Multiple Features Based Object-Oriented Buildings Extraction from VHR Image

Xu Song *, Guoying Liu, Hongyu Zhou

School of Computer and Information Engineering, Anyang Normal University, Anyang, 455000, Henan, China
* Corresponding author: Xu Song

Abstract: A multiple features based object-oriented method is proposed to solve the issue of automatically extracting buildings in high resolution remote sensing images. This method uses multi-scale segmentation to get image objects, and then executes classification; in this process, using spectral, texture and context features information to extract buildings from image. At last, by using the spatial features of the buildings, it can optimize the building recognition results. Experiments show that this method can improve the precision of buildings extraction.

Keywords: object-oriented; buildings extraction; remote sensing; multiple features; segmentation

1. Introduction

The goal of remote sensing is to extract information from image. Man-made objects (such as roads, squares and buildings, etc.) are important content of target identification and information extraction for remote sensing. High-resolution satellite remote sensing observer ground objects in a very fine way, the obtained images can more clearly express the spatial structure features and texture information of surface, and distinguish finer internal composition of ground objects. The edge information of ground objects is also clearer. It provides an effective way for extraction of Man-made objects.

With the rapid development of urban construction, the building has become one of the most vulnerable update targets in geographic data, and the updating workload is huge. Therefore, the automatic extraction of buildings from very high resolution remote sensing image is necessary. People have proposed many models and methods for building extraction, which can be divided into the following three categories:

1. Extracting buildings based on the underlying linear feature [1-2]. Such methods first extract the edge of the image, then extract straight lines according to the edge, and group them according to certain criteria, at last extract buildings by linear graphs.

2. Extracting buildings based on LiDAR data combined with high-resolution remote sensing image [3-4]. These methods use the elevation information provided by LiDAR data to separate non-ground points, then separate the building points from the non-ground points, at last, combine with edge information provided by remote sensing data to extract outlines of building;

3. Extracting buildings based on object-oriented method [5]. Such methods first execute image segmentation, and use the homogeneity of pixels to constitute different sizes of image objects, and then use the spectral information, shape features, texture features and context of the objects to extract targets.

In the high resolution remote sensing images, due to the "noise" produced by increased resolution, it causes the edge information is not obvious, and the complexity of the image background will also generate a lot crushing edge which is difficult to deal with. So there will be a lot of error detection and missing detection of buildings in the first class which use edge detection method, and there is a great gap for the practical application. The second class, which using elevation information provided by LiDAR data to separate building points, use the remote sensing data to extract the outer outlines of the buildings. Because of such auxiliary information is not always available, these methods also have some limitations.

An object-oriented image analysis is a new information extraction method in recent years, which extracts homogeneous regions by image segmentation, and then extracts targets area by analyzing of feature of each region. The image processing is not based on pixel-by-pixel, but rather based on combination the homogeneous regions. This method extracts the target objects of remote sensing image in a higher level, in order to reduce loss rate which is high in the traditional pixel-based classification method. It can overcome the limitations of pixel-based analysis. This approach avoids the problem of edge detection, which has obvious advantages for the first class method, and it does not require any auxiliary prior knowledge. This paper studies the features of buildings, adding human's thinking, fully exploits the abundant spectral feature, shapes feature and texture feature and implicit richer deep spatial semantic features of objects in remote sensing images, builds classification system of building extraction. Experiments show that this method makes the classification results more precise.

2. Algorithm Description

In this paper, the method of buildings extraction is as follows:

1) Using multi-scale segmentation for the image to create image object layer from large-scale to small-scale;

2) In the large-scale image object layer. Using spectral and other features to extract potential areas of buildings based on knowledge rules fuzzy classification;

3) In the small-scale image object layer, using shape, texture and other features to separate buildings from potential areas of buildings based on the CART (Classification And Regression Trees) classification;

4) Using the spatial features of buildings to correct the building recognition results, and get more accurate buildings information.

2.1. Image Segmentation

Segmentation is an operation which merges homogeneous characteristics image pixel into a meaningful image objects. In this paper, we use FNEA (Fractal Net Evolution Approach) multi-scale image segmentation algorithm, which is a region merging algorithm for local optimal adaptation, which uses a bottom-up region merging operation to form an object, the image is divided into homogeneous objects with more characteristic, these objects combine with spectral features and spatial features under the merger criterion of minimum object heterogeneity.

The purpose of the algorithm is to achieve the minimize weight heterogeneity of the image objects after image segmentation. Only considering the minimum heterogeneity of the spectrum will lead to the polygon boundary of image objects relatively broken, so it is often using the spectral heterogeneity and the spatial heterogeneity to achieve the minimize weight heterogeneity. Before segmentation, it also needs to determine the impact of heterogeneity of tightness. Only guarantee the spectral heterogeneity, the spatial heterogeneity and the tightness heterogeneity are minimum, it can ensure that average heterogeneity of all objects in the entire image is minimum.

The heterogeneity f of any image object is calculated by four variables: w_{color} (spectral factor), w_{shape}(shape factor), h_{color} (spectral heterogeneity), h_{shape} (shape heterogeneity), and $w_{color} + w_{shape} = 1$. w is the spectral factor, which its range is 0-1.

$$f = w \cdot h_{color} + (1-w)h_{shape} \quad (1)$$

Spectral heterogeneity h_{color} is not only related to the number of pixels in objects, but also depends on the standard deviation of each band.

$$h_{color} = \sum_c w_c (n_{Merge} \cdot \sigma_c^{Merge} - (n_{Obj1} \cdot \sigma_c^{Obj1} + n_{Obj2} \cdot \sigma_c^{Obj2})) \quad (2)$$

σ_c is the standard deviation of internal pixel value of objects, n is the number of pixels.

Shape includes two sub-factors: the tightness h_{cmpct} and smoothness h_{smooth}

$$h_{shape} = w_{cmpct} \cdot h_{cmpct} + (1 - w_{cmpct}) \cdot h_{smooth} \quad (3)$$

h_{cmpct} and h_{smooth} are depended on pixel number of the object n, side length of the polygons l and the smallest side length b in the same area of polygons.

$$h_{conpct} = n_{Merge} \cdot \frac{l_{Merge}}{\sqrt{n_{Merge}}} - (n_{Obj1} \frac{l_{Obj1}}{\sqrt{n_{Obj1}}} + n_{Obj2} \frac{l_{Obj2}}{\sqrt{n_{Obj2}}}) \quad (4)$$

$$h_{smooth} = n_{Merge} \cdot \frac{l_{Merge}}{b_{Merge}} - (n_{Obj1} \frac{l_{Obj1}}{b_{Obj1}} + n_{Obj2} \frac{l_{Obj2}}{b_{Obj2}}) \quad (5)$$

This algorithm needs appropriate parameters for multiple-scale image segmentation to build the network hierarchy of image objects, also considers the multi-level of ground surface and uses multi-scale object-level network structure to reveal surface features. It overcomes the shortcomings of the original data source which is fixed in scale.

2.2. Feature Space

Feature space as an abstract feature set of a variety feature elements in geospatial and image space is a key to understand remote sensing images. Taking advantage of the spectral, shape and other features of the object, it makes the resulting segmentation "homogeneous object" having these features. After building image object layer in different scales on multi-scale image segmentation, it reflects spectral information, texture information and spatial structure information between objects in the real world through a variety of features of image object. It objectively describes the characteristics of the real object, and the target object and the actual feature have high consistency, which provides the basis for subsequent classification. In this study, we use spectral features, shape, texture and spatial feature of buildings to extract targets.

2.2.1. Spectral features

The primary classification of remote sensing image is based on spectral features of ground object. Spectral features reflect the electromagnetic radiation of ground object in different bands, which can be used as original features variable for remote sensing image classification. It includes mean, standard deviation, brightness, gradation ratio, etc., and the analysis of the spectral information of the image can be more effective to extract and classify targets.

Mean spectral $\bar{C}_L$

$$\bar{C}_L = \frac{1}{n} \cdot \sum_{i=1}^{n} C_{L_i} \quad (6)$$

Standard deviation δ_L

$$\delta_L = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} \left(C_{L_i} - \bar{C}_L \right)^2} \quad (7)$$

Brightness b

$$b = \frac{1}{n_l} \cdot \sum_{i=1}^{n_l} \bar{C}_i \quad (8)$$

Where C_{L_i} is the gray value of band, n is number of pixel, n_l is band number of image.

2.2.2. Shape features

The shape information of a remote sensing image is a very important basis for visual discrimination, but almost impossible to be extracted in the traditional methods. In computer vision, with respect to the underlying features, such as color or texture, the shape is a important feature to portray the object of image and region features, also is an important means to describe high-level visual features (such as target, object, etc.), which contains more semantic information than the underlying features. It delivers more specific and accurate semantic information. Geometric feature of the building is on the basis of shape and geometry, the purpose of constructing shape parameters is to use the mathematical tools to accurately reflect the shape in human describing.

Shape factor s

$$s = \frac{e}{4 \cdot \sqrt{A}} \tag{9}$$

Tightness c_1

$$c_1 = \frac{nm}{a} \tag{10}$$

Tightness c_2

$$c_2 = \frac{4\pi S}{l^2} \tag{11}$$

Where e is the length of the boundary of the object, A is area of the object, m and n respectively represent the length and width of the image object, a is the pixel number of the object, S is the area of all the polygon, l is area of a circle which the perimeter is equal.

2.2.3. Texture features

Texture feature is often seen as a local nature in image, or a measure of the relationship between pixels in the local area. Texture commonly is using the method GLCM (Grey Level Concurrence Matrix) to describe, which is proposed by Haralick. The object-oriented texture extraction is calculated for pixels in each object itself, and includes all edge pixels in order to reduce the edge effect. The size and shape of the calculating window varies for each other. Description of texture features based on the object is closer to the true performance of different geographical phenomena. The texture features include homogeneous texture, contrast, entropy, angular second moment and correlations.

Moment of Inertia I

$$I = \sum_i \sum_j (i-j)^2 P(i,j) \tag{12}$$

Angular second moment E

$$E = \sum_i \sum_j [P(i,j)]^2 \tag{13}$$

Entropy H

$$H = \sum_i \sum_j [P(i,j)] \lg P(i,j) \tag{14}$$

Where i and j respectively represent the ranks of the image, $P(i,j)$ is the joint probability of grayscale i and j.

2.2.4. Spatial features

When building a network of object, it needs to analyze the spatial relationships between objects at various levels. Spatial relationships can be summarized into two categories:

- Topological relationships between objects in the same layer
- Inheritance relationships between objects in the adjacent layers

According to changes of the relative positions between objects, the topological relations can subdivided into disjoint, meet, contains; According to the number and scale factor between objects, the inheritance can subdivided into one-to-two (parent with two sub-objects) and one-to-more (parent with more than two sub-objects), as shown in Figure 1.

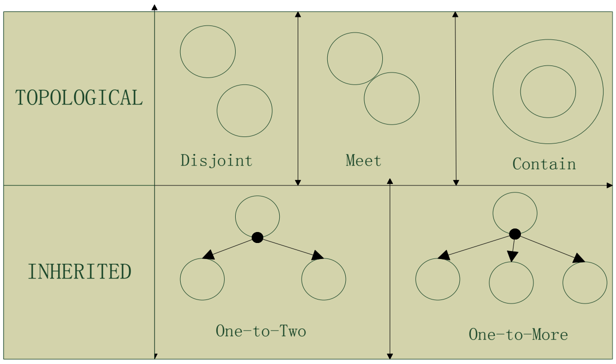


Figure 1. Examples of spatial relations in objects network

3. Experiments and Analysis

3.1. Multi-Scale Segmentation

Before buildings extraction, first, we execute multi-scale segmentation to obtain image objects; (1) Set segmentation parameters, including the weight of each band, that is to determine the importance of each band in the segmentation process. In the color factor, it required to determine the weight of spectral factor and shape factor according to image texture features and thematic information. In the shape factor, it required to determine the weight of the tightness factor and smoothness factor according to structural properties of the most feature classes. (2) Use any one image pixel as the center began to execute segmentation, in the first time, the single pixel is as a smallest polygon object involved in calculation of heterogeneity; after the completion of first segmentation, it use the generated polygon to execute second segmentation and calculate the difference between the value of heterogeneity and the threshold which set in advance, if the value of heterogeneity is less than the threshold, multiple segmentation is executed, otherwise the image segmentation works is stop, it forms a fixed scale image object layer.

The scale parameter depends on the purpose and the resolution of image. According to several experiments, it is suitable to select large-scale 40 and small-scale 10 to create two image object layer. From the segmentation results in Fig. 2(b), Fig. 3(b), the green space, buildings and roads are divided into larger objects polygons in the large-scale, which well reflect the features of the objects. But in Fig. 3(c), Fig. 3(c), when the split-scale becomes 10, the green space, buildings and roads become even more broken on the basis of large scale, most of the buildings image becomes smaller. The integrity of the building should be one or several polygons; the gray value inside the boundary is relatively homogeneous.

3.2. Obtain Potential Building Region

Based on the above segmentation, according to the significant difference between background of ground objects and features of target objects, these objects can be separated under large scale segmentation. It is suitable to adopt fuzzy classification based on knowledge rules by using the object feature. First, we execute rough classification, followed by refine the classification.

First, use the spectral features to create vegetation and non-vegetation classes. Building knowledge rules as follows:

IF (blue band ratio values < 0.2009) AND (Near infrared band ratio value > 0 .3154)

THEN

object = vegetation class

ELSE

object = non-vegetation classes.

Second, use spectral and spatial features to create a subclass of non-vegetation classes: potential buildings and other classes. Building knowledge rules are as follows:

IF (blue band mean > 241.5) AND (near-infrared mean < 0.1376) AND (blue band ratio values > 0.2298) AND (mean difference near-infrared and adjacent objects < 54.47) AND (tightness < 2.2) AND (tightness > 1. 4)

THEN

object = potential building class

ELSE

object = other classes.

After the classification, vegetation, water and shadows can be distinguished; the potential building region can be obtained, which is the basis for the next precise extraction.

3.3. Extraction Building

On the small scale, the potential building regions which are extracted in the previous include buildings, roads and squares. Because these ground objects have overlapping part in spectral feature, using shape, texture and other features can separate them. It is suitable to use the length-width ratio, shape index and entropy of GLCM (GLCM) to separate buildings from other ground objects. Since the shape of the building is regular, so the shape index is small, after repeated comparisons, the shape index is set to below 2.80; due to length-width ratio of building is generally large, it is suitable to set the length-width ratio above 1.26; Buildings generally have a certain texture, so its entropy is large, the entropy is set to above 7.90. According to the above feature parameters, CART classification method is used to extract buildings.

3.4. Optimization Results

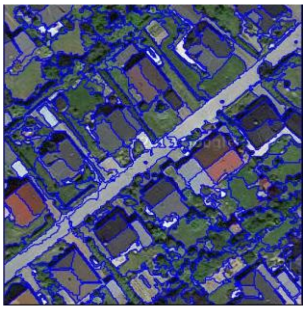
The spatial feature of the building contains spatial relationship among buildings and spatial relationship between buildings and surrounding objects. The feature of spatial feature of objects and its corresponding optimization as follows:

Some buildings have the same shape and arrange in neat rows. Due to the impact of improved resolution and complex background, it will be produce a large number of crushing edge in image. It can carry out intersection calculation in the regions which the relative boundary between them is 0 to extract the missing information.

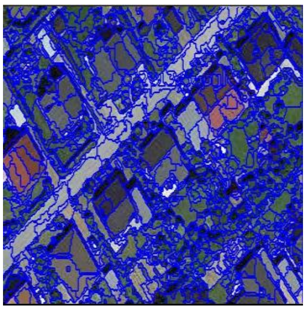
When buildings is extracted, there will produce some cracks, and these objects are surrounded by buildings. These objects can be eliminated by using morphology algorithms

After buildings extraction, there may have some small objects confused with the building extraction, such as car, boat, etc.. These objects can be eliminated by using area feature.

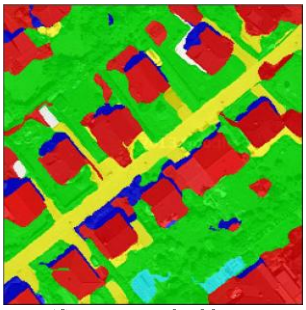
(a) Given image



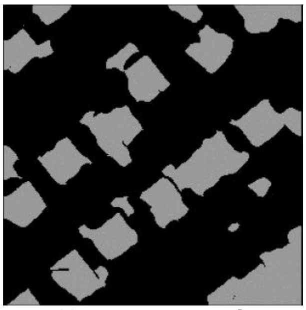
(b) Multi-scale segmentation -40



(c) Multi-scale segmentation 10



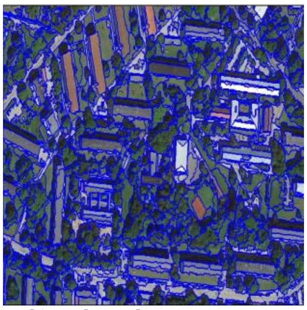
(d) Extracting buildings



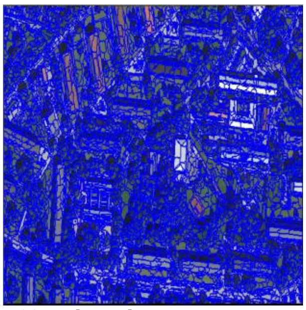
(e) Optimization results

Figure 2. Segmentation results

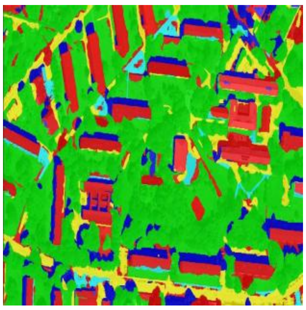
(a) Given image



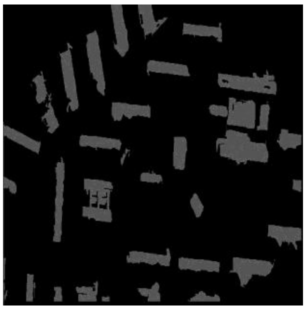
(b) Multi-scale segmentation -40



(c) Multi-scale segmentation 10



(d) Extracting buildings



(e) Optimization results

Figure 3. Segmentation results

3.5. Accuracy Assessments

In order to quantitatively evaluate performance of the algorithm, the manual extracting result is used to be a reference to evaluate the results of the building extraction. Using the correct rate, misclassification rate and missing rate as the evaluation index to statistic results. Let the area counting by this method is S_1, and area counting by manual extraction is S_2,the correctly classified area is T, the misclassification area is F, the missing area is M, and the area is calculated by pixels. The correct rate $C_T=T/S_1$, the misclassification rate $C_F=F/S_1$, the missing rate $C_M=M/S_2$. The statistics is as follows Table 1:

Table 1. Comparison of the building extraction

	Correct rate	Misclassification rate	Missing rate
Fig.1	92.54%	7.46%	8.13%.
Fig.2	93.87%	6.17%	7.52%

4. Conclusion

In this paper, we propose an object-oriented method, which first obtains homogeneous object by image segmentation, and then uses the spectral, shape and other feature to extract building objectives, at last, uses the spatial features as auxiliary information to optimize building recognition results. This method effectively reduces the interference of other targets. Its accuracy is significantly improved over traditional methods. Meanwhile, how to get more feature and information of image and improve the efficiency of interpreting is worth pursuing in the future research.

Acknowledgment

This work was financially supported by the National Natural Science Foundation of China under Grant 41001251, U1804153; the Science and Technology Development Project of Henan Province of China under Grant 202102110115, 212102210502; the Science and Technology Development Project of Anyang of China under Grant 2021C01GX018.

References

[1] Tao, W.B.; Liu, J.; Tian, J.W. A New Approach to Extract Rectangle Building Automatically from Aerial Images. Chinese Journal of Computers, 2003, 26(7):866-873.

[2] Cha, J.; Cofer, R.H.; Kozaitis S.P. Extended Hough Transform for Linear Feature Detection. Pattern Recognition, 2006, 39(6):1034-1043.

[3] Cheng, L.; Gong, J.Y. Building, Boundary Extraction Using Very High Resolution Images and LiDAR. Acta Geodaetica et Cartographical Sinica, 2018, 37(3):391-393.

[4] Sampath, A.; Shan, J. Building Boundary Tracing and Regularization from Airborne Lidar Point Clouds. Photogrammetric Engineering & Remote Sensing, 2017, 73(7):805-812.

[5] Qiao, C.; Luo J.C.; Wu Q.Y. Object-Oriented Method Based Urban Building Extraction from High Resolution Remote Sensing Image. Geography and GeoInformation Science, 2018, 24(5):36-39.