

FUSION OF HIGH-RESOLUTION AERIAL IMAGERY AND LIDAR DATA FOR OBJECT-ORIENTED URBAN LAND-COVER CLASSIFICATION BASED ON SVM

LI Haitao ^{a,*}, GU Haiyan ^{a,b}, HAN Yanshun ^a, YANG Jinghui ^a

^aInstitute of Photogrammetry and Remote Sensing, Chinese Academy of Surveying and Mapping, Beijing, 100039, China.- lhtao@casm.ac.cn

^bInstitute of Surveying and Geography Science, Liaoning Technical University, Liao Ning, 123000, China.

KEY WORDS: Aerial imagery, Light Detection and Ranging (LIDAR), Image Fusion, Object-oriented Classification, Land-cover Classification, Support Vector Machine (SVM)

ABSTRACT:

We present a new object-oriented land cover classification method based on Support Vector Machine (SVM) by fusing spectral and textural information of High-Resolution (HR) aerial imagery and Digital Surface Model (DSM) information of Light Detection and Ranging (LIDAR) data in the same urban area. It synthesizes the advantages of digital image processing (image fusion, image segmentation), Geographical Information System (GIS) (vector-based feature selection) and Data Mining (intelligent SVM classification) to interpret image from pixels to segments and then to thematic information. Compared with the pixel-based SVM classification where the multi-source imageries were used, the results showed that the proposed object-oriented method by fusing of multi-source imageries could share redundant and complementary information, and also provided greater classification detail and accuracy, especially, various kinds of shadows were correctly identified, and the computational performance for classification were improved. Moreover, the land cover classification map could update GIS database in a quick and convenient way.

1. INTRODUCTION

For years, many efforts have been made to develop automatic procedures for integrating raster analysis and vector analysis. Numerous researches have been focusing on updating GIS databases using remote sensing data and land cover classification technique. The High Resolution (HR) imagery has the advantage of high user interpretability, rich information content, sharpness, high image clarity, and integrity, which provides the unique tool for classification. However, it is not practical to classify the HR imagery using traditional pixel-based classification methods (e.g., Maximum Likelihood, ISODATA, K-means). Recently, researchers have explored new classification techniques, such as multi-classifier systems that integrate the outputs of several underlying classifier algorithms, and the use of structural, spatial, context information from imagery. However, these pixel-based classification techniques improve classification accuracy with some limitations: 1) they have considerable difficulties dealing with the abundant information of HR data, 2) they produce a characteristic and inconsistent salt-and-pepper classification, 3) they are not capable of extracting objects of interest, and 4) they can not update GIS database expediently (Baatz, M., et al., 2004). Thus, the pixel-based classification methods are no longer applicable for HR imagery.

In this situation, commercial software, eCognition, brought forward the revolutionary of object-oriented approach, which adopts a bottom-up region-merging segmentation technique and fuzzy logical classification method. Recently, some researchers have been succeeded in classification with new object-oriented classification methods adopting other segmentation and classification techniques. Furthermore, in object-oriented image analysis, multi-source data fusion technique is helpful to land cover classification. Cuzzo, G. (2004) succeeded in classification with new object-oriented classification methods adopting Markov Random Field (MRF) segmentation

techniques and MLC algorithm. Bruzzone, L.(2004) proposed a multi-level hierarchical approach to classify high spatial resolution images with Support Vector Machines, and the results and comparisons confirmed the effectiveness of the approach. Cui, L.L. (2006) succeeded in classification of typical artificial objects based on various features and object-oriented classification. Park, J.Y. (2001) fused ALSM (LIDAR) and colour digital photography for land-cover classification, and the results showed that the fusion of multi-source imageries could improve accuracy and create more consistent recognition of land cover patterns. In this paper, an object-oriented classification scheme based on SVM is developed where LIDAR data and HR aerial imagery are used.

2. OBJECT-ORIENTED IMAGE CLASSIFICATION BASED ON SVM

2.1 SVM Classification Method

SVM is a classification technique developed by Vapnik and his group at AT&T Bell Laboratories. It is based on Vapnik-Chervonenkis (VC) dimension theory and Structural Risk Minimization (SRM) rule. SVM can solve sparse sampling,

non-linear, high-dimensional data, and global optimum problems. The performance of SVM has been proved as good as or significantly better than that of other competing methods in most cases (Burges, C. J. C., 1998. Zhang, X. G., 2000).The SVM method separates the classes with a hyperplane surface to

maximize the margin among them (see figure 1), and m is the distance between H_1 and H_2 , and H is the optimum separation plane which is defined as:

$$w \cdot x + b = 0, \quad (1)$$

where x is a point on the hyperplane, w is a n-dimensional vector perpendicular to the hyperplane, and b is the distance of the closest point on the hyperplane to the origin. It can be found as:

$$w \cdot x_i + b \leq -1, \text{ for } y_i = -1 \quad (2)$$

$$w \cdot x_i + b \geq 1, \text{ for } y_i = +1 \quad (3)$$

These two inequalities can be combined into

:

$$y_i [(w \cdot x_i) + b] - 1 \geq 0 \quad \forall i. \quad (4)$$

The SVM attempts to find a hyperplane (1) with minimum $\|w\|^2$ that is subject to constraint (4).

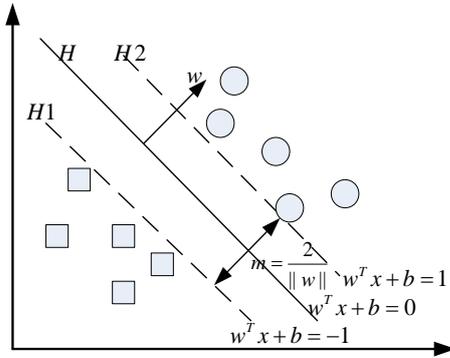


Fig. 1. Optimum separation plane.

The processing of finding optimum hyperplane is equivalent to solve quadratic programming problems:

$$\begin{aligned} \min & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \\ \text{s.t.} & y_i [w \cdot \phi(x_i) + b] \geq 1 - \xi_i \\ & \xi_i \geq 0, i = 1, 2, \dots, l \end{aligned} \quad (5)$$

where C is penalty parameter, which is used to control the edge balance of the error ξ . Again, using the technique of Lagrange Multipliers, the optimization problem becomes:

$$\begin{aligned} \min & \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j K(x_i, x_j) - \sum_{i=1}^l \alpha_i \\ \text{s.t.} & \sum_{i=1}^l y_i \alpha_i = 0 \\ & 0 \leq \alpha_i \leq C, i = 1, 2, \dots, l \end{aligned} \quad (6)$$

where $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$ is kernel function. There are three major kernel functions including Gaussian Radius Basis

Function (RBF), Polynomial, and Sigmoid function. The optimum classification plane is solved through chunking, Osuna, and SMO algorithms, and then we will only need to compute $K(x_i, y_j)$ in the training process. The final decision function is:

$$f(x) = \text{sgn} \left(\sum_{sv} y_i \alpha_i K(x, x_i) + b \right) \quad (7)$$

When multi-class SVM is concerned, three basic methods are available to solve the classification: One-Against-All (OAA), One-Against-One (OAO), and Directed-Acyclic-Graph (DAG). Recent remote sensing literature has shown that SVM methods generally outperform traditional statistical and neural methods in classification.

2.2 The Processing of Object-oriented Classification Based on SVM

In this section, a new procedure of object-oriented image classification based on SVM is presented. It synthesizes image preprocessing (e.g., information fusion, image segmentation), geographical information system (vector-based feature selection) and data mining (intelligent SVM classification) technology to interpret image from pixels to segments and then to thematic information, and it is an integrative, iterative classification system. The flowchart of the object-oriented classification algorithm is shown in figure 2, and there are three main processes: the multi-source imagery information fusion progress, the improved Color Structure Code (CSC) segmentation progress, and the object-oriented SVM classification progress.

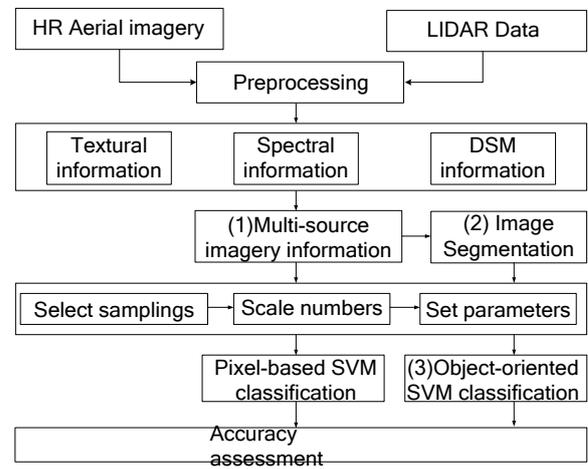


Fig.2. The flowchart of object-oriented classification based on SVM.

(1) In the multi-source imagery information fusion progress, aerial imagery and LIDAR derived DSM are fused for urban land cover classification. The potential advantage for fusing multi-source data is the possibility to obtain more accurate information concerning features that cannot be achieved with individual single-source data. LIDAR collection system includes a powerful laser sensor comprising a transmitter and receiver, a geodetic-quality Global Positioning System (GPS)

receiver, and Inertial Navigation System (INS) unit. Therefore, LIDAR system records elevation data directly, and LIDAR data provide very accurate position and height information, but less direct information on the object's geometrical shape, while HR aerial imagery offers very detailed information on objects, such as spectral signature, texture, shape, etc, but is relatively insensitive to variation in their height. Therefore, the fusion of LIDAR data and aerial imagery can greatly improve the accurate of land-cover classification (Syed, S., 2005).

(2) In the segmentation progress, the fusion imagery is segmented by use of the improved CSC segmentation method, which is a hierarchical region growing method, and inherently parallel, and therefore independent of the choice of the starting point and the order of processing. Moreover, it combines spectral and boundary information, and uses local and global information, and produces robust and accurate segmentation results (Li, H.T., 2007). Through the segmentation process, the image is subdivided into a huge number of homogeneous elementary objects.

(3) In the object-oriented SVM classification progress, firstly, the features (e.g., spectral, textural) of sample objects from different classes are calculated. Typical features of some objects act as the training samples of SVM, and the whole objects act as the testing samples. Secondly, the training samples and the testing samples are scaled. The main purpose of this step is to avoid the domination of greater numeric ranges on smaller numeric ranges. Another purpose is to avoid numerical difficulties during the calculation. Thirdly, the parameters are set. In this step, the training samples are mapped to the higher dimensional feature space to find the optimum space of various classes using SVM, thus the support vectors and VC are gained reliably, and the decision function and training model are formed. The testing samples acting as the input of decision function are tested to get the classification map. At last, the misclassified objects detected by the proposed classification method are classified again. This step is iterative until getting the best classification result. For the implementation of SVM, the software package LIBSVM by Chang and Lin (2001) was adopted.

3. CLASSIFICATION EXPERIMENT

3.1 Study Area

The experimental data are LIDAR data and HR aerial imagery at 0.4-m spatial resolution, which were acquired in Odense of Denmark. The LIDAR derived DSM image and HR aerial image were geo-referenced to the same coordinate and projection with 376 by 339 pixels size. The digital surface model was created from the first and last returns LIDAR data by applying height thresholds (figure 3). Figure 4 shows the HR aerial imagery of bands 3, 2, and 1. Seven land cover classes were identified: 1 road, 2 building, 3 shadow of building, 4 tree, 5 shadow of tree, 6 grass, 7 bare land.

3.2 Classification Process

In the data pre-processing stage, the textural features (variance, contrast) of aerial imagery were calculated using Gray Level Co-occurrence Matrix (GLCM) filtering, and the texture and spectral information of HR aerial imagery were fused with the LIDAR derived DSM imagery to get multi-source image information.

The segmentation was achieved by the improved CSC segmentation method, which combines spectral and boundary information, and got accurate segmentation results. Figure 5 shows the segmented fusion imagery.

After the segmentation, we selected several objects from each class as the training samples, and defined the whole objects as the testing samples.

During the classification processing, first, we scaled the training samples and the testing samples before applying SVM. Secondly, we chose RBF kernel, which is more suitable to land cover classification. Thirdly, we used cross-validation and grid-search method to get the best parameters of RBF. The highest overall accuracy on the multi-source image information was obtained with $g=0.125$ and $C=32768$, where g is the width of the kernel function, and C is the regularization parameter. At last, the training samples were used to train the support vector machine, and the resulted model was used to classify the entire image with the OAO method to get the preliminary classification result which is perhaps not the best one. So we selected samples and classified sequentially for various classes by analyzing the classification errors of omission and commission until getting the best result. The final classification map is shown in figure 6. For the purpose of getting the preferable classification result, it is better to consider small misclassified objects.

To evaluate the classification performance of the proposed object-oriented method, we also classified the fusion imagery using pixel-based SVM method under ImageInfo (www.image-info.com) software environment where the multi-source image information was used. The overall accuracy was obtained with $g=0.25$ and $C=100$, and the result is shown in figure 7.



Fig.3. The LIDAR derived DSM imagery.

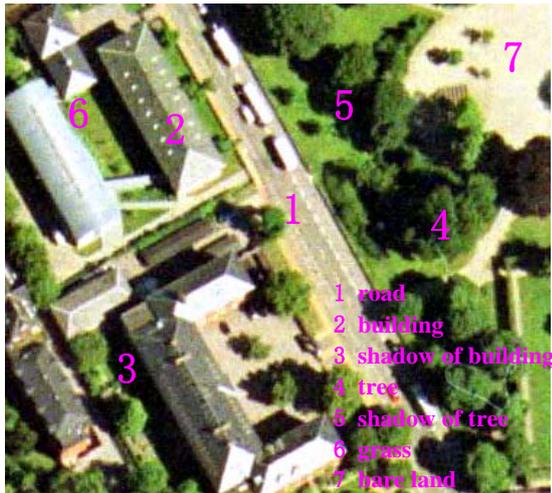


Fig.4. The high resolution aerial imagery.



Fig.5. The improved CSC segmentation imagery.

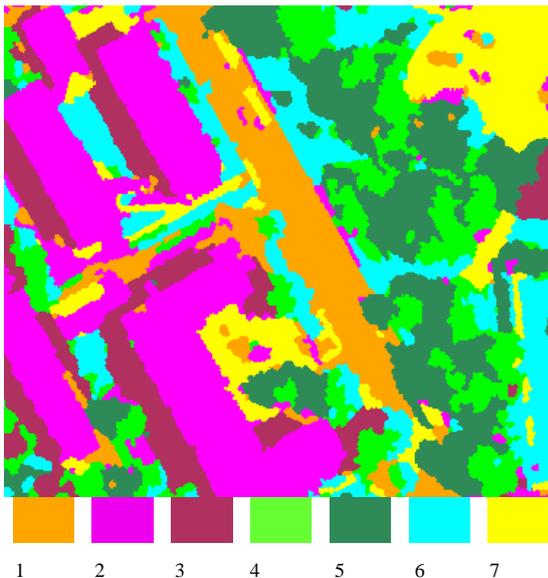


Fig.6. The object-oriented classification map based on SVM.

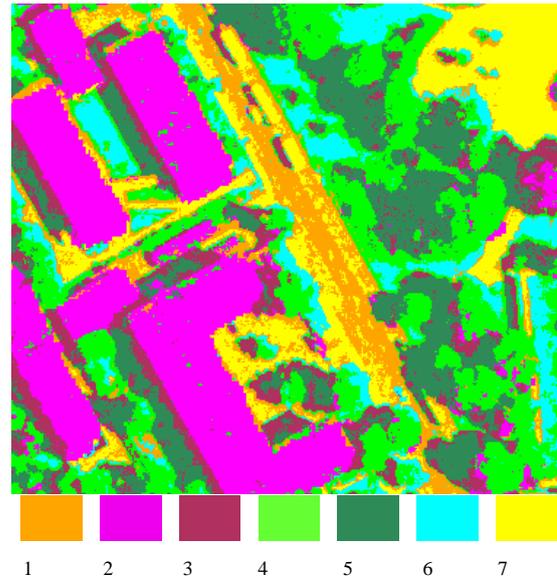


Fig.7. The pixel-based classification map based on SVM.

3.3 Accuracy Assessment

Comparing Fig.6 and Fig.7, we notice that different shadow species are easier to distinguish when both aerial images and LIDAR data are used together. Fig.7 appears “salt and pepper” phenomenon, and shadows of buildings are confused with shadows of trees. Table 1 shows the confusion matrix of object-oriented classification based on SVM, and table 2 shows the confusion matrix of pixel-based classification based on SVM.

The results show that the overall accuracy of pixel-based SVM is 82.92%, whereas the overall accuracy of the object-oriented classification based on SVM is 95.97%. The overall accuracy is improved by 13.05%. In addition, the producer’s accuracy of each class using object-oriented classification is higher than the pixel-based SVM method except building class. The user’s accuracy of each class using object-oriented classification is higher than the pixel-based SVM method. In particular, the land cover class, building and road, are obviously distinguished by the object-oriented classification method by fusing of DSM information of LIDAR data and spectral, textural information of aerial imagery.

4. CONCLUSION AND DISCUSSION

The study proposed a new object-oriented land cover classification method based on SVM using HR aerial imagery and LIDAR data, which built connection between the domains of raster analysis and vector analysis. The SVM classification method could solve sparse sampling, non-linear, high-dimensional data, and global optimum problems. The fusion of multi-source data could obtain more accurate information concerning features that can not achieve with individual single-source data. It was expected that the addition of LIDAR derived DSM information would result in a significantly higher accuracy than using just aerial imagery, especially, object-oriented SVM classification has proven effective in correctly identifying various kinds of shadows.

Compared with the pixel-based SVM classification method, the results indicated that the proposed object-oriented classification method can improve the efficiency of training and classification, get accurate classification results, and update GIS database in a

quick and convenient way. Because the proposed classification method could detect and reclassify the small misclassified objects in an iteration way, we may get better classification result. Additionally, this method has the potential to apply in land cover monitoring at regional and global scale, object detection, feature extraction, etc. From this object-oriented approach, it was envisaged that incorporating LIDAR data into the HR aerial imagery could make the automatic feature extraction more effective.

For the purpose of accurately classifying different ground objects, it is important to consider scale effects and additional effective features in the near future.

REFERENCES

- Baatz, M., et al., 2004. eCognition User Guide 4, www.definiens-imaging.com (Last visited March 10, 2007).
- Bruzzone, L., Carlin, L. and Melgani, F., 2004. A multilevel hierarchical approach to classification of high spatial resolution images with support vector machines. IGARSS '04. pp. 20-24.
- Burges, C. J. C., 1998. A tutorial on support vector machines for pattern recognition. Knowledge Discovery and Data Mining, 2(2), pp. 121-167.
- Chang, C. C. and Lin, C. J., 2001. LIBSVM: a library for support vector machines. <http://www.csie.ntu.edu.tw/~cjlin/>, (Last visited March 28, 2007).
- Cuozzo, G., Elia, C. D. and Puzzolo, V., 2004. A method based on Tree-Structured Markov Random Field for forest area classification. IGARSS'04. pp. 2352-2354.
- Cui, L.L., Tang, P. and Zhao Z.M., 2006. Study on object-oriented classification method by integrating various features. Journal of remote sensing, 10(1), pp. 104-110.
- Li, H.T., Gu, H.Y., Han, Y.S. and Yang, J.H., 2007. Object-oriented classification of very high-resolution remote sensing imagery based on improved CSC and SVM. Geoinformatics'2007.
- Park, J.Y., Shrestha, R.L., Carter, W.E. and Tuell, G.H., 2001. Land-cover classification using combined ALSM (LIDAR) and color digital photography. Presented at ASPRS conference, St. Louis, Missouri, April, pp. 23-27.
- Syed, S., Dare, P. and Jones, S., 2005. Automatic classification of land cover features with high resolution imagery and lidar data: an object-oriented approach. Proceedings of SSC2005 Spatial Intelligence, Innovation and Praxis: The national biennial Conference of the Spatial Sciences Institute, September, 2005. http://www.definiens.com/resources/es_scientific_paper2005.php, (Last visited March 10, 2007).
- Zhang, X. G., 2000. Introduction to statistical learning theory and support vector machines. Journal of Automatization, 26(1), pp. 32-42.

ACKNOWLEDGEMENTS

This work was supported by Key Laboratory of Geoinformatics of State Bureau of Surveying and Mapping (Grant No.200601). We acknowledge Definiens, the image intelligence company, who provided us the LIDAR data and HR aerial imagery.

Class	Road	Building	Shadow of building	Tree	Shadow of tree	Grass	Bare land	Row Total	User's Accuracy (%)
Road	70	0	0	0	0	0	0	70	100
Building	0	88	0	0	0	0	0	88	100
Shadow of building	0	0	67	0	0	0	0	67	100
Tree	0	0	0	68	0	0	0	68	100
Shadow of tree	0	0	0	0	65	0	0	65	100
Grass	0	0	0	6	0	76	0	82	92.68
Bare land	9	6	0	0	0	0	66	81	81.48
Column Total	79	94	67	74	65	76	66	521	
Producer's Accuracy (%)	88.61	93.62	100	91.89	100	100	100		

Note: Total samples = 521 pixels, correctly classified samples =500 pixels, overall accuracy 95.97%, kappa coefficient $\hat{k} = 0.95$.

Table 1. The confusion matrix of object-oriented classification based on SVM

Class	Road	Building	Shadow of building	Tree	Shadow of tree	Grass	Bare land	Row Total	User's Accuracy (%)
Road	61	0	0	0	0	0	17	78	78.21
Building	0	93	0	0	1	0	0	94	98.94
Shadow of building	0	0	40	0	4	0	0	44	90.91
Tree	0	0	0	62	0	9	0	71	87.32
Shadow of tree	0	1	27	4	60	0	0	92	65.22
Grass	0	0	0	8	0	67	0	75	89.33
Bare land	18	0	0	0	0	0	49	67	73.13
Column Total	79	94	67	74	65	76	66	521	
Producer's Accuracy (%)	77.22	98.94	59.70	83.78	92.31	88.16	74.24		

Note: Total samples = 521 pixels, correctly classified samples =432 pixels, overall accuracy 82.92%, kappa coefficient $\hat{k} = 0.80$.

Table 2. The confusion matrix of classification based on pixel-based SVM