

Extracting Urban Ground Object Information from Images and LiDAR Data

Lina Yi^{a*}, Xuesheng Zhao^a, Luan Li^a, Guifeng Zhang^b

^aChina University of Mining & Technology (Beijing), College of Geoscience and Surveying Engineering, Ding No.11
Xueyuan Road, Haidian District, Beijing 100083, China

^bAcademy of opto-electronics, Chinese academy of sciences

KEY WORDS: Urban Ground Objects; Multi-scale Image Segmentation; High Resolution Image; Object-oriented Classification; LiDAR

ABSTRACT:

To deal with the problem of urban ground object information extraction, the paper proposes an object-oriented classification method using aerial image and LiDAR data. Firstly, we select the optimal segmentation scales of different ground objects and synthesize them to get accurate object boundaries. Then, this paper uses ReliefF algorithm to select the optimal feature combination and eliminate the Hughes phenomenon. Eventually, the multiple classifier combination method is applied to get the outcome of the classification. In order to validate the feasible of this method, this paper selects two experimental regions in Stuttgart and Germany (Region A and B, covers 0.21km² and 1.1km² respectively). The aim of the first experiment on the Region A is to get the optimal segmentation scales and classification features. The overall accuracy of the classification reaches to 93.3%. The purpose of the experiment on region B is to validate the application-ability of this method for a large area, which is turned out to be reaches 88.4% overall accuracy. In the end of this paper, the conclusion shows that the proposed method can be performed accurately and efficiently in terms of urban ground information extraction and be of high application value.

1. INTRODUCTION

As the development of Remote sensing technology, the approach to extract information from high resolution images has become the key to monitor the development of urban area, including the changes of land usage. However, for a complex urban area, the spectrum, texture and other characteristics of buildings, vegetation and other surface features on high resolution image is liable to be affected by the shadow, which will affect the information extraction accuracy. To improve the accuracy, more researchers are attempting to combine multi-source images for urban areas information extraction. For example, Haala (1997) conducted urban land use classification using aerial images, DSM (digital surface model) and the existing 2D Geographic Information Systems (GIS). Grigillo (2012) integrated the LiDAR (Light Detection and Ranging) and aerial image to extract building and vegetation. Charaniya (2004) classified aerial LiDAR height data into roads, grass, buildings, and trees using a supervised parametric classification algorithm. However, all these researches use the pixel-based classification methods, which will result in the serious Pepper-Salt effect.

Object-based classification technique separates similar pixels at different scale levels as segmented objects, this technique can use spatial information of objects which is helpful to extracting urban ground information, which was outlined by Moran E F (2010). At present, the relevant research mainly focused on the multi-scale image segmentation and classification methods. However, the single scale of segmentation result is liable to be over- or under-segmented, For a complex urban ground field, to

get a high quality segmented result, it is essential to use multi-scale segmentation considering the scale characteristics of various objects. For example, a multi-scale segmentation method based on QuickBird imagery and nDSM (normalized digital surface model) data was proposed by Chen etc (2009), in which different land cover classes were extracted using different segmentation parameters (such as scale, color and shape parameters) based on different images. However, the optimal segmentation scales are mainly identified by trial and error in these researches, such as Syed et al (2005). It not only results in low degree of automation but also is more or less subjective. On the other hand, the urban ground object classification methods are mainly based on the complicated hierarchical classification rules. For example, Alireza (2014) established the classification rule sets to divide urban ground objects into 13 categories. It is of low reuse ability because the classification rules depend on the study areas. The rule sets especially rely on expertise and the researcher's experience.

In order to solve these problems that the urban ground object boundary is difficult to extract accurately and the classification method is hard to be reused for other application, this paper proposes an object-oriented method integrating high spatial resolution image and LiDAR data. Firstly we select the optimal segmentation scale by the supervised segmentation evaluation method, and the segmentation accuracy assessed the quality of the delineated urban ground objects. Then a scale synthesis segmentation method is used to get the segmentation result which considers both the local and global optimal segmentation results. After that, we select the optimal combination of features through ReliefF Feature

Selection Algorithm (Zhang G, 2012). Then the multiple classifiers combination method is proposed to classify the urban ground objects. To validate our method, three classification experiments were conducted. The first experiment on region A is to select the optimal segmentation scales and classification features. The overall accuracy of the classification reaches to 93.3%, and experiment on region B is designed to validate the application-ability of this method for large area, the overall accuracy reaches to 88.42%, the last experiment is based on hierarchical object oriented classification, the overall accuracy reaches to 79.75%.

In the following text, the study area and data information is introduced in Section 2. The methodology of the data pre-processing, multi-scale segmentation, classification and verification in large urban area are described in Section 3. The results of the experiment are discussed and the conclusions are shown in Section 4.

2. STUDY AREA AND DATA DESCRIPTION

The data set was captured over Vaihingen in Germany, it is the test data provided by the ISPRS-Commission Group Working III /4, including LiDAR data (the mean point density is 6.7 points/m), DSM, digital aerial images (resolution 0.08m, NIR/R/G) and true orthoimages. The DSM was generated from the last return measurements with 0.25m ground resolution. LiDAR point cloud data includes x, y, z coordinate information, multiple echo information, intensity information etc. Two study regions as shown in Fig. 1(A and B) are selected for classification experiments. There are seven land cover classes both in region A and B including: building, road, open spaces (including parking lot, bare soil), tree, grassland, water and shadow. Region A is a smaller study area for training the segmentation, selecting feature and classification method and getting the appropriate parameters which can be used in each step. Experiment on region B was implemented to validate the applicability of this method. In order to improve classification accuracy on shadow, we also defined shadow into pure shadow, shaded building, shaded road, shaded open space and shaded vegetation.



Figure1. The study area

3. METHODOLOGY

There are three steps in the urban ground object information extraction process: image segmentation (Zhang G, 2010), classification and post-classification. The workflow of our method is presented in Fig. 2. At first, we fuse aerial image and LiDAR data by image registration and band overlay method. Then, the multi-scale segmentation is implemented, the supervised segmentation evaluation method is used to select the optimal segmentation scales for different objects, and a scale synthesis method is used to extract the final segmentation result, in which the spectral, texture, spatial information and LiDAR elevation information, intensity and other characteristics of ground objects are extracted. After that, the ReliefF feature selection algorithm is used to select the optimal feature combination. The study area is finally classified through the multiple classifier combination method, and post-classified by rules with the help of eCognition software. The steps are detailed in the following sub-sections.

3.1 Data pre-processing

The original point cloud data are provided by ISPRS, and geographical coordinates have been corrected by the adjustment of the plane. Terrascan software is used to remove noise, then the Digital Terrain Model (DTM) is generated with ENVI LIDAR data processing functions, and the projected coordinate system is set as UTM. The datum is identified as WGS84. Subsequently, the nDSM was generated by subtracting the DTM from the DSM through band math function of ENVI5.1, it provided the accurate height information of all above ground objects (such as buildings, vegetation, etc.). Then, we make geometric registration of imagery by corresponding image points from the point cloud data and image. Finally, the orthophoto and nDSM were synthesized using the layer stacking function of ENVI5.1.

3.2 Multi-scale segmentation and scale synthesis

For urban ground objects with different shape and size, the single segmentation scale cannot reflect the shape information feature, and the classification accuracy decreased when over- and under-segmentation occurred. In this paper, we used a multi-scale edge embedded watershed algorithm for remote sensing image segmentation (Zhang G, 2010), which started with a scale parameter of 3 and ended in 87(the scale interval is 3, a total of 28 segmentation scales). Usually, the scale over 90 will not be considered, which have serious under-segmentation phenomenon. Aiming at quantitatively analyze the optimal segmentation scale of features, we used supervised segmentation accuracy evaluation method to select the optimal segmentation scale of each type of ground objects, and synthesize the optimal scales to get the final segmentation result (Zhang

G, 2010). Fig. 3 display the initial and the synthesized segmentation result of region A.

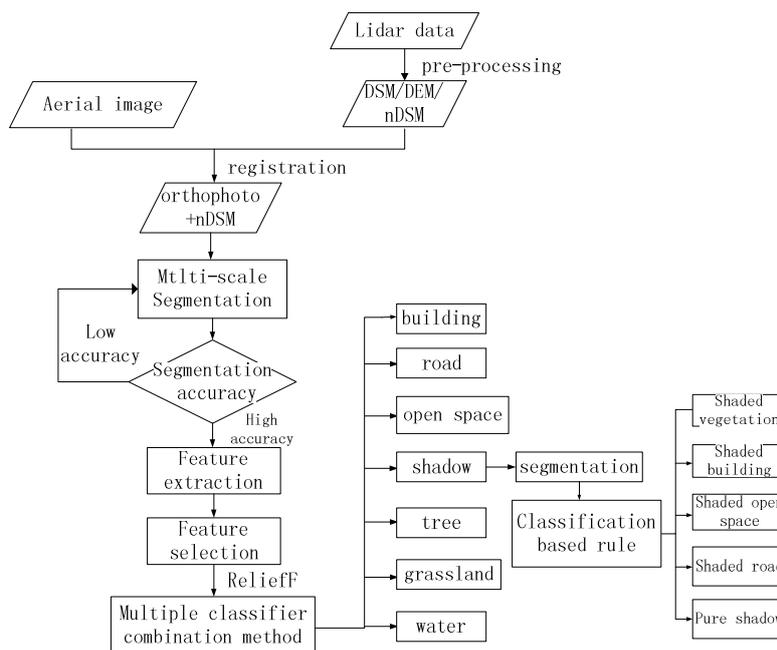


Figure2. Workflow of urban ground object extraction

Precision and recall are supervision evaluation methods which can effectively reveal over- and under-segmentation. The precision measure indicates the ratio of the intersection area and the segmented object (equation 1). While the recall measure implies the ratio of the intersection area and the reference object. The m2 measure shows the ratio of the intersection area and the union area.

$$p = \frac{|S_i^m \cap S_{I(i)}^h|}{|S_i^m|} \setminus * \text{MERGEFORMAT} ($$

$$r = \frac{|S_i^m \cap S_{I(i)}^h|}{|S_{I(i)}^h|} \setminus * \text{MERGEFORMAT} ($$

$$m2 = \frac{|S_i^m \cap S_{I(i)}^h|}{|S_i^m \cup S_{I(i)}^h|} \setminus * \text{MERGEFORMAT} ($$

where S_i^m = Pixels conforming a region in the segmented image.

$S_{I(i)}^h$ = Pixels conforming a region in the ground truth image.

To evaluate the accuracy of each scale, we used a stratified random sampling strategy to choose the testing data, the Fig. 4 presents the curve of the segmentation

quality. The construction land includes woodland, building, road and open space; the woodland includes trees and grassland. The shadow includes shadows cast by buildings and trees in high-density urban environments. This figure shows that when segmentation scale increases, the p value reduces and the r value increases, and a large recall value and a small precision value indicate under-segmentation. On the contrary, over-segmentation has a small recall value and a large precision value. The value m2 is the weight sum of the two metric and without deviation. As a result, the optimal segmentation scale can be chosen in the place where the m2 reaches the largest value. If the m2 values are close, the segmentation can sketch the object boundary effectively when the p value get the largest. As shown in Fig.4, the selected optimal scale of tree and grassland is 33, the road and open space is 45 and the shadow is 18. With consideration that the obvious water characteristic in experimental area and there is no echoed signal in water area in LiDAR, we selected image segmentation primitives as the optimal segmentation scale of water.

After obtaining the optimal segmentation scales of the different ground objects, then scale synthesis is conducted to get the final segmentation result as shown in Fig.4. this method takes the optimal segmentation scale into consideration, consequently, the best segmentation result could be achieved to present the real distribution of the ground object.

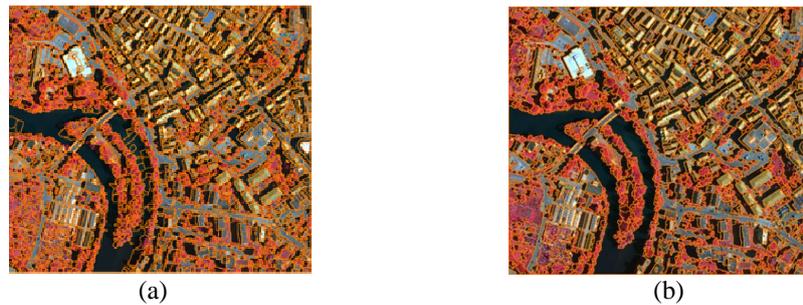


Figure3. (a) The initial segmentation result of A shadow (b)The synthesized segmentation result of A construction land

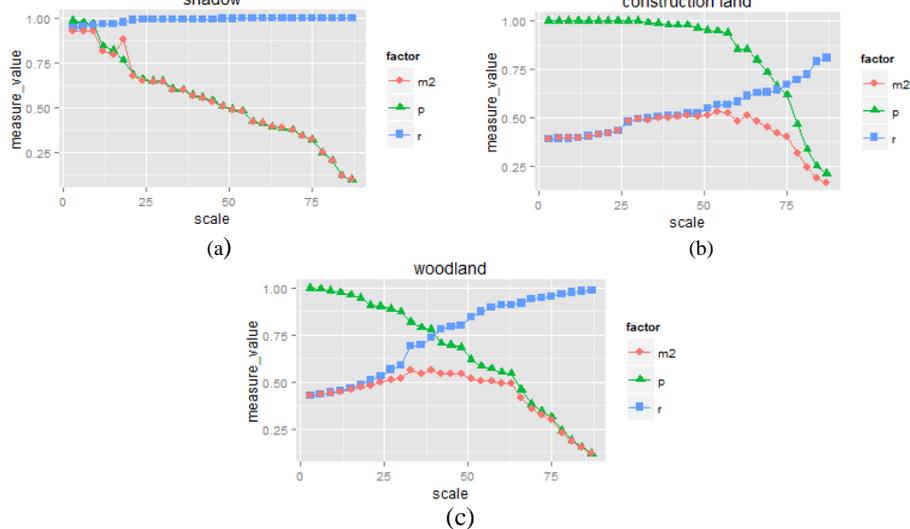


Figure4. Segmentation accuracy evaluation chart of shadow(a), construction land(b) and woodland(c)

3.3 Classification

Before classification, we selected the typical training samples of each type of ground objects, we compiled the statistics of more than 40 features on their spectral, index characters, texture information, shape features, elevation information, and intensity information and so on. In order to remove redundant features without reducing the recognition performance, we measured the importance of features based on ReliefF index (Zhang G, 2010), and classified the data based on the top 12 features sequentially as well as the area and compactness features (distinguishing between roads and open spaces). The selected classification features include: NIR band mean value, nDSM mean value, NDVI(normalized differential vegetation index), RVI(ratio vegetation index), TSAVI(transformed soil adjusted vegetation index), NDWI(normalized difference water index), SVI(shadow vegetation index), custom 3, entropy(nDSM), homogeneity(NIR), Angular Second Moment (nDSM), area, compactness. Chen Y (2007) provided Chen(3) which is the shadow related index, and $Chen(3) = \frac{NIR+G-2*R}{NIR+G+2*R}$.

After features selected, we classified urban targets by using multiple classifiers combination method given by Chen B (2005), which based on the voting weights to improve the classification accuracy, the classifier used in this experiment including: k-Nearest Neighbour (Beckonert O, 2003), Neural Network (Shao Y, 2012) and

SVM_RBF (V. Kecman, 2001). Fig. 6 is the classification result of study area A. considering that there are high-rise buildings and trees in urban areas, the image classification was extremely affected by the shadow, we dealt with this problem by firstly segmenting the shadow areas with the software of eCognition Professional 5.0, and then dividing the shadow areas into five parts on the basis of rules: shaded vegetation, shaded building, shaded road, shaded open spaces and pure shadow. We merged building and shaded building, grassland and shaded vegetation, open spaces and shaded open space, road and shaded road, refined rules of shadow areas as Fig. 5. After supervised classification, there are some confused objects, such as building and road, tree and grassland, so we figured out the suitable rule set and further refined the classification result. Classification rules are as follows:(1)The region with elevation above 1.5m and the area above 10 m² is building.(2)The value of NDVI over 0.17 is vegetation, then the elevation of 1m is selected to separate grassland from tree. (3)After merging open spaces and shaded open spaces, road and shaded road, the compactness of 3.5 is used to discriminate road and open spaces and increase the classification accuracy.

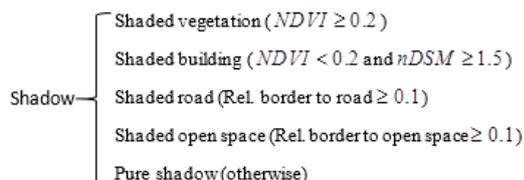


Figure 5. The rule set of shadow areas

There are 269 validation samples used to make the classification assessment in region A: tree(80), grassland(29), building(91), road(48), open spaces(7), water(2) and pure shadow(12), which were generated randomly using ground truth ROIs (region of interest). We assessed the classification accuracy by using confusion matrix (Yi, 2011) (Table 1, the brackets content for the legend color), the overall accuracy was 93.1%. The table revealed that road and open spaces are easy to be define as the same object due to the fact that the road and open spaces have the same spectral and height information. The pure shadow misrecognized as building and shaded road because the shadow covers the road which makes the discontinuous of the road for getting the accurate shape information. At the same time, the shaded road and open space were reclassified by adjacency relationship analysis (the border to neighbors), the wrongly classified neighbor objects (such as road and open spaces) also cause classification errors. In this

experiment, shaded vegetation and shaded building were effectively distinguished by NDVI and nDSM. And the height information can make it easy to differentiate the buildings from road and open spaces the user accuracy and producer accuracy of building, road, water, grassland and tree are over 85%. However the pure shadow was not clearly defined, the user accuracy and producer accuracy are only 55.56% and 41.67%.

3.4 Verification

To verify the applicability of this method, the corresponding classification procedure of region A was applied to region B, the classification results of region B, as shown in Fig.7, the result of confusion matrix as shown in Table 2. All kinds of ground objects have the similar classification accuracy, and automatic classification accuracy reached at 88.42%. This experimental result indicated that this method can be extended to similar research area with high applicability and efficiency in extracting the urban ground objects in a large area. The part of water was misclassified to building, this is because the coastal waters of region B has shadow, so it is easy to segment the shadow to the water, and different water depth and sediment concentration are also the big reasons.

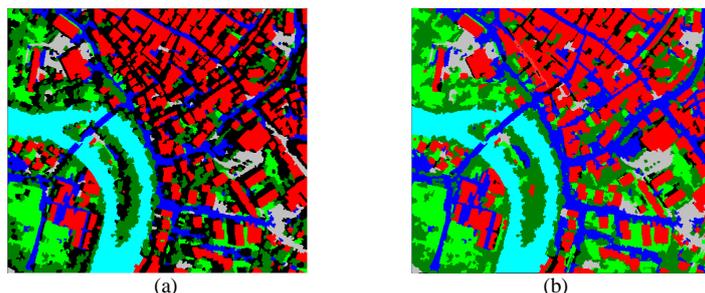


Figure 6 (a)The initial classification result of A (b)The final classification result of A

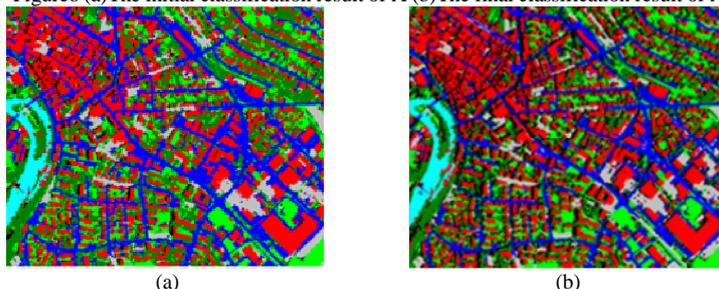


Figure 7. (a)The initial classification result of B (b) The final classification result of B (the legend color is on Table 1)

class	Tree	Grassland	Building	Road	Open space	Water	Pure shadow	User accuracy (%)
Tree(dark green)	80	2	1	0	0	0	1	95.24
Grassland(green)	0	26	0	0	0	0	1	96.3
Building(red)	0	0	90	1	0	0	2	96.77
Road(dark blue)	0	0	0	43	2	0	3	89.58
Open space(grey)	0	0	0	1	5	0	0	83.33
Water(blue)	0	0	0	0	0	2	0	100
Pure shadow(black)	0	1	0	3	0	0	5	55.56
Producer accuracy(%)	100	89.66	98.9	89.58	71.43	100	41.67	

Total accuracy: 93.3% Kappa coefficient: 0.9102

Table 1. The confusion matrix of classification on region A

class	Tree	Grassland	Building	Road	Open space	Water	User accuracy (%)
Tree	346	33	7	7	0	0	87.15
Grassland	8	202	0	2	2	0	92.66
Building	1	0	421	11	1	2	96.12
Road	0	5	3	247	14	0	86.97
Open space	0	0	2	9	18	0	62.07
Water	0	0	0	0	0	3	100
Pure shadow	0	10	5	17	0	0	54.29
Producer accuracy(%)	97.46	80.8	96.12	84.3	51.43	60	

Total accuracy: 88.42% Kappa coefficient: 0.8497

Table 2. The confusion matrix of classification on region B

class	Tree	Grassland	Building	Road	Open space	Water	Pure shadow	User accuracy(%)
Tree	104	13	14	0	0	0	0	79.39
Grassland	7	83	6	4	0	0	5	79.05
Building	0	1	204	7	0	2	2	93.58
Road	0	3	15	61	21	0	0	58.1
Open space	0	1	1	2	22	4	7	84.62
Water	0	0	1	5	0	14	0	70
Pure shadow	0	1	0	7	0	4	20	62.5
Producer accuracy(%)	93.7	81.37	84.64	70.93	50	63.63	64.51	

Total accuracy: 79.75% Kappa coefficient: 0.7403

Table 3. The confusion matrix of classification on region B (Hierarchical method)

We also compared this method with hierarchical object oriented classification which is provided by Chen Y (2009). Block body was extracted by using brightness at first, we used NDVI and mean_SD(the mean of SD_NIR, SD_R, SD_G and SD_nDSM) distinguish water and shadow. Secondly, vegetation was masked by NDVI, tree and grassland was obtained by nDSM, and the LBW and shape index were used to separate road and open space, the result of confusion matrix as shown in Table 4, the user accuracy and producer accuracy of objects(except building) are below 80%, which can validate that our method more effective and automatic.

4. CONCLUSIONS

An object-oriented classification method which integrates aerial image and LiDAR data was proposed in this paper. The precision and recall measures were used to select the optimal segmentation scales of different urban ground object, which made a quantitatively analysis and reduced errors of subjective judgments. Then we synthesize the optimal segmentation results to get object boundaries. ReliefF algorithm was an effective method to choose the optimal feature subset, which measured the feature importance of the objects and improved automaticity. Finally, the multiple classifier combination method was used to get the classification result. These classification experiments of two regions prove that this method is particularly suited for urban ground object extraction. Especially this method subdivides the shadow area by doing the post classification with rules and effectively reduces classification errors. Our accuracy assessment results show that the method has strong applicability in extracting urban ground object information, this method is worth promoting.

However, there are still some issues that should be studied further. In segmentation, some mixed-objects were generated, it could affect the classification accuracy to some extent. For example, due to the differences of shape

and size of buildings, segmentation boundary doesn't match with natural boundary of complex buildings in shape; there are also some problem when height data as the extra wave band was segmented because of the inevitable vertical error of nDSM; moreover because of the same spectral characteristics and elevation information, the contour of low trees couldn't extract accurately, some low trees and grassland couldn't be separated, some roads and adjacent open spaces are always divided into another object. To deal with these problems, it may be necessary to extract other LiDAR derived features (such as curvature, multi-returns information and so on) for classification and it requires further study.

5. ACKNOWLEDGE

The paper was supported in part by the Fundamental Research Funds for the Central Universities under Grant 2011QD03, the 2014 geographical ground monitoring open source project in the state key laboratory of mapping and geographic information (2014NGCM), the National Natural Science Foundation of China under Grant (61405204), the Innovation Training project of University Students in China University of Mining & Technology (Beijing)(C201502030).

REFERENCES

- Beckonert O, Bollard M E, Ebbels T M D, et al. NMR-based metabonomic toxicity classification: hierarchical cluster analysis and k-nearest-neighbour approaches. *Analytica Chimica Acta*, 2003, 490(1): 3-15.
- Chen B, Wang J, "Combining multiple classifiers for thematic classification of remotely sensed data," *Journal of Remote Sensing*, 2005, 9(5): 555-563.
- Chen Y.,D. Wen, L.Jing, et al, "Shadow information recovery in urban areas from very high resolution satellite imagery," *International Journal of Remote Sensing*. 2007, 28(15): 3249-3254.

- Chen Y, Su W, Li J, et al, "Hierarchical object oriented classification using very high resolution imagery and LIDAR data over urban areas," *Advances in Space Research*, 2009, 43(7): 1101-1110.
- Charaniya A P, Manduchi R, Lodha S K, "Supervised parametric classification of aerial LiDAR data," //Computer Vision and Pattern Recognition Workshop. CVPRW'04. Conference on. IEEE, 2004: 30-30.
- Grigillo D, Kanjir U, "Urban object extraction from digital surface model and digital aerial images," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2012, 1(3): 215-220.
- Haala N, Anders K H, "Acquisition of 3D urban models by analysis of aerial images, digital surface models," and existing 2D building information//AeroSense'97. International Society for Optics and Photonics, 1997: 212-222.
- Hamedianfar A, Shafri H Z M, Mansor S, et al, "Improving detailed rule-based feature extraction of urban areas from WorldView-2 image and LiDAR data," *International Journal of Remote Sensing*, 2014, 35(5): 1876-1899.
- Moran E F, "Land cover classification in a complex urban-rural landscape with QuickBird imagery," *Photogrammetric engineering and remote sensing*, 2010, 76(10): 1159.
- Shao Y, Lunetta R S. Comparison of support vector machine, neural network, and CART algorithms for the land-cover classification using limited training data points. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2012, 70: 78-87.
- Syed S, Dare P, Jones S, "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. Melbourne: Melbourne: Spatial Science Institute, 2005: 512-522.
- V. Kecman, *Learning and Soft Computing, Support Vector machines, Neural Networks and Fuzzy Logic Models*, The MIT Press, Cambridge, MA, 2001.
- Yi , "Uncertainty analysis in Object-based Remote Sensing Image Classification," Wuhan: Wuhan University, 2011.
- Zhang G, Yi L, "Feature selection using rough set theory for object-oriented classification of remote sensing imagery" //Geoinformatics (GEOINFORMATICS), 2012 20th International Conference on. IEEE, 2012: 1-7.
- Zhang G., "Multi-scale remote sensing image segmentation based on granularity theory," Wuhan: Wuhan University, 2010.