

## LAND COVER CLASSIFICATION METHOD ORIENTED TO GEOGRAPHIC NATIONAL CONDITIONS INVESTIGATION

CHENG Tao \*

National Geomatics Center of China, China – [chengtao@nsdi.gov.cn](mailto:chengtao@nsdi.gov.cn)

**KEY WORDS:** Land Cover, Classification, Geographic National Conditions Investigation, High Spatial Resolution Remote Sensing Imagery, Information Extraction, Method Research

### ABSTRACT:

Developing the project of geographic national conditions investigation is in order to obtain land cover change information which is caused by natural and human social and economic activities, and serve the information for government, enterprise and public. Land cover is an important method to describe the geographic national conditions information, which can truly reflect the land surface material type and its natural attribute. It has been contained in the content system preliminary scheme as an important portion. In this paper, it discusses and analyzes on the method and key technology, with according to the land cover content that geographic national conditions watches on. Then it evaluates the applicability of automatic classification method using in land cover information extraction, and comprehensively analyzes various extraction methods' maximum effectiveness. Finally, it proposes a method that is based on high spatial resolution remote sensing imagery and can be used in engineering applications, which provides a reference method for geographic national conditions investigation.

### 1. INTRODUCTION

Land cover information obtaining is the primary mission of the project of geographic national conditions investigation; it's also the basis of the project. Utilizing some classification technology and methods, we can get the range, area, and spatial distribution information of land cover, such as plough, garden plot, woodland, lawn, building, road, construction, artificial stack, bare land, water, and so on, which constitute the classification system of the project and can provide important evidence for index information's statistic and proclamation.

There are five series of global land cover products issued so far, correlative research is also developed in China. However, the scale of global land cover research is too small, and it aims at satisfying the requirement of global change and earth's mode simulation, which uses lower spatial resolution images. Secondly, the land cover classification systems in different areas are established only for the requirement of each area

individually.

Oppositely, the classification system established in the project of geographic national conditions investigation is from the point of view of monitoring the changing information of geographic objects throughout the country, which has synthetically considered the requirements of multi-areas, and has referenced the national standards, technology norms and land cover classification systems that have been issued and applied broadly. Besides, in this project, it uses middle and high spatial resolution remote sensing images to extract information.

There are many methods to obtain land cover information based on remote sensing images, such as automatic classification, automatic classification by assistant thematic data, artificial interpretation and editing, outside investigation, and so on. In which the most efficient method is automatic classification.

However, mostly successful cases are oriented to thematic objects' extraction, or belong to researches in small areas,

---

\* Corresponding author.

which can't satisfy the requirements of wide range and synthetically objects' extraction in the project of geographic national conditions investigation. The paper proposes a feasible method based on analyzing the theory of classification methods, which can provide references for the project of geographic national conditions investigation.

## 2. INFORMATION EXTRACTION'S PRINCIPLES

The project of geographic national conditions investigation defines ten first-level classes according to certain principle and theory, including plough, garden plot, woodland, lawn, building, road, construction, artificial stack, bare land and water. Of course, it also defines some second-level and third-level classes.

As mentioned in the introduction, there are many methods for obtaining land cover information. However, each method has its respective process flow. Which method and process flow will be selected is mainly decided by the factors of image's quality, land cover's type, the abundance of fundamental geography information data and thematic data and so on.

### 2.1 Principle which is decided by image's quality

(1) If the image's quality is very good, which includes that the image is multispectral, clear and has no cloud, the temporal phase of the image is suitable, we can firstly use the method of automatic classification to obtain the preliminary results, and then use the method of artificial interpretation and editing to edit and perfect the results.

(2) Inversely, if the image's quality is not good, such as the image is not clear or has cloud, or the temporal phase is not suitable, we should use the method of artificial interpretation and editing; just rely on the experiential knowledge.

### 2.2 Principle which is decided by land cover's type

(1) For the types of natural land cover, such as water and vegetation, which can be classified automatically comparatively simple, we should use the method of automatic classification firstly; then, obtain the lower classes' types of plough, garden plot, woodland, lawn and so on based on the results.

(2) For the types of artificial land cover, such as building, construction, artificial stack, and some type of natural land cover, such as bare land, which cannot be classified

automatically easily and accurately, we should firstly use the method of artificial interpretation and editing.

(3) Significantly, for some especial type of artificial land cover, such as road, which has especial geometry characters, for example, the width is uniform at each point in the same road, and the ratio of length to width is always very large. In this instance, we could use the method of automatic classification to obtain road's geometric information, and also we could use the method of artificial interpretation and editing. The two methods can be complementary for each other.

### 2.3 Principle which is decided by the abundance of fundamental geography information data and thematic data

(1) If the fundamental geography information data or the thematic data is abundant and the quality is good, we can make full use of these data for classification. Especially for the land cover types which must be obtained by the method of artificial interpretation and editing, we can firstly obtain them from fundamental geography information data or thematic data; then, we can update it on the basis of images; based on these results, we could use the method of automatic classification to obtain other land cover types' information.

(2) Inversely, if the fundamental geography information data and the thematic data are not abundant, we can firstly use the method of automatic classification to obtain the preliminary results, and then use the method of artificial interpretation and editing to edit and perfect the results.

### 2.4 Results' integration

For the results of land cover classification, no matter which method is used to obtain the information, the final results must cover the total region, whatever we divide the unit into parts of image's scene, frame or administrative district. Besides, all the polygons must not overlap in the spatial coordinate system. So, for the same unit region, the result of land cover information obtained frontally should be as a basis for the following type. The results of land cover classification must be integrated to ensure the spatial relationships are reasonable and have no loopholes.

## 3. INFORMATION EXTRACTION METHOD

Different information extraction method exhibits different

efficiency and precision. The guideline is furthest using automatic classification methods to obtain land cover information, and then using some other methods when the results' precision can't achieve the request. It's best to take on high efficiency and high precision.

Automatic classification method mainly uses the objects' characters of spectrum, texture and shape to build adaptive extraction regulations and automatically extract all levels' land cover information. The process flow of inside processing remote sensing images is:

(1) Pre-process remote sensing image

This process includes images' bands stacking, merge, ortho-rectification, and so on, which would construct fine foundation for the following processing.

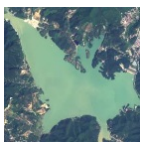
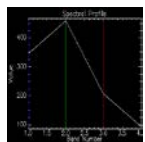
(2) Segment image

In order to obtain feature objects, firstly, we need to analyse pixels' spectral and shape characters and calculate eigenvalue for each band's spectral heterogeneity and shape heterogeneity; then, we need to define some parameters including scale, image layer weights, compactness, and segment image to feature objects by using multi-resolution segmentation arithmetic.

(3) Obtain objects' original characters in images

Objects' original characters normally include spectrum, texture, shape, spatial position, etc. They all play important role in information extraction. It's very useful to analyse the function of them in information extraction.

Take the spectral character analysis for example, we should analyse the reflectivity of land cover on the bands the image has. Based on the reflectivity, some indexes can be built for automatic classification. The following figure shows the reflectivity of water, vegetation, other artificial surface, bare land and road on the multispectral remote sensing image. It uses WorldView2 multispectral image, which has 4 bands, including blue, green, red and near-infrared in turn.

types	image	spectral	indexes
water			NDWI, ratio B, standard deviation B


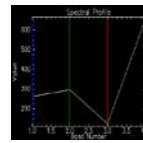

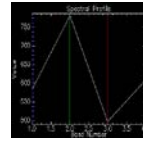

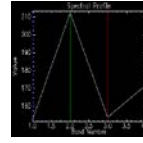

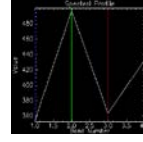
vegetation			NDVI, mean B
artificial surface			artificial surface indexes
bare land			bare land indexes
road			asymmetry, length/width, density

Figure 1. R, G, B layer stacking images and spectral profile of land cover on each band

It shows that water exhibits strong reflex and strong absorbing characters respectively in bands of blue and near-infrared, vegetation exhibits the same characters respectively in the bands of red and near-infrared, artificial surface exhibits the same characters respectively in the bands of blue and red-edge, bare land exhibits the same characters respectively in the bands of red and red-edge, and road exhibits the same characters respectively in the bands of blue and near-infrared. These characters would play an important role in the indexes building and automatic classification.

Figure 1 exhibits that some indexes can be built on the basis of spectral characters, for example, NDWI (Normal Difference Water Index) =  $(G - N) / (G + N)$ ; Ratio B =  $B / (B + G + R + N)$ ; Standard Deviation B =  $\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2}$ ; NDVI (Normal Difference Vegetation Index) =  $(N - R) / (N + R)$ ; where B, G, R, N respectively stands for the band of blue, green, red and near-infrared,  $x_i$  is the digital number of pixel  $i$  in the image of band blue,  $\bar{x}$  is the average value of digital number of all pixels in the image of band blue.

(4) Build indexes for objects' extraction

As we know, some indexes such as NDVI, NDWI and

LAI (Leaf Area Index) are valuable for extracting land cover information. Similarly, we can build indexes for objects' extraction according to all kinds of original characters of the objects, which would compose the decision regulations for information extraction.

(5) Inset assistant datum for objects' extraction

It is very useful for improving precision by importing assistant datum in classification, such as DEM (Digital Elevation Model) and DSM (Digital Surface Model), which would reduce mixing probability of objects.

(6) Select samples for classes

Select typical samples for classes which can be extracted automatic according to the objects' characters, which would become to the symbols in classification. The samples' characters would stand for the corresponding class's characters.

In the process flow of information extraction, on one hand, we could use extraction rules for automatic classification; on the other hand, we can use supervised classification model of object-oriented automatic classification method.

(7) Build extraction rules

In the process of automatic classification by computer, it is very important to build extraction rules on the basis of original characters' analysis, indexes building and assistant datum. The rules are corresponding with the characters respectively. Besides, we should set threshold for each parameter. It's better and more exact to use multi-rules than use singular rule when we extract land cover information.

(8) Extract land cover information

Extract land cover information by utilizing methods of supervised or decision-tree or object-oriented classification based on the original characters of objects, index information, and assistant datum.

(9) Analyse and evaluate the results' quality

Collect some test samples to analyse the extraction results' quality. Evaluate the precision of the results and make the results' precision achieve the requirement of the project by artificial interpretation and editing.

To sum up, the work flow of land cover information extraction based on remote sensing image is as shown below.

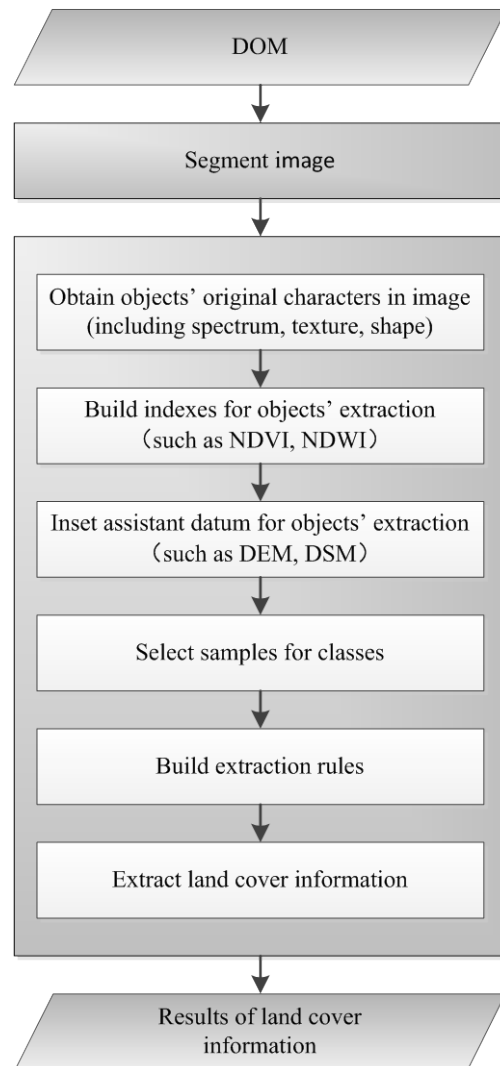


Figure 2. Work flow of automatic classification by computer

#### 4. KEY TECHNOLOGY ANALYSIS

(1) Requirement of data source

With the rapid development of information technology and satellite sensors, the class of satellite remote sensing images is increasing gradually, which provides abundant source data for geographic national conditions investigation. Therefore, it could use different spatial and spectrum resolution images for different classes' land cover information extraction.

It's suitable to use multi-spectrum images of 10 meters resolution to extract first-level classes, 5 meters resolution to extract second-level classes, and 1 meter resolution to extract third-level classes. Furthermore, the higher spectrum resolution of remote sensing images, the better to improve the results' precision. Thus, if conditions permit, it's best to use high spectrum resolution images as far as possible.

(2) Analyse objects' characters

Objects on earth have many kinds of characters. The primary task of automatic classification is to analyse these characters. At present, the main characters used in automatic classification are spectrum characters, which are the basis of texture and shape characters and are the most direct information in images. To achieve automatic classification, it could directly build decision regulations or build some indexes based on objects' characters.

### (3) Build regulations for information extraction

It's very important to build suitable decision regulations in information extraction, which would exhibit good capability in automatic classification. For example, NDVI index is always used to extract vegetation information, NDWI index is always used to extract water information. They are both built by ratio computing. The theory of this arithmetic is to find out the bands which the objects' exhibit strongest reflex and weakest reflect characters, and then put the strongest one on the numerator, the weakest one on the denominator, which can enhance the researching objects and bate the others.

Influenced by the factor of hypsography, the spectrum reflex would change. And land covers' grown area is restricted by the height above sea level, slope and aspect. Therefore, it could improve the results' precision by using assistant DEM and DSM datum.

According to the present research status, with the improvement of images' spatial resolution, the character difference between each land cover class is very tiny, which brings difficulty on building decision regulations for information extraction. The extraction result of single class always mixes the other land cover information. So, it should build detailed decision regulations as far as possible to obtain perfect results.

### (4) Select classification methods

Traditional Pixel-Based classification method based on remote sensing images is on the conditions that the spectrum information is very abundant and the difference of spectrum between different land cover classes is obvious. In high resolution remote sensing images, the land covers' construction and texture are very outstanding, but spectrum character differences between different land covers are not that outstanding. Traditional Pixel-Based method would not satisfy the present requirement for information extraction.

For satisfying the new application situation nowadays, the Object-Oriented classification method provides new solution in high resolution remote sensing images' information extraction.

This method not only utilizes spectrum characters, but fully utilizes abundant spectrum, texture, shape, and spatial position characters to extract land cover information. The core of this method is images' segmentation. Segmentation arithmetic includes multi-resolution, grey-based, texture-based, and knowledge-based and so on, in which the multi-resolution arithmetic is used mostly. This arithmetic synthetically utilizes land covers' spectrum and texture characters, calculates the spectrum and shape differences between each bands, then sets certain threshold, when all the segmentation objects' characters' value extends the threshold, the segmentation process is completed. On the base of these processes, it could build proper decision regulations to extract land cover information.

Another method named decision-tree based on expert's knowledge is also a commonly used method. This method classifies pixels mainly according to spectrum characters, spatial relation, and other context relations between pixels. It's generated by analysing and concluding of mass samples' attributes.

A decision-tree consists of one root node, a series of inner nodes and offshoots, and some leaf nodes. Each inner node stands for an attribute will be tested in decision process. Each offshoot stands for a result in test. Different attribute forms different offshoot. And each leaf node stands for one class, which is the classification result.

In the project of geographic national conditions investigation, it's better to select proper classification method according to the images' spatial resolution and spectrum resolution, and improve the automatic classification's precision as far as possible, which could reduce the following workload and improve classification results' efficiency.

### (5) Deal with special circumstances

Firstly, if there is only panchromatic image in some regions, which would influence the information extraction of the types of natural land cover seriously. In this instance, we should collect some middle and lower resolution remote sensing images, such as 30 meters, 10 meters and 5 meters resolution and so on, whose spectral information would be very useful and can be assistant in land cover classification.

Secondly, if there are a lot of clouds in the high resolution optical image in some regions; furthermore, it is very difficult to obtain some good quality images in these regions, which are also influence the information extraction very seriously. In this instance, we should collect some high and middle resolution SAR remote sensing images whose resolution should be higher

than 5 meters, and some middle and lower resolution remote sensing images, which can help classification.

From SAR remote sensing images, we could extract the types of artificial land cover; besides, water exhibits strong mirror reflex characters, based on this point, and we could extract water information from SAR remote sensing images. Then, extract other types of land cover by using multi-source optical remote sensing images.

## 5. CONCLUSIONS

(1) At present, there are many automatic classification methods based on remote sensing images. However, because of the images' complexity, there isn't one method whose applicability is universal and has high efficiency. In practical application, it should find out a most proper classification method to achieve high precision as far as possible according to the type of images used in project.

(2) The classification efficiency and results' precision demanded in the project of geographic national conditions investigation are very high. Therefore, it's inevitable to interpret and edit by artificial process when the automatic classification can't satisfy the requirement.

(3) Traditional Pixel-Based classification method is suitable for middle, low spatial resolution multi-spectrum images. Aim at extracting information from high spatial resolution images, it's better to utilize Object-Oriented method. This method's technology flow mainly includes pre-processing remote sensing images, obtaining objects' original characters, building indexes, inset assistant datum if necessary, selecting samples, extracting land cover information, analyse and evaluate the results' quality.

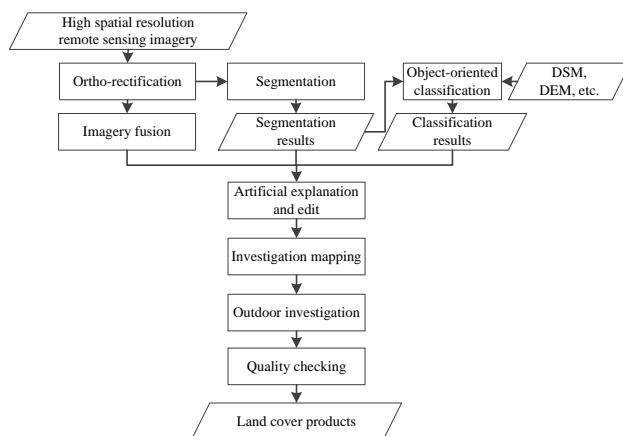


Figure 3. Flow charts of classification based on high spatial resolution remote sensing imagery

Based on this method, the first-level classes land cover information would be extracted by automatic classification, partial second-level classes should be added artificial interpretation and editing, and even outside investigation. This method is suitable for the project of geographic national conditions investigation and can provide references for this project.

## 6. REFERENCES

- XU Han-qiu, 2005. A Study on Information Extraction of Water Body with the Modified Normalized Difference Water Index (MNDWI). *Journal of Remote Sensing*, 9(5), pp. 589-595.
- LUO Xiao-bo, ZHAO Chun-hui, PAN Jian-ping, ZHOU Chun-yan, WANG Yun-an, 2011. *Remote sensing image Intelligent classification and its application*. Publishing House of Electronics Industry, pp. 19-27.
- XIAO Peng-feng, FENG Xue-zhi, 2012. *Segmentation and Information Extraction based on High Resolution Remote Sensing Images*. Science Press, pp. 5-19.
- CAO Li-qin, LI Ping-xiang, ZHANG Liang-pei, CEN Yi, 2010. Classification of Remote Sensing Image Based on Multi-parameters of Land Surface. *Remote Sensing Technology and Application*, 25(1), pp. 38-44.
- CHEN Yun-hao, FENG Tong, SHI Pei-jun, WANG Jin-fei, 2006. Classification of Remote Sensing Image Based on Object Oriented and Class Rules. *Geomatics and Information Science of Wuhan University*, 31(4), pp. 316-320.
- HU Zhang-wu, 2011. Research on Methods of Road Extraction from High-resolution Remote Sensing Image. *Bulletin of Surveying and Mapping*, 8, pp. 13-16.
- MA Xue-mei, CHEN Liang, YU Bing, XU Feng, 2006. An Automatic Classification Model of City Expansion Based on Pixel Information Decomposition Combined with Decision Tree. *Bulletin of Surveying and Mapping*, 10 pp. 9-12.