhauck, G., 2002. *eCognition 3.0 USER GUID*, DEFINIENS imaging,

Benz, U., 2001. Definiens Imaging GmbH: Object Oriented Classification and Feature Detection, *IEEE Geoscience and Remote Sensing Society Newsletter*.

Darwish, A., Leukert, K. and Reinhardt, W., 2003. Image Segmentation for the Purpose Of Object-Based Classification, *IGARSS 2003*, Paris.

Esch, T., Roth, A., Strunz, G. and Dech, S., 2003. Object-oriented classification of Landsat-7 data a for regional planning purpose, *The international archives of the photogrammetry, remote sensing and spatial information sciences*, vol. XXXIV-7/W9, Regenburg.

Hong, W. H. and Kim, N. C., 1991. Image coding using HVS-based segmentation, *J. KITE*, 28(B), pp. 45-52.

Jensen, J. R., 1996. *Introductory Digital Image Processing : A Remote Sensing Perspective*. Prentice Hall, New Jersey, pp. 205-208.

Mitri, G. H and Gitas, I. Z., 2002. The development of an object-oriented classification model for operational burned area mapping on the Mediterranean island for Thasos using LANDSAT TM images, *Forest Fire Research & Wildland Fire Safety*, Millpress, Rotterdam.

Niemeyer, I. and Canty, M. J., 2001. Knowledge-Based Interpretation of Satellite Data by Object-Based and Multi-Scale Image Analysis in the Context of Nuclear Verification, *IGARSS 2001*, Sydeny.

Park, S. M., Im, J. H. and Sakong, H. S., 2001. Land Cover Classification of a Wide Area through Multi-Scene Landsat Processing, *Korean Journal of Remote Sensing*, 17(3), pp.189-197.

Sande, C. J., Jong, S. M. and Roo, A. P. J., 2003. A segmentation and classification approach of IKONOS-2 imagery for land cover mapping to assist flood risk and flood damage assessment, *International Journal of Applied Earth Observation and Geoinformation*, 4, pp. 217-229.