Object-oriented Land Cover Mapping in China National Geographical Conditions Census

Liang Zhai, Huiyong Sang, Qinghua Qiao

Chinese Academy of Surveying and Mapping, Lianhuachi West Road 28, Haidian District, Beijing, 100830, China-(zhailiang, huiyong.sang, qiaoqh)@casm.ca.cn

ABSTRACT:

Remote sensing image classification is one of the important methods of obtaining land cover/use information for nationwide general survey of geographical conditions in China. In this study, an object-oriented decision tree algorithm is utilized for land cover and land use mapping from high resolution WorldView-2 satellite remote sensing images. The major steps of this approach include image segmentation to get image objects and pixel-based classification based on the image objects. Multi-resolution segmentation approach was adopted in this study, which utilizes spatial and spectral information of land covers in the image to segment into different small objects with particular structural, spectral and texture attributes. A new decision tree classifier called AdaTree. Weight Leaf (AdaTree. WL) was applied to conduct classification process based on the segmented objects, which is modeled by combining the algorithms C4.5 and AdaBoost. The decision algorithm is integrated in the software--GLC (Global Land Cover Classification) classifier. Several study sites are selected from northwest and southeast in China to test GLC classifier, and compared to SVM (Support Vector Machine), the GLC classifier could reaches better results (mean kappa coefficient is 84.61%) in different scenarios with WorldView-2 images.

1. INTRODUCTION

In order to investigate national geographical conditions for the purpose of economic society development and ecological civilization construction, Chinese government initiated the program of the first national geographical conditions census. Land cover mapping is one of the main objectives in this program (NGCC, 2013). Moreover, to better understand the basic conditions of natural ecology and human activity, the timely mapping land-use/land-cover distribution and detecting its change is required (Zhang, J, 2013). Remote-sensing technologies are useful tools in providing these services. Current remote sensing image classification models commonly used pixel-based and object-oriented method according to different studying objects and purposes. Pixel-based classifiers, such as maximum likelihood, decision tree, SVM and neural networks, can be extensive applied for land use classification with moderate or low resolution remote sensing images (Lackner M, 2014, Dengsheng Lu, 2013, Vieira, I. C. G, 2003). With the increase of spatial resolution of satellite image, the traditional per-pixel classification results in the well-known salt and pepper effect, while the object-oriented method has provided a new method for this situation (Lackner M, 2014, Sun Z P, 2010, Chaudhary R C F N, 2008). Meanwhile, decision tree algorithm becomes more mature (Zhai L, 2015, Wang P, 2015, Chasmer L, 2014). GLC tree classifier integrated the improved C4.5 and AdaBoost algorithm (Zhai L, 2015, Dou P, 2013). In the global surface coverage classification of Oceania, GLC tree has been successfully applied to Landsat satellite images (2000 and 2010, 2,796 scenes) category, and the average classification accuracy reached more than 85%. Therefore, the objectives of this research are (1) to develop a new method to improve land use classification accuracy through the integrated use of object-oriented method and GLC tree algorithm with high resolution WorldView-2 satellite remote sensing images and (2) to propose the automatic classification technology for the national geographical conditions census.

2. METHOD

2.1 Data Source

WorldView-2 remote sensing images are the main data source of nationwide general survey of geographic conditions. It provides eight bands (coastal zone, blue, green, red, yellow, red edge band, and two near infrared bands) with 2 m pixel size, and one panchromatic band with 0.46 m pixel size. While in the census，there are four bands (blue, green, red and near infrared) mainly used in extracting the land-cover information. Therefore, these four bands are adopted in this study.

2.2 Method

The major steps of this approach include image segmentation to get image objects and pixel-based classification based on the image objects. The strategy of object-oriented decision tree algorithm for land cover classification is illustrated in Figure 1. The major steps include (1) taking image segmentation on WorldView-2 image to get homogeneous polygons; (2) selecting and extracting features for classification, taking cross-validation on sample quality after collecting specimens; (3) achieving automatic classification by using GLC tree classifier; (4) mapping land cover with manual editing.
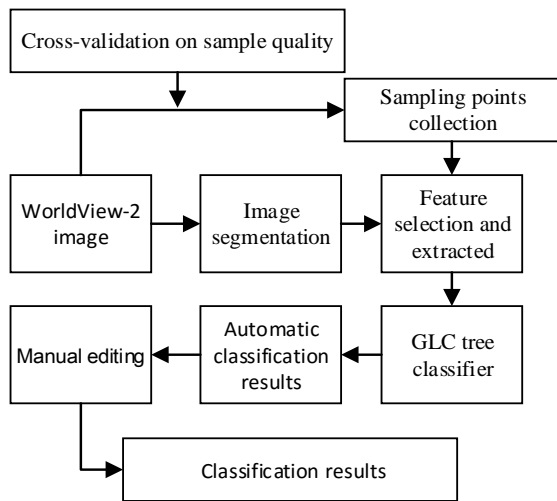
Figure 1. Strategy of object-oriented decision tree algorithm for land cover classification

2.2.1 Image segmentation and feature selection

Multi-resolution segmentation approach is adopted in this study, which utilizes spatial and spectral information of land covers in the image to segment into different small objects with particular structural, spectral and texture attributes. Table 1 shows some of the most common features among the classification based on object-oriented method.

Table 1. Features of remote sensing images

Type	Characteristic
Spectrum	Brightness, Mean, Ratio, Standard Deviation, Normalized Difference Vegetation Index (NDVI), etc
Structure	Aspect Ratio, Perimeter, rectangle fit, Density, Area, Compactness, etc
Texture	Homogeneity (HOM) , Entropy (ENT), Contrast (COM), Angular Second Moment (ANG), etc

2.2.2 Sampling points collection

The classifications of land cover are defined in *Contents and indexes of national geographical conditions census*, and classified into three levels (the first level has 12 second classes, and the second level are divided into 49 third classes, and then reclassified into 86 subclasses). It's suggested to flexibly define the land coverage in this research. User-defined land cover types are also utilized in the sample collection. Croplands and fallow lands with evident spectral difference are further differentiated from the land cover class "agricultural lands". There are several suggestions that may help extracting land cover samples:

(1)The classification accuracy is determined by the quality of training samples. It is better to collect samples by the technicians with rich experience.
(2)Training samples should generally be no less than five in each image.
(3)The classification accuracy of land cover classes is related with the number of training samples.
(4)The samples are selected from the centres of land cover types.
(5)Usually the collected samples have a uniform spatial distribution.

2.2.3 GLC tree classifier

A new decision tree classifier called AdaTree. Weight Leaf (AdaTree. WL) was applied to conduct classification process based on the segmented objects. AdaTree. WL is modeled by combining the algorithms C4.5 and AdaBoost. Unlike C4.5, this algorithm develops a single decision tree by conducting a binary recursive partitioning process and predefining a depth threshold for this tree. AdaTree. WL also modifies the final hypothesis of AdaBoost by setting a prediction weight for each leaf. The prediction weight is given by its prediction accuracy and the weight of the decision tree it belongs to. The decision algorithm is integrated in the software--GLC (Global Land Cover Classification) classifier. There are three steps to achieve automatic classification results (Figure 2). Based on this method, it takes a set of rules in process of training samples, and the single one can be expressed as follows:

count = "2" cover = "6" correct = "6" pre = "1.0518" class = "8";
att = "4" cut = "1170.5" result = "<=";
att = "4" cut = "601" result = ">"

where   count = number of judgment condition
            cover = sample size of the single rule
            correct = correct number of predicted value
            pre = weighted prediction of the single rule
            class = class code
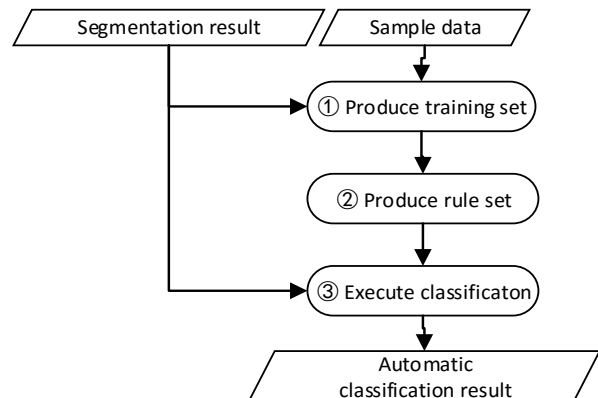   att = attributive judgment
            cut = threshold

Figure 2. Workflow of GLC tree classifier for land cover classification

3. RESULT AND DISCUSSION

In order to test the object-oriented decision tree algorithm, several study sites are selected from northwest northeast and south east in China. Meanwhile, a comparison of the results with SVM-based method shows that whether the GLC classifier is creditable.

(1) Experiment 1 - Lingtong Shanxi

In this experiment, segmentation scale is set to 100, and training samples and test samples are 368 and 201, respectively.
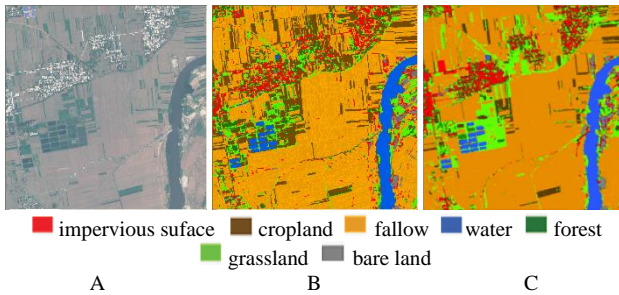
Figure 3. A comparison of classification results from different classifiers. A is the WorldView-2 image somewhere in Lingtong. B, C are classified images based on GLC and SVM classifier, respectively.

(2) Experiment 2 - Northeast China

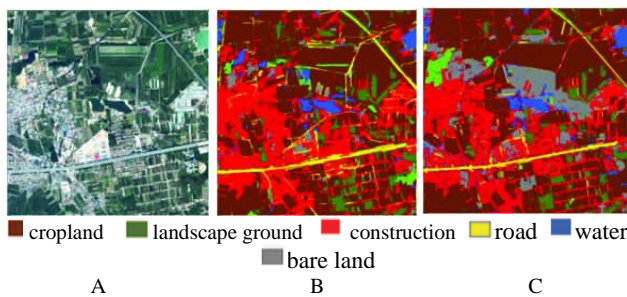For the second experiment, parameter settings are 200, 285 and 236.

Figure 4. A comparison of classification results from different classifiers. A is the WorldView-2 image somewhere in Northeast China. B, C are classified images based on GLC and SVM classifier, respectively.

(3) Experiment 3 - Hangzhou Zhejiang

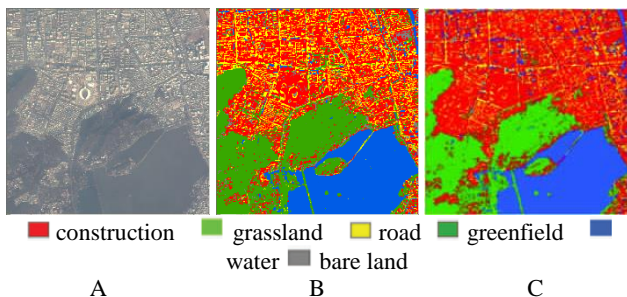The segmentation scale is 150, training samples and test samples are set in 406, 206.

Figure 5. A comparison of classification results from different classifiers. A is the WorldView-2 image somewhere in Hangzhou. B, C are classified images based on GLC and SVM classifier, respectively.

Compared to the distinction of land cover from original images among different study areas, classification difficulty increases gradually. Comparative analysis of the land cover classification results (Figure 3, 4 and 5) shows that GLC tree algorithm is helpful in improving classification accuracies. The land cover classification results (Table 2) based on image segmentation indicate that, compared to SVM method, GLC increased OA (Overall Accuracy) by 1.06%-4.65%. In particular, the precision are willing to decline as the difficulty of classification. Meanwhile, based on AdaBoost lifting scheme, GLC classifier

can improve classification accuracy，but there existed some differences each experiment for that algorithm uses random function to reselect different samples based on a new weight. While there is no remarkable difference between each classification，so this approach remains suitable for national geographical conditions census.

Table 2. Comparison of classification results from different classifiers based on WorldView-2 data

Evaluation Index	GLC-Tree		SVM	
	Kappa Coefficient	Overall Accuracy	Kappa Coefficient	Overall Accuracy
Experiment 1	93.53%	92.42%	89.55%	87.77%
Experiment 2	87.29%	85.19%	84.65%	82.70%
Experiment 3	81.30%	79.23%	80.19%	78.17%
Mean	87.37%	85.61%	84.80%	82.88%

4. CONCLUSIONS

The method developed in this paper, which is based on the object-oriented decision tree algorithm that combine image segmentation and GLC-based algorithm, can effectively extract land cover with WorldView-2 images. Meanwhile, three images from different regions in China were choose and processed with GLC classifier, and the mean kappa coefficient of result reaches 84.61%. In particular, the automatic classification technology is recommended in geographical conditions census for detailed land cover classification. In view of improving work efficiency in the project, we will concentrate on setting up a rule database for land cover classification based on GLC tree algorithm in the future.

REFERENCES

NGCC, 2013. The notification of first national geographical conditions census.
http://ngcc.sbsm.gov.cn/article/kjcj/zdxmcg/201307/201307000 02626.shtml.

Zhang, J., 2013. Understanding geographical conditions monitoring: a perspective from China. *International Journal of Digital Earth*, 8:38-57.

Zhang J., 2014. A comparison of pixel-based and object-based land cover classification methods in an arid/semi-arid environment of Northwestern China[C]// Earth Observation and Remote Sensing Applications (EORSA), 2014 3rd International Workshop on. IEEE, 2014:403-407.

Lackner M., 2014. Determining land-use information from land cover through an object-oriented classification of IKONOS imagery. Canadian Journal of Remote Sensing, 34(2):77-92.

Dengsheng Lu., 2013. A comparison of multisensor integration methods for land cover classification in the brazilian amazon. *Giscience & Remote Sensing*, 48(3):345-370.

Vieira, I. C. G., 2003. Classifying successional forests using landsat spectral properties and ecological characteristics in eastern amazônia. *Remote Sensing of Environment*, 87(4):470-481.

Sun Z P., 2010. Object-oriented land cover classification using HJ-1 remote sensing imagery. *Science China*, 53(S1):34-44.

Chaudhary R C F N., 2008. Multi-scale Image Segmentation and Object-Oriented Processing for Land Cover Classification. *Giscience & Remote Sensing*, 45(4):377-391.

Zhai L., 2015. A new approach for mapping regional land cover and the application of this approach in Australia. *Remote Sensing Letters*, 6(4):267-275.

Dou P., 2013. Research and application of object-oriented remote sensing image classification based on GLC. *Geomatics & Spatial Information Technology*, 36(11):68-71.

Wang P., 2015. A Land Cover Classification Method for Antarctica Using Support Vector Machine and Decision Tree. *Open Cybernetics & Systemics Journal*, 9(1):2920-2928.

Chasmer L., 2014. A decision-tree classification for low-lying complex land cover types within the zone of discontinuous permafrost. *Remote Sensing of Environment*, 143(10):73-84.