Automated hierarchical segmentation of high-resolution remote sensing imagery with introduced relaxation factors

DENG Fuliang, TANG Ping, LIU Yuan, YANG Chongjun

State Key Laboratory of Remote Sensing Science, Jointly Sponsored by the Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences and Beijing Normal University, Beijing 100101, *China*

Abstract: This paper proposes a new automated hierarchical segmentation method with introduced relaxation factors for processing high-resolution remote sensing imagery, which aims to provide a theoretical framework in setting the scale parameters and reducing the influence of human factors. The first relaxation factor is used to adjust the heterogeneity between the image-objects to be merged, thus improving the speed of the entire segmentation by controlling the number of image-objects in each recursive merging. With the mean of the heterogeneity between image-objects taken as the cardinality, the second relaxation factor is introduced to control the scaling parameter of the levels exported in the process of segmentation, automatically producing multi-scale hierarchical segmentation results. The experimental results show that this method produces segmentation with higher quality, which meets the accuracy requirements of further image analysis and geographic object extraction. Other theoretical and practical contributions of this method include reducing the influence of human factors and improving the level of automation in segmentation. Further investigation is still required with respect to processing the boundaries of geographic objects with complex image, and increasing the compactness and smoothness of image-objects.

Key words: relaxation factor, high-resolution satellite imagery, automated hierarchical segmentation, criterion of heterogeneity, region adjacency graph

CLC number: TP751.1 **Document code:** A

Citation format: **Deng F L, Tang P, Liu Y and Yang C J. 2013. Automated hierarchical segmentation of high-resolution remote sensing imagery with introduced relaxation factors. Journal of Remote Sensing, 17(6): 1492–1507 [DOI: 10.11834/jrs.20133031]**

1 INTRODUCTION

High-Resolution Satellite Imagery (HRSI) possesses many new characteristics, such as presenting geographic objects with more visible geometrical structures, clearer location layout, and more precise information about texture and size, containing multi-scale target objects and so on. Because of these new characteristics, the traditional pixel-based techniques of processing and analyzing low-to-middle-resolution remote sensing imagery cannot effectively decode HRSI (LI, et al., 2012; Sun, et al., 2011; Liu, et al., 2009). Posing the question "what's wrong with pixels?" Blaschke, et al. (2008) pointed out that traditional pixel-based methods could barely meet the practical requirements in processing and analyzing HRSI. Baatz, et al. (2006) also stated that important semantic information cannot be represented in single pixels, and semantic interpretation should rely on meaningful image-objects and their mutual relations.

Consequently, new object-based image analysis techniques were created (Blaschke, et al., 2008; Hay, et al., 2006), incorporating the spectra, shape, texture, and topological information contained in the image (Blaschke, et al., 2010; Korting, et al., 2011). In most cases, they start with image segmentation, and then proceed with classification and identification of the basic processing units in the form of image-objects (segments) for geographic object extraction (Blaschke, et al., 2010). Therefore, as a primary method of object acquisition, image segmentation has become a critical step in object-based and object-oriented image analysis (Baatz, et al., 2006; Dezsö, et al., 2012a; Murthy, et al., 2012).

Extensive and in-depth research has been carried out regarding the segmentation of HRSI, and Hierarchical Segmentation (HS) is a hot topic in current research. The renowned commercial software eCognition provides a multi-scale segmentation tool based on a typical HS method. However, it has been found (Lin, et al., 2012; Wang, et al., 2009) that with this method, a high degree of human interference exists in setting the scale parameters without the support of a theoretical framework. Moreover, the HS result is not obtained in one implementation of segmentation, but through multiple implementations defined with

---

Received: 2013-02-22; **Accepted:** 2013-06-29; **Version of record first published:** 2013-07-06

Foundation: High Resolution Earth Observation System of Major National Science and Technology projects (No. 02-Y30A04-9001-12/13)

First author biography: DENG Fuliang (1982—), male, Ph. D. candidate, his research interests are remote sensing image processing and application. E-mail: fldeng8266@gmail.com

Corresponding author biography: TANG Ping (1968—), female, professor, her research interest is remote sensing image processing. E-mail: tangping@irsa.ac.cn different scale parameters, which also determine the scale of the output levels in constructing the HS results.

This paper proposes a new method of constructing HS results for HRSI. Unlike the method provided in eCognition, a relaxation factor is used to adjust the heterogeneity between the regions/image-objects to be merged and improve the speed of merging; with the mean heterogeneity between regions taken as the cardinality, another relaxation factor is introduced to control the scale parameters of the levels to be exported during segmentation, automatically generating multi-scale HS results.

2 HIERARCHICAL SEGMENTATION METHOD AND PROCESS ANALYSIS

HRSI contain semantic multitudes about geographic objects and the space, which can only be represented and depicted using multi-scale approaches (Marr, 1982; Trias-Sanz, et al., 2008). HS is a widely used multi-scale method of constructing a variety of geographic objects in remote sensing imagery. "Hierarchical" in this paper refers to clusters of segments with multiple scales and different levels of detail acquired from the segmentation of a single image. These segment clusters can be used to form hierarchical network representations corresponding to different scales in the same image, as shown in Fig. 1. A layer of segments is called one image-object level, containing multiple regions/image-objects; multiple levels form a group of image-object levels. In such a group, the image-objects are linked in a multi-level network structure: the objects within one level are related by topological relations; and the objects between the upper and lower levels are connected by their inter-level inheritance. Such a network representation of remote sensing image is exactly the goal of HS.

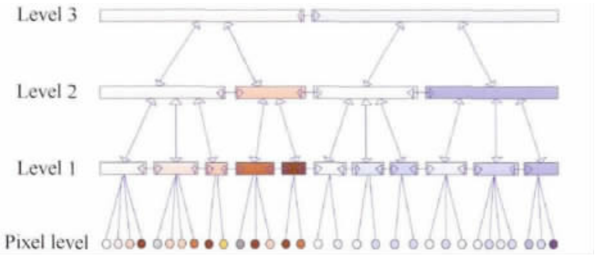


Fig. 1 Multi-scale hierarchical network representation of remote sensing imagery

From the topological relations within each level and the inter-level inheritance, this hierarchical representation of image-objects can well reflect the context information of geographical objects. It is apparently superior to the results obtained from a single segmentation, and regarded as an effective means of fully discovering the application potentials of HRSI (Plaza, et al., 2005).

Currently, the most widely used multi-scale segmentation method is provided by one of the most successful commercial software products, eCognition. The hierarchical network structure introduced above is constructed from the results of multiple implementations of segmentation with different scale parameters. The result derived in each implementation corresponds to each level of image-objects in the hierarchical network structure, with the upper levels corresponding to coarser scales and lower levels to finer scales (Trimble, 2012). Based on the principle of merging regions with the minimum heterogeneity, the regions are merged from the bottom to the top level, and the image-objects are extracted accordingly. This HS result is achieved not through one implementation of segmentation, but through multiple implementations of segmentation with different scale parameters that determine the output levels of regions or image-objects. In eCognition, different scale parameters are usually tried in carrying out the segmentation, and by comparing their segmentation results, one level or multiple levels of image-objects are used in later image analysis and geographic object extraction. Practical experience of using this HS method has shown that because setting the scale parameters lacks the support of a theoretical framework, the segmentation result is subject to significant influence of human factors, and the resultant low-level automation can hardly satisfy the requirements of practical applications.

Careful analysis of the HS method provided by eCognition shows that it mainly comprises three steps: first, choosing the criterion of heterogeneity used in region merging; second, developing the strategy for region merging; finally, defining the output rules and the conditions to terminate merging.

3 NEW HIERARCHICAL SEGMENTATION (HS) METHOD WITH INTRODUCED RELAXATION FACTORS

To improve the level of automation in the three primary steps introduced in Section 2, a new automated HS method with introduced relaxation factors is proposed in this paper. Major innovations in this new method are summarized as below. (1) A new criterion of heterogeneity is proposed; (2) A relaxation factor is used to adjust the heterogeneity between regions to be merged; (3) With the mean heterogeneity between regions or image-objects as the cardinality, a second relaxation factor is introduced to control the scale of the levels to be exported during region merging, so that multi-scale HS results can be generated in a single implementation of segmentation.

3.1 Criterion of heterogeneity for region merging

The definition of the criterion of heterogeneity is important to the quality of segmentation results. The features selected in the criterion are not only depended on the specific problem but are also related to the type of RS image data (Zhang, et al., 2012). Segmentation with superior quality normally satisfies three conditions (Haralick, et al., 1985): (1) homogeneity of the selected features within regions; (2) sufficient heterogeneity between adjacent regions; (3) concise, smooth, and accurate region boundaries. The first two conditions are called feature criteria, and the last is called a semantic criterion and it is also followed in human perception (Zhang, et al., 2008). In this paper, spectra and shape are selected to define feature criteria.

Spectral heterogeneity is used to represent the difference between the spectra of all pixels within regions or image-objects. The method of representation, using mean square error and its modifications, has been the dominant quantitative performance index and criterion of assessment in image processing for over 50 years (Wang, et al., 2009). In this paper, mean square error is also used in calculating the criterion of spectral heterogeneity, which is represented as

$$h_{\mathrm{MSE}} = \left[\frac{n_i n_j}{n_i + n_j}\sum_{b=1}^{B}(\mu_{ib} - \mu_{jb})^2\right]^{1/2} \tag{1}$$

where n_i and n_j represent the number of pixels in regions i and j respectively, B is the number of bands in the RS image, and μ_{ib} and μ_{jb} represent the means of the spectra in the b^{th} band of regions i and j respectively.

The other feature in defining spectral heterogeneity is the maximum normalized standard deviation, which is computed as

$$\sigma_i = \max\{\sigma_{ib} : b = 1, 2, \cdots, B\} \tag{2}$$

where σ_{ib} represents the normalized standard deviation of the b^{th} band of the spectra in region i, which is computed as

$$\sigma_{ib} = \frac{1}{\mu_{ib}}\sqrt{\frac{1}{n_i - 1}\left[\sum_{x_p \in \boldsymbol{X}_i}(x_{pb})^2 - n_i(\mu_{ib})^2\right]} \tag{3}$$

where $\boldsymbol{X}_i$ represents the i^{th} region, μ_{ib} represents the mean of the b^{th} band of the spectra in region i, n_i is the number of pixels in Region i, x_p represents the p^{th} pixel in Region i, x_{pb} is the b^{th} band of the spectrum of pixel x_p. If a region contains only one pixel, the value of σ_{ib} is 0.0.

Combine Eq. (1) and Eq. (3), and the criterion of spectral heterogeneity for merging two regions are computed according to

$$h_{\mathrm{spectral}} = h_{\mathrm{MSE}}\left[1.0 + w_{\mathrm{stddev}}\frac{|\sigma_i - \sigma_j|}{\sigma_i + \sigma_j}\right] \tag{4}$$

where σ_i and σ_j represent the maximum normalized standard deviation of regions i and j respectively, w_{stddev} represents the weighting of the maximum normalized standard deviation, which ranges from 0 to ∞. with default value of 1.0.

Larger values of Eq. (4) suggest that the spectral information in the two regions is more similar and has greater likelihood to be merged. Vice versa, two regions are less likely to be merged if a large difference exists in their spectral information. Normalized standard deviation of the spectra is primarily used to increase the impact of assessing the spectral similarity between regions. Proper weighting factors can also lead to better spectral heterogeneity between the image-objects in the segmentation result.

Shape heterogeneity is used to represent the differences in the shape of regions or image-objects. To obtain concise, smooth, and accurate region boundaries, shape compactness is included in computing the criterion of heterogeneity in this paper. As an important shape property, compactness may be represented in multiple ways in terms of the geometrical parameters of the object and related to the scale parameter (Zhang, 2012). Compactness of the image-objects or regions can be indirectly depicted in the comparison of regions with ideal shapes such as square and circle (Marchand-Maillet, et al., 1999). In this paper, Aspect Ratio (AR) is selected as one criterion of shape heterogeneity, which is calculated according to

$$AR = l/h \tag{5}$$

where l and h represent the length and width of the minimum box enclosing the object. The AR of square and circle objects has the minimum value of 1, narrow and long objects have AR values larger than 1, and AR increases with the degree of narrowness of the object.

The other feature selected in the criterion of shape heterogeneity is Circle Index (CI), which assesses the smoothness of the region boundaries. It is computed as

$$CI = a_{\mathrm{c}}/a_{\mathrm{p}} \tag{6}$$

where a_{p} is the area of the object, and a_{c} is the area of the circle of equivalent circumference to the object. Circular objects with the smoothest boundary have minimum CI of 1, the CI values of other shapes are greater than 1, and increase with the degree of roughness in the boundaries.

Combining the two feature criteria defined in Eq. (5) and Eq. (6), we have the criterion of shape heterogeneity for merging regions i and j

$$h_{\mathrm{shape}} = w_{\mathrm{compact}}\left|\frac{l_{ij}}{h_{ij}} - \left(\frac{l_i}{h_i} + \frac{l_j}{h_j}\right)\right| + (1 - w_{\mathrm{compact}})\left|\frac{a_{\mathrm{c}ij}}{a_{\mathrm{p}ij}} - \left(\frac{a_{\mathrm{c}i}}{a_{\mathrm{p}i}} + \frac{a_{\mathrm{c}j}}{a_{\mathrm{p}j}}\right)\right| \tag{7}$$

where w_{compact} represents the weighting of compactness, and $(1 - w_{\mathrm{compact}})$ is the weighting of smoothness. The lower index ij indicates the new region or image-object obtained from merging regions i and j.

Based on the criterion of spectral and shape heterogeneity introduced above, the general criterion of heterogeneity incorporating both aspects is computed as

$$h_{\mathrm{merge}} = h_{\mathrm{spectral}} + w_{\mathrm{shape}} h_{\mathrm{shape}} \tag{8}$$

where w_{shape} is the weighting of shape heterogeneity, which ranges from 0 to ∞ and has the default value of 1.0. Smaller heterogeneity h_{merge} suggests that the two regions are more similar, and have larger likelihood to be merged. Vice versa, larger h_{merge} indicates that the two regions are less likely to be merged, because of the greater distinction that exists between them.

Spectra contain the most important information in RS data, so spectral heterogeneity normally has the most critical influence on the quality of segmentation results. Shape heterogeneity is included for more compact and smoother boundaries in the segmentation results, reducing the occurrence of regions with incomplete shape; however, the weighting of shape heterogeneity should not be set too large in practical applications (Lv, et al., 2012).

3.2 Optimized merging and segmentation using region adjacency graph with introduced relaxation factors

In HS, constructing the multi-scale details in each level of image-objects usually employs the region-growing method (Cardelino, et al., 2009) with either local (Benz, et al., 2004; Sarkar, et al., 2000) or global merging strategies (Bealieu, et al., 1989). The former locally searches and merges the adjacent regions meeting the criteria for region merging as introduced above, while the latter searches and merges the adjacent regions optimized globally at the whole level, and is capable of being adapted to the changes brought by each merging, so it is more accurate and stable (Dezsö, et al., 2012b). In this paper, the global strategy is used, and optimization is realized in a region adjacency graph illustrating the relationship between image-objects or regions (Tupin, et al., 2005; Xia, et al., 2006).

A region adjacency graph is defined as $\boldsymbol{G} = (\boldsymbol{V}, \boldsymbol{E})$, wherein V_i is a node representing an image-object or region; E_{ij} is a portion of arc with the heterogeneity value between two adjacent regions i and j as its weighting. In the global strategy used in this paper, the portion of arc with the minimum weighting is searched, and the regions adjacent to the arc are merged; in the meanwhile, local topological relations and arc weightings are adjusted. The procedure is repeated until the condition of the terminating region merging is satisfied, and the segmentation is completed. Although this globally optimized merging strategy is highly accurate and strict in theory, it is relatively less efficient, especially when large RS images are processed, so it cannot meet practical requirements. Therefore, a parameter λ is introduced to control the heterogeneity value

$$w_h = \lambda w_{\min} \tag{9}$$

where λ is the introduced relaxation factor in each region merging iteration; $w_{\min}$ is the weighting of the arc with minimum heterogeneity in the region adjacency graph. Adjusting the value of λ can control the number of regions to be merged recursively, i. e. , the regions adjacent to the arcs with weighting $w_{ij} \leqslant w_h$ will be merged. It can be shown that for smaller λ, a smaller number of regions is merged in each recursive iteration; for $\lambda = 1$, the regions merged each time have the minimum heterogeneity. However, larger λ would affect the final accuracy of region merging, so it should not be set too large.

As each level in the HS result corresponds to a scale with different accuracy adapted to the extraction and recognition of different types of geographic objects, the intermediate states of the scales to be recorded in the course of region merging need to be determined. This is usually realized in HS by setting the scales multiple times (Tzotsos, et al., 2006). In this paper, the mean of inter-regional heterogeneity is taken as the cardinality, and another relaxation factor is introduced to automatically control the scale of the level to be exported

$$s_1 = \delta\left(\frac{1}{n}\sum w_{ij}\right)^{1/2} (i \neq j) \tag{10}$$

where δ is the relaxation factor used to control the number of levels to be exported; n is the number of arc portions in the region adjacency graph; w_{ij} represents the heterogeneity value between regions i and j, or the weighting of the arc portion. In the automated HS method proposed in this paper, only the scales of the starting level S_{start} and the final level S_{stop} need to be preset. The scale mentioned here refers to the square root of the mean of all inter-regional heterogeneity values. Taking the square root of the mean is for the purpose of increasing the sensitivity of scale control. During region merging, multiple segmentation results with scales ranging from S_{start} to S_{stop}, and determined by the δ value, are exported in forming the final HS result cluster. In practical applications, the value of δ may be determined according to the particular data and application conditions. Larger δ would lead to fewer levels and vice versa.

3.3 The HS procedure with introduced relaxation factors

Before introducing the HS procedure, the primary segmentation parameters are first listed in Table 1.

Table 1 Primary parameters in the HS method

Parameter	Meaning
S_{start}	Scale of the first output level
S_{stop}	Scale of the last output level scale
w_{stddev}	Weighting of the normalized standard deviation of spectral feature
w_{shape}	Shape weighting
w_{compact}	Compactness weighting
λ	Relaxation factor in each region merging iteration
δ	Relaxation factor used to control the number of levels to be exported

The HS procedure is illustrated in Fig. 2, with the following specific steps.

Step 1 Import the original RS image, and set the segmentation parameters, the majority of which are listed in Table 1.

Step 2 Generate the initial regions or image-objects based on a particular method, or take each pixel as an initial image-object or region.

Step 3 Transform the initial regions in a region adjacency graph, with arc weightings equal to the inter-regional heterogeneity value computed with the criteria introduced above.

Step 4 Merge the adjacent regions satisfying the merging conditions. The scale of the previous level is assumed to be S_{prev}, and the scale of current segmentation is S. If S is larger than $\delta \times S_{\text{prev}}$, execute Step 5; otherwise, obtain the minimum arc weighting $w_{\min}$ in the region adjacency graph, and compute w_h according to Eq. (9). Then put the arc portions with weighting less than and equal to w_h in a temporary group or dimension and sequence them in ascending order. If two arc portions have equal weightings, they are ordered according to the pixel number of the adjacent region with lower number of pixels. If equal numbers of pixel exist in both regions, then they are sequenced according to their region number. The above procedure is repeated until the condition of level exporting is satisfied.

Step 5 Export the segmentation result of the current scale, and record the current scale as S_{cur}, then let S_{prev} equal S_{cur}, which is taken as the reference value of the next level to be exported. If S_{start} is larger than S_{stop}, skip to execute Step 6; otherwise, return to Step 4.

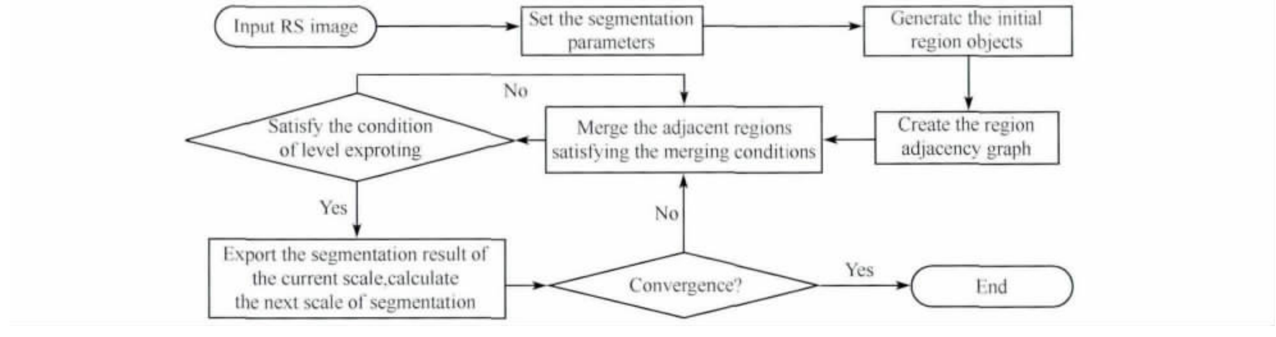

Input RS image
Set the segmentation parameters
Generate the initial region objects
No
Satisfy the condition of level exproting
Merge the adjacent regions satisfying the merging conditions
Create the region adjacency graph
Yes
No
Export the segmentation result of the current scale,calculate the next scale of segmentation
Convergence?
Yes
End

Fig. 2 The HS procedure

Step 6 At the completion of the segmentation procedure, for the convenience of further HS based applications such as image analysis, geographic object extraction, etc., the inherited relations between the sub-objects and super-objects, and the topological adjacency relations within each level, are constructed in the order from fine grains to coarse grains, forming a hierarchical network representation of the same RS image.

In general, the above procedure includes two critical components: (1) region merging, as described in Step 4; (2) level exporting control-based on the method of computing and controlling segmentation scales introduced above, intermediate results satisfying the condition of level exporting are exported during region merging, and the scale of next level to be exported is computed.

4 EXPERIMENTAL RESULTS AND DISCUSSION

Based on the HS method proposed in this paper, a high-resolution satellite image is chosen to conduct HS experiments, for testing the validity of this new method. This image data is an aerial photograph taken in July 2011 in Beijing, with a size of 913 × 811 pixels, spatial resolution of 0.2 m, and three bands of visible light. The original satellite image is shown in Fig. 3.

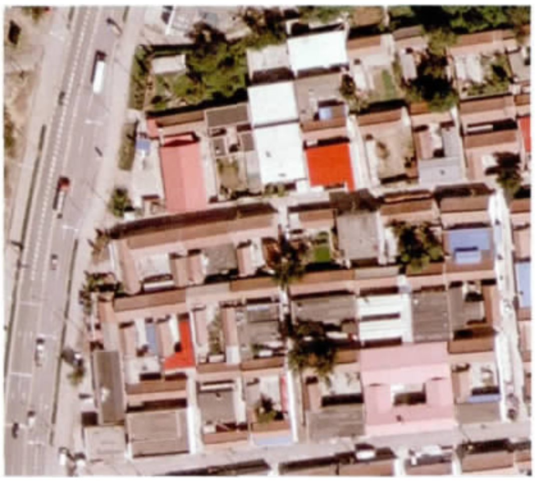
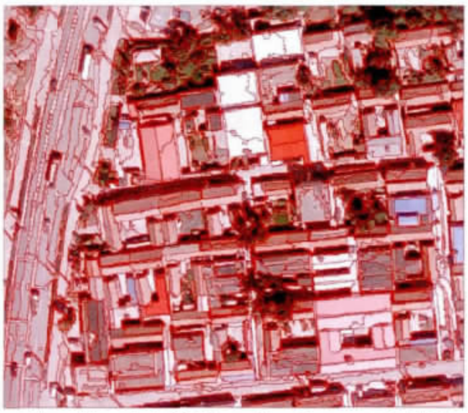


Fig. 3 Original satellite image

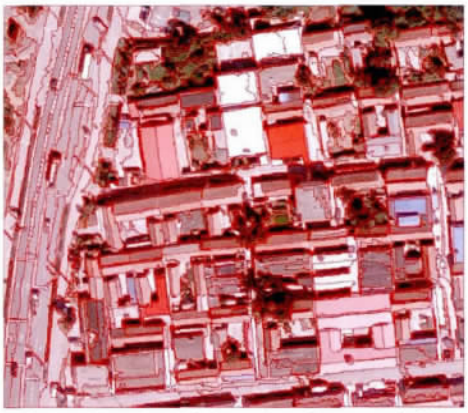
The segmentation parameters are set as below: $S_{\text{start}} = 200.0$, $S_{\text{stop}} = 382.2$, $w_{\text{stddev}} = 0.1$, $w_{\text{shape}} = 40.0$, $w_{\text{compact}} = 0.5$, $\lambda = 1.01$ and $\delta = 1.3$. Despite the large number of parameters to be set, experimental results show excellent segmentation results for a variety of HRSI. In practical applications, the relaxation factor λ controlling region merging is set to be 1.01, to maintain required segmentation accuracy.

Fig. 4 shows the HS result of the original image, including four levels with different scales. The specific segmentation scales and number of regions can be found in the legends.

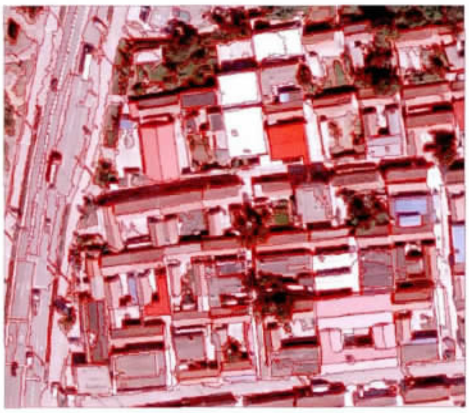
The pictures clearly show the fine-to-coarse segmentation results of the geographic objects, such as the white (dashed) bicycle line on the road, the spots on the roofs of residential buildings, etc. From Level 1 to Level 4, regions of larger and larger scales are merged level by level, and the number of regions/image-objects is gradually reduced, with small regions



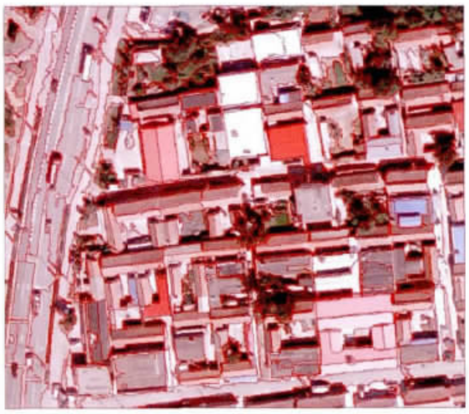
(a) Level 1:scale is 200.0, number of segments is 1166



(b) Level 2:scale is 260.0, number of segments is 766



(c) Level 3:scale is 338.2, number of segments is 500



(d) Level 1:scale is 382.2, number of segments is 408

Fig. 4 The HS results

being merged into large regions. For example, at Level 3, the roofs of many residential buildings have been merged into relatively complete image-objects with boundaries matching more closely to the actual boundaries. At Level 4, the shadow of road ancillary facilities and small blocks on the road have been further merged into bigger image-objects, and the segmentation result appears to be more concise. These experimental results show that the HS method proposed in this paper can be applied to multiscale segmentation of HRSI and generate from fine-to-coarse hierarchical segmentation results.

This segmentation result is also compared with that obtained from eCognition (Version 8.7) by means of visual and quantitative analysis. The segmentation result of Level 4 has 408 image-objects, so the same number of objects is generated in eCognition through trials of different scale parameters (Fig. 5).

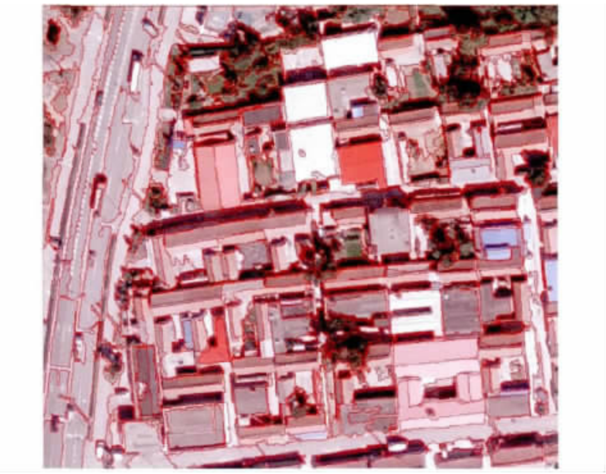
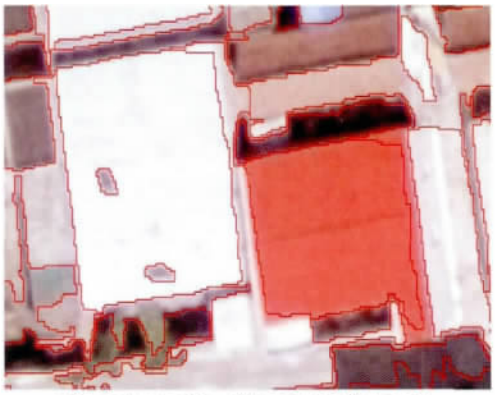
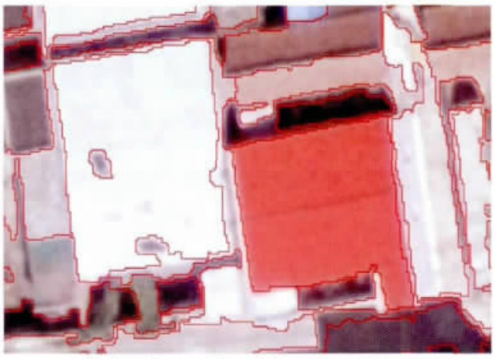


Fig. 5 The segmentation result generated in eCognition

Visual analysis shows that the location layout of objects is similar in both segmentation results, and larger geographical objects are also well segregated in both results, such as the residential buildings, roads, trees, etc. However, some localized differences lie in the small objects between these larger objects and the transitional edge pixels between them. To illustrate these differences, both segmentation results are expanded in Fig. 6. In the segmentation result generated with the proposed method (referred to as "Method 1" or "Result 1"), two gray colored spots in a larger white colored roof of a residential building are well segmented as two small objects, but in the result obtained with eCognition (referred to as "Method 2" or "Result 2") the lower spot is missed. Furthermore, the boundaries of the roof are more regular-shaped and agreeing with the edges of the roof in Result 1. The nearby larger red colored roof is more regular-shaped in Result 2, with neater and more concise edge pixels. However, in both segmentation results, narrow and long regions with colors lighter than the larger center region show up along the boundaries of the roof profile, and they are segmented out of the roof object. This tendency is more apparent in Result 1. This phenomenon is probably caused by the region-growing method adopted for merging regions in both segmentation methods. In the region-growing method, adjacent pixels (regions) are not considered in pixel (region) merging, so it is hard to generate consistent boundaries for the profiles of geographic objects. Therefore, the region-growing method needs to be improved in both segmentation methods.



(a) Local expansion of the segmentation result generated with the proposed new method



(b) Local expansion of the segmentation resualt genrated with the eCognition software

Fig. 6 Local expansion of the segmentation result generated with the proposed new method and eCognition

For further assessing the accuracy of the proposed segmentation method, statistical analysis results of the shape characteristics of the regions or objects in both segmentation results, including compactness and smoothness, are shown in Table 2, with the cells with white colored background representing Method 1, and those with gray colored background representing Method 2.

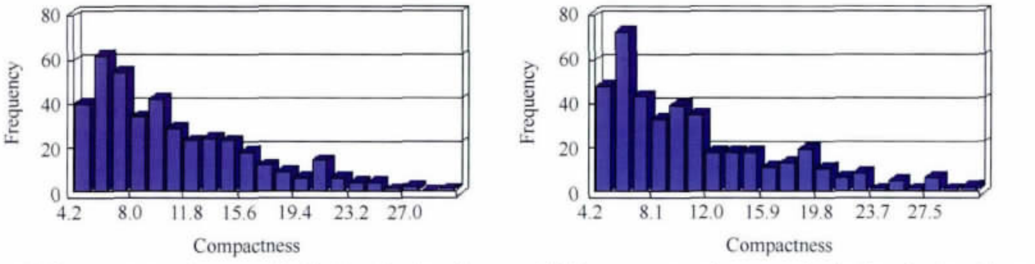
Table 2 Statistics of the characteristics of the segmented regions generated with the proposed new method and eCognition

Parameter	Minimum	Maximum	Average	Standard deviation
Compactness	4.179107	30.786345	11.028516	5.333127
	4.215839	31.384937	11.157507	5.897363
Smoothness	0.923593	2.546151	1.360699	0.312535
	0.909091	2.982122	1.394478	0.376648

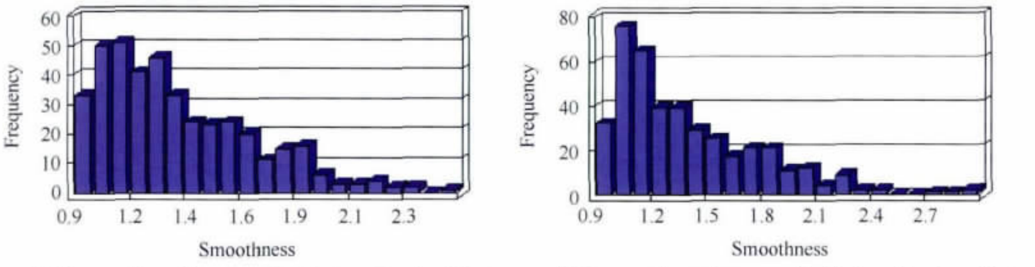
Note: Cells with white colored backgroud represent Method 1, Cells with gray colored background represent Method 2.

The frequency distributions of both characteristics are shown in Fig. 7, with Result 1 on the left and Result 2 on the right.

Comparison of the statistics and frequency distribution of each characteristics belonging to both methods shows that (1) the mean and standard deviation of the compactness of the region polygons in Result 1 are both smaller than Result 2, suggesting that Result 1 has better compactness than those in Result 2, and with smaller individual differences or closer to the mean value; (2) the mean and standard deviation of the smoothness of the region polygons in Result 1 are both smaller than those in Result 2, suggesting that Result 1 has better smoothness than Result 2, and with smaller individual differences or closer to the mean value.



(a) The compactness frequency distribution of region objects (b) The compactness frequency distribution of region objects



(c) The smoothness frequency distribution of region objects (d) The smoothness frequency distribution of region objects

Fig. 7 The frequency distributions of compactness and smoothness of the segmentation result generated with the proposed new method and eCognition

In sum, when segmenting an RS image with the proposed new method and eCognition, if the number of image-objects in the segmentation result is kept constant, both methods can segment different geographic objects well in general by adjusting the segmentation parameters. The location layouts of objects are generally similar in both segmentation results, with differences only in localized details, but these differences cannot be strictly graded with respect to particular geographic objects. Quantitative analysis shows that the compactness and smoothness of the region polygons in the segmentation result generated with the proposed new method are better, suggesting that the boundaries of geographic objects are located more accurately, and the boundary profiles of image-objects are relatively smoother and more concise.

5 CONCLUSIONS

A new automated HS method with introduced relaxation factors is proposed for processing HRSI. Based on the region adjacency graph and optimized region merging principle, a relaxation factor is used to adjust the heterogeneity between the regions to be merged, and the speed of the entire segmentation is improved by controlling the number of objects to be merged in each recursive merging. With the mean of the heterogeneity between regions taken as the cardinality, another relaxation factor is used to control the scale parameter of the levels to be exported, automatically generating multi-scale HS results.

Experimental results show that this new segmentation method provides results of good quality, and can satisfy the accuracy required for further image analysis and geographical object extraction. Comparison with the segmentation result generated in eCognition further demonstrates this point. However, further investigation is still required on processing the boundaries of geographic objects with complex image contents and reducing the occurrence of regions of narrow and long shape.

REFERENCES

Baatz M, Arini N, Schäpe A, Binnig G and Linssen B. 2006. Object-oriented image analysis for high content screening: detailed quantification of cells and sub cellular structures with the cellenger software. Cytometry Part A, 69(7): 652 - 658

Baatz M. and Schäpe A. 1999. Object-oriented and multi-scale image analysis in semantic networks [A]. In: Proc of the 2nd International Symposium on Operationalization of Remote Sensing [C]. Enschede ITC

Beaulieu J M and Goldberg M. 1989. Hierarchy in picture segmentation: a stepwise optimization approach. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11(2): 150 - 163 [DOI: 10.1109/34.16711]

Benz U C, Hofmann P, Willhauck G, Lingenfelder I and Heynen M. 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for gis-ready information. ISPRS Journal of Photogrammetry and Remote Sensing, 58(3/4): 239 - 258 [DOI: 10.1016/j.isprsjprs.2003.10.002]

Blaschke T. 2010. Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing, 65(1): 2 - 16 [DOI: 10.1016/j.isprsjprs.2009.06.004]

Blaschke T, Lang S and Hay G. 2008. Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications. New York: Springer

Cardelino J, Caselles V, Bertalmío M and Randall G. 2009. A contrario hierarchical image segmentation. In: Proceedings of the 2009 16th IEEE International Conference on Image Processing (ICIP). Cairo: IEEE: 4041 - 4044 [DOI: 10.1109/ICIP.2009.5413723]

Dezsö B, Fekete I, Gera D, Giachetta R and László I. 2012a. Object-based image analysis in remote sensing applications using various segmentation techniques. In: Annales Univ. Sci. Budapest, Sect. Comp, 37: 103 - 120

Dezsö B, Giachetta R, László I and Fekete I. 2012b. Experimental study on graph-based image segmentation methods in the classification of satellite images. EARSeL eProceedings, 11(1): 12 - 24

Haralick R M and Shapiro L G. 1985. Image segmentation techniques. Computer Vision, Graphics, and Image Processing, 29(1): 100 - 132

Hay G J and Castilla G. 2006 OBIA. Object-based image analysis: strengths, weaknesses, opportunities and threats (Swot). The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences

Korting T S, Castejon E F and Fonseca L M G. 2011. Divide and Segment-an Alternative for Parallel Segmentation. In: Procssings XII GEOINFO, November 27 - 29, 2011, Campos do Jordao, Brazil. p 97 - 104

Li D R, Tong Q X, Li R X, Gong J Y and Zhang L P. 2012. Current issues in high-resolution Earth observation technology. Science China Earth Science, 42(6): 805 - 813 [DOI: 10.1007/s11430-012-4445-9]

Lin H, Liu P, Du P J and Xia J G. 2012. Research on multi-scale segmentation algorithm of high-resolution remote sensing imagery based on improved statistical region merging. Computer Engineering and Applications, 48(18): 159 - 163

Lv Z Y, Zhang X L, Gao L P and Jin J X. 2012. The study and application analysis of FNEA segmentation algorithm based on high resolution remote sensing image data. Geomatics & Spatial Information Technology, 35(10): 13 - 16

Liu J H and Mao Z Y. 2009. A Survey on high spatial resolution remotely sensed imagery segmentation techniques and application strategy. 0 (6): 95 - 101

Marchand-Maillet S and Sharaiha Y M. 1999. Binary Digital Image Processing: A Discrete Approach. San Diego: Academic Press

Marr D. 1982. Vision: A Computational Approach. San Francisco: Freeman & Co.

Murthy Y S S R, Krishna I V M and Gupta S. 2012. Object oriented image analysis for feature extraction from high resolution multispectral imagery. Indian Journal of Electronic and Electrical Engineering, 1 (1/2): 1 - 9

Plaza A J and Tilton J C. 2005. Automated selection of results in hierarchical segmentations of remotely sensed hyperspectral images. In: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. [s. n.]: IEEE: 4946 - 4949 [DOI: 10.1109/IGARSS.2005.1526784]

Sarkar A, Biswas M K and Sharma K M S. 2000. A simple unsupervised Mrf model based image segmentation approach. IEEE Transactions on Image Processing, 9(5): 801 - 812 [DOI: 10.1109/83.841527]

Sun X, Fu K and Wang H Q. 2011. High resolution remote sensing image understanding. Beijing: Science press

Trias-Sanz R, Stamon G and Louchet J. 2008. Using colour, texture, and hierarchial segmentation for high-resolution remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing, 63(2): 156 - 168 [DOI: 10.1016/j.isprsjprs.2007.08.005]

Trimble. 2012. Ecognition Developer 8. 7. 2 Reference Book. Trimble Germany GmbH, Trappentreustr. 1, D-80339 Munchen, Germany

Tupin F and Roux M. 2005. Markov random field on region adjacency graph for the fusion of sar and optical data in radargrammetric applications. IEEE Transactions on Geoscience and Remote Sensing, 43 (8): 1920 - 1928 [DOI: 10.1109/TGRS.2005.852080]

Tzotsos A and Argialas D. 2006. Mseg: a generic region-based multiscale image segmentation algorithm for remote sensing imagery. In: Proceedings of ASPRS 2006 Annual Conference. Reno, Nevada: ASPRS: 1 - 13

Wang A P, Wang S G and Wu H Z. 2009. Multiscale Segmentation of High Resolution Satellite Imagery by Hierarchical Aggregation. Geomatics and Information Science of Wuhan University, 34(9): 1055 - 1058

Wang Z and Bovik A C. 2009. Mean squared error: love it or leave it? A new look at signal fidelity measures. IEEE Signal Processing Magazine, 26(1): 98 - 117 [DOI: 10.1109/MSP.2008.930649]

Xia G S, He C and Sun H. 2006. An unsupervised segmentation method using markov random field on region adjacency graph for SAR images. In: Proceedings of CIE'06 International Conference on Radar. Shanghai, China: IEEE: 1 - 4 [DOI: 10.1109/ICR.2006.343148]

Zhang H, Fritts J E and Goldman S A. 2008. Image segmentation evaluation: a survey of unsupervised methods. Computer Vision and Image Understanding, 110(2): 260 - 280 [DOI: 10.1016/j.cviu.2007.08.003]

Zhang Y J. 2012. Image engineering (Ⅱ)-image analysis. Beijing: Tsinghua University press

引入松弛因子的高分辨率遥感影像自动多层次分割

邓富亮，唐娉，刘源，杨崇俊

遥感科学国家重点实验室 中国科学院遥感与数字地球研究所，北京 100101

摘　要：针对当前高分辨率遥感影像多层次分割尺度参数设置缺少理论框架支持、人为因素影响较多等缺点，提出一种引入松弛因子的高分辨率遥感影像自动多层次分割方法。该方法利用1个松弛因子调节引导区域对象合并的异质性值大小，通过控制每次递归合并区域的对象个数，提高了整体分割的速度；以区域对象间异质性平均值作为基数，引入另一个松弛因子控制分割过程中层次输出的尺度参数，使整个分割过程自动得到不同尺度的多层次分割结果。实验结果表明，该方法具有较高的分割质量，能够满足遥感影像分析及地物提取的精度要求，并且减少了人为因素影响，提高了自动化程度。但是，对于复杂图像内容的地物目标边界处理和减少狭长区域对象的出现还需要进一步深入研究和实践。

关键词：松弛因子，高分辨率遥感影像，自动多层次分割，异质性测度，区域邻接图

中图分类号：TP751.1　　**文献标志码**：A

引用格式：邓富亮，唐娉，刘源，杨崇俊. 2013. 引入松弛因子的高分辨率遥感影像自动多层次分割. 遥感学报，17(6)：1492－1507

Deng F L，Tang P，Liu Y and Yang C J. 2013. Automated hierarchical segmentation of high-resolution remote sensing imagery with introduced relaxation factors. Journal of Remote Sensing，17(6)：1492－1507 [DOI：10.11834/jrs.20133031]

1 引　言

高分辨率遥感影像由于具有地物的几何结构更加明显、位置布局更加清晰、纹理和尺寸更加精细，地物目标多尺度化等特征，使传统的基于像素的中低分辨率遥感影像处理和分析体系陷入困境难以有效地解译高分辨率遥感影像(李德仁 等,2012;孙显等,2011;刘建华 等,2009)。Blaschke等人(2008)发出了“像素出了什么问题”的疑问，指出传统的基于像素的遥感影像处理和分析方法很难满足当前高分辨率遥感影像的应用需求。Baatz等人(1999)也指出重要的语义解释更需要用有意义的影像中对象及对象之间的相互关系而不是用一个个像素来表示。由此，面向对象的影像分析技术应运而生。基于对象的分析方法是一种能够融合影像的光谱信息、形状信息、纹理信息、空间拓扑关系等的研究方法(Blaschke 等,2010;Korting 等,2011)，其主要思路是首先将影像进行分割，然后以分割的影像块作为基本处理单元进行分类识别与地物提取(Blaschke 等,2010)。作为对象获取的主要方法，影像分割已经成为面向对象的影像分析的关键步骤(Baatz 等,2006;Dezsö 等,2012a;Murthy 等,2012)。

众多学者对高分辨遥感影像的分割方法进行了广泛而深入地研究，其中多层次分割方法是当前研究的热点。著名的商业软件eCognition提供的多尺度分割是一种典型的多层次图像分割方法，但它存在尺度参数设置缺少理论框架支持、人工干预程度高等缺点，并且它的多层次分割结果不是一次分割形成，而是通过人为设定多个尺度参数来确定分割结果的输出层次，从而构建多层次的分割结果(林卉 等,2012;王爱萍 等,2009)。

2 多层次分割方法及其过程分析

高分辨率遥感影像中丰富的地物目标与空间语

收稿日期：2013-02-22；**修订日期**：2013-06-29；**优先数字出版日期**：2013-07-06

基金项目：高分辨率对地观测系统重大专项(编号：02-Y30A04-9001-12/13)

第一作者简介：邓富亮(1982—)，男，博士研究生，现从事遥感图像处理和应用研究。E-mail：fldeng8266@gmail.com

通信作者简介：唐娉(1968—)，女，研究员，博士生导师，主要从事遥感图像处理研究。E-mail：tangping@irsa.ac.cn

义信息必须在多个尺度下才能充分表达和描述（Marr，1982；Trias-Sanz 等，2008）。多层次分割是构建遥感影像中不同种类地物多尺度表达常用的方法。本文中的多层次是指经由一幅影像分割生成多个尺度不同层次细节的分割结果集，这些分割结果集会形成对同一幅影像、不同尺度的多层次网络结构表达，如图 1 所示。其中一层分割结果称为区域对象层，一个区域对象层是由多个区域对象组成，多个区域对象层形成区域对象层组。在区域对象层组中，通过对同一区域对象层中区域对象间构建拓扑邻接关系，上下区域对象层中区域对象间构建层次继承关系，则形成区域对象的多层次网络结构。在区域对象的多层次网络结构中，每一层区域对象层对应一个分割尺度，多层次分割的目标既是生成对一幅遥感影像的不同尺度的多层次网络结构表达。

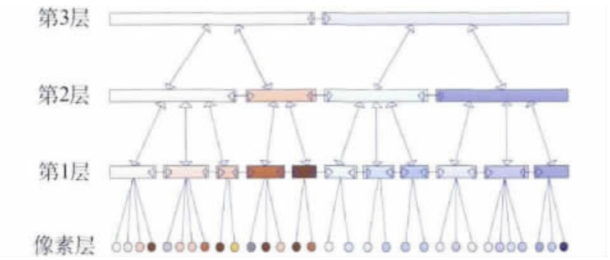


图 1　遥感影像的不同尺度多层次网络结构表达

从图 1 可以看出，多层次分割结果中同层区域对象之间具有邻接拓扑关系，上下层的区域对象之间具有隶属关系。这种区域对象的多层次表达可以很好地反映地物的上下文信息，相比单一分割结果的表达具有非常明显的优势，是充分挖掘高分辨率遥感影像应用潜力的一种有效途径（Plaza 和 Tilton，2005）。

目前最成功的商业软件 eCognition 提供的多尺度分割是应用最为广泛的多层次分割方法。它通过设定不同的尺度参数，进行多次分割，形成网络层次结构，每一层分割结果对应一个分割解决方案，层次越往上对应尺度越粗，相反，层次越往下，则尺度越细（Trimble，2012）。其基于异质性最小区域合并原则，采用自底向上进行区域合并，从而完成区域对象的提取。这种多层次分割结果不是一次分割形成，而是通过人为设定多个尺度参数来确定分割结果的输出层次。一般地，在 eCognition 软件中可以通过不断尝试不同尺度参数对影像进行多次分割，然后比较多个分割结果，选择其中的一层或多层应用于具体影像分析或地物提取。实践发现，这种多层次分割方法存在尺度参数难以设置，缺少理论框架支持，人为因素影响较大，自动化程度低等缺点，很难达到完全实用的目的。

eCognition 软件提供的多尺度分割方法主要包括 3 个环节：首先，要选择区域对象合并的异质性测度方法；其次，制定区域合并过程中区域对象的合并策略；最后，确定层次输出规则以及合并终止条件。

3　引入松弛因子的多层次分割方法

针对多层次分割方法 3 个主要环节，提出一种引入松弛因子的高分辨率遥感影像自动多层次分割方法，主要包括以下步骤：提出一种新的区域合并异质性测度方法；通过利用 1 个松弛因子调节引导区域对象合并的异质性值的大小；以区域对象间异质性平均值作为基数，引入另一个松弛因子控制区域合并过程中层次输出的尺度，使通过一次分割自动生成具有不同尺度的多层次分割结果集。

3.1　区域合并异质性测度

区域合并异质性测度方法对分割结果的优劣有至关重要的影响，其特征选取不仅依赖于具体问题本身，也和所用遥感影像数据的种类有关（章毓晋，2012）。而优质的分割一般需满足以下几个条件（Haralick 和 Shapiro，1985）：（1）区域内具有特定测度的同质性；（2）相邻区域间具有足够的异质性；（3）区域边界简洁、平滑、准确。前两个条件称为特征准则，第 3 个条件称为语义准则，即符合人类知觉原理（Zhang 等，2008）。因此，本文从光谱特征和形状特征两方面选择区域合并异质性测度准则。

光谱异质性是用来表示区域对象内部各像素之间的光谱差异性。其表示方法有很多种，而均方误差及其改进方法 50 多年来一直占据图像处理领域定量性能指标和评价标准的主导地位（Wang 等，2009），本文选择均方误差作为区域合并的光谱异质性测度准则之一

$$h_{\mathrm{MSE}} = \left[\frac{n_i n_j}{n_i + n_j} \sum_{b=1}^{B} (\mu_{ib} - \mu_{jb})^2 \right]^{1/2} \qquad (1)$$

式中，n_i 和 n_j 分表代表区域对象 i 和 j 的像素数；B 为遥感影像波段数；μ_{ib} 和 μ_{jb} 分表代表区域对象 i 和 j 第 b 波段的光谱均值。

光谱异质性测度准则的另外一个特征是规范化最大标准差

$$\sigma_i = \max\{\sigma_{ib} : b = 1,2,\cdots,B\} \tag{2}$$

式中,σ_{ib}代表区域对象 i 第 b 波段的规范化光谱信息标准差

$$\sigma_{ib} = \frac{1}{\mu_{ib}}\sqrt{\frac{1}{n_i - 1}\left[\sum_{x_p \in X_i}(x_{pb})^2 - n_i(\mu_{ib})^2\right]} \tag{3}$$

式中,X_i 为第 i 个区域对象;x_p 代表区域对象 i 中第 p 个像素;x_{pb}代表像素 x_p 的光谱值。当区域对象只包含一个像素时,σ_{ib}的值为 0.0。

联合式(1)和式(3),给出两个区域合并的光谱异质性测度准则

$$h_{\text{spectral}} = h_{\text{MSE}}\left[1.0 + w_{\text{stddev}}\frac{|\sigma_i - \sigma_j|}{\sigma_i + \sigma_j}\right] \tag{4}$$

式中,w_{stddev}代表规范化最大标准差的权重,取值范围为 [0,∞),默认值为 1.0。

式(4)的取值越大,表示两个区域的光谱信息越相似,合并的可能性越大,反之,两个区域的光谱信息差异越大,合并的可能性越小。规范化光谱信息标准差主要用来增加两个区域对象光谱相似性判断的筹码,通过设置合适的 w_{stddev} 权重,可使分割结果的区域对象内光谱异质性更好。

形状异质性是用来表示区域对象形状的差异性。为使分割出的区域对象边界简洁、光滑、准确,本文将形状紧致性融入异质性计算准则。紧致性是一个重要的形状性质,有许多描述方式,这些描述基本上都对应目标的几何参数,所以均与尺度有关(章毓晋,2012)。目标区域的紧致性可以通过与正方形或圆等理想形状的区域进行比较来间接描述(Marchand-Maillet 和 Sharaiha,1999)。本文选择外观比 AR 作为区域合并的形状异质性测度准则之一

$$AR = l/h \tag{5}$$

式中,l 和 h 分别表示目标最小包围盒的长和宽。对正方形或圆形目标,外观比值最小,为 1,对于狭长目标,外观比的值大于 1,并且随目标的狭长程度而增加。

本文选择的另外一个形状异质性测度准则是圆形指数 CI,用来衡量区域对象轮廓边界的光滑程度即

$$CI = a_{\text{c}}/a_{\text{p}} \tag{6}$$

式中,a_{p} 为目标的面积,a_{c} 为与目标周长相等的圆的面积。对圆形目标,圆形指数最小,为 1,表示最光滑,对于其他形状其值大于 1,并且随目标边界的不规则程度增加。

联合式(5)(6),给出两个区域对象合并的形状异质性测度准则

$$h_{\text{shape}} = w_{\text{compact}}\left|\frac{l_{ij}}{h_{ij}} - \left(\frac{l_i}{h_i} + \frac{l_j}{h_j}\right)\right| + (1 - w_{\text{compact}})\left|\frac{a_{\text{c}ij}}{a_{\text{p}ij}} - \left(\frac{a_{\text{c}i}}{a_{\text{p}i}} + \frac{a_{\text{c}j}}{a_{\text{p}j}}\right)\right| \tag{7}$$

式中,w_{compact}表示紧凑度的权重;$1 - w_{\text{compact}}$表示光滑度的权重;其中下标 ij 表示区域对象 i 和 j 合并后的新区域对象。

区域合并总体异质性测度主要考虑光谱特征和形状特征,故包括光谱异质性测度和形状异质性测度两个方面即

$$h_{\text{merge}} = h_{\text{spectral}} + w_{\text{shape}}h_{\text{shape}} \tag{8}$$

式中,w_{shape}表示形状异质性权重,取值范围为 [0,∞),默认值为 1.0。h_{merge}取值越小表示两个区域对象的相似性越好,合并的可能性越大,反之,表示两个区域对象的差异性越大,合并的可能性越小。

由于光谱信息是遥感影像数据中最重要的数据,通常情况下光谱异质性最重要,其对分割结果的优劣有至关重要的影像。引入形状异质性对分割结果具有指示作用,使分割结果的区域对象的边界更加紧凑和光滑,有助于避免区域对象形状的不完整,但是,在实际使用中其权重不宜设置过大(吕志勇 等,2012)。

3.2 引入松弛因子的区域邻接图最优合并分割

多层次分割通常采用区域生长法构建多尺度不同层次细节的区域对象(Cardelino 等,2009),区域生长的合并策略主要有基于局部(Benz 等,2004;Sarkar 等,2000)和面向全局(Beaulieu 和 Goldberg,1989)两类。前者在局部范围内搜索满足合并准则的相邻区域对象对进行合并。后者为层次合并,每次搜索全局最优相邻区域对象对进行合并,能及时针对每次合并产生的变化作出调整,精度较高,稳定性较好(Dezsö 等,2012b)。本文选择面向全局的区域生长策略,利用区域邻接图来表达区域对象间的空间关系(Tupin 和 Roux,2005;Xia 等,2006)。

区域邻接图的定义为:$G = (V,E)$,其中 V 为节点的集合,V_i 为第 i 个节点,每个节点代表一个区域对象,E 为弧段的集合,E_{ij}为第 i 个节点和第 j 个节点间的弧段,弧段的权值为两相邻区域对象间的异质性测度值。由此,面向全局的区域生长策略即转化为每次搜索区域邻接图中权值最小的弧段,然后将该弧段的两个相邻区域对象进行合并,同时调整局部拓扑关系及弧段权值,迭代至满足合并终止条件则分割过程结束。这种全局最优合并策略理论严密,精度较高,但是效率相对较低,尤其是处理大幅

面遥感影像时很难满足实际需求。为此，本文提出在合并过程中区域对象间异质性值大小控制参数

$$w_h = \lambda w_{\min} \tag{9}$$

式中，λ 为松弛因子，用于控制每次迭代区域合并的异质性值；$w_{\min}$ 表示区域邻接图中异质性最小的弧段权值。可以通过调节 λ 的值控制每次递归合并的区域对象个数，即弧段权值满足 $w_{ij} \leqslant w_h$ 的相邻区域对象都将被合并。可见，λ 值越小每次递归合并的区域对象个数越少，当为 1.0 时退化为每次合并异质性最小的相邻区域对象。但是，λ 值较大时会影响最终区域合并的精度，不宜设置过大。

多层次分割结果中每一层次对应一个尺度，不同尺度具有不同精度，适合不同种类的地物提取及目标识别。那么，合并过程中记录哪些尺度的中间状态是一个要解决的问题，一般多尺度分割是通过多次设定尺度来实现的（Tzotsos 和 Argialas，2006）。本文提出采用区域间异质性值的平均值作为基数，引入另一个松弛因子来自动控制层次输出尺度即

$$s_1 = \delta\left(\frac{1}{n}\sum w_{ij}\right)^{1/2} (i \neq j) \tag{10}$$

式中，δ 为松弛因子，用于控制层次输出个数；n 表示区域邻接图中弧段个数；w_{ij} 表示区域对象 i 和 j 间的异质性值，即弧段权值。本文自动多层次分割，需要预先设置起始层次输出尺度 S_{start} 和结束层次输出尺度 S_{stop}，这里的尺度既是指所有区域间异质性值的平均值的平方根，对平均值进行开方主要是为了增加尺度控制的敏感性。在区域合并过程中，将根据 δ 值输出尺度在 S_{start} 和 S_{stop} 范围内的多个分割结果，形成最终的多层次分割结果集。在实际应用中，δ 的值可以根据具体应用及特定数据设置，值越大输出的层次个数越少，反之越多。

3.3 引入松弛因子的多层次分割过程

在给出多层次分割方法具体过程前，先给出其主要分割参数项，如表 1 所示。

表 1 多层次分割方法主要参数列表

参数名称	参数意义
S_{start}	起始层次输出尺度
S_{stop}	结束层次输出尺度
w_{stddev}	规范化标准差权重
w_{shape}	形状异质性权重
w_{compact}	紧凑度权重
λ	区域合并控制松弛因子
δ	层次输出控制松弛因子

多层次分割过程如图 2 所示，具体描述如下：

步骤 1 输入原始遥感影像，设置分割参数，主要包括表 1 参数列表项。

步骤 2 基于某种方法生成初始区域对象或每一个像素作为一个初始区域对象。

步骤 3 将初始区域对象转化为区域邻接图，采用区域合并异质性测度计算相邻区域间异质性值作为弧段权值。

步骤 4 合并异质性值满足合并条件的相邻区域对象。假设当上一层次输出尺度为 S_{prev}，该步骤的执行过程如下：计算当前分割状态的尺度，记为 S。如果 S 大于 $\delta \times S_{\text{prev}}$，执行步骤 5；否则，获得区域邻接图中最小的边权值 $w_{\min}$，根据式(9)计算区域合并条件 w_h。将区域邻接图中边权值小于等于 w_h 的边记录至一个临时数组，并按照权重大小升序排列，如果权重相等则按照边相连的区域对象中包含的较少像素个数的大小升序排列，如果像素个数再相等则按照区域编号大小升序排列。依次合并临时数组中边连接的相邻区域对象。重复上述过程，直至满足层次输出尺度条件。

步骤 5 输出当前尺度的分割结果，记录当前分割状态的尺度，记为 S_{cur}，将 S_{cur} 赋给 S_{prev}，作为下一层次输出尺度的参照。如果 S_{cur} 大于 S_{stop} 跳转到步骤 6；否则回到步骤 4。

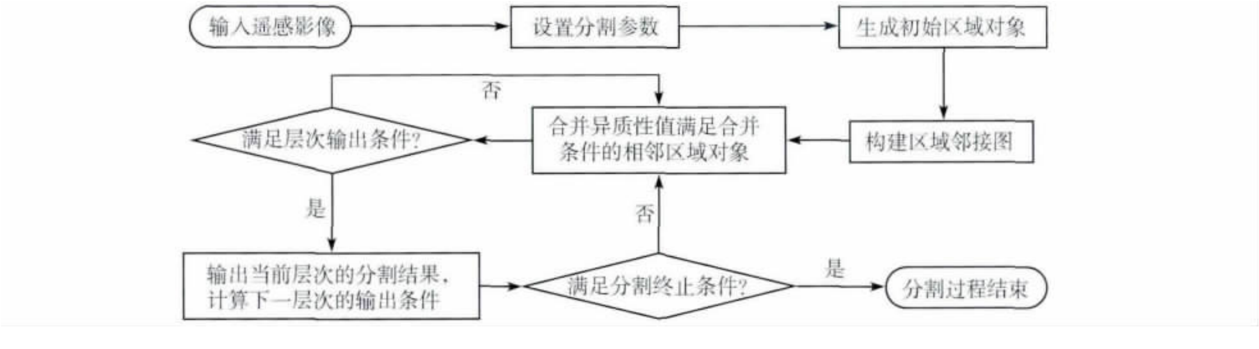


图 2 多层次分割过程

步骤 6 分割过程结束，为便于后续基于多层次分割结果进行图像分析、地物提取等深化应用，从细粒度到粗粒度依次构建上下层区域对象之间的隶属关系及层内区域对象的拓扑邻接关系，形成对同一幅遥感影像的多层次网络结构表达。

综上所述上述过程主要包括两个关键环节：第 1 个环节是区域合并，即根据每次迭代统计出满足合并条件的相邻区域对象，然后依据相应规则依次进行区域合并，直至满足层次输出尺度条件，本轮区域合并停止。第 2 个环节是层次输出控制，根据本文多层次分割尺度计算方法及控制方式，在区域合并过程中输出满足层次输出尺度的中间结果，然后计算下一个层次输出尺度。重复以上两个环节，直到满足分割终止条件，即得到不同尺度的多层次分割结果。

4 实验结果与分析

依据本文提出的多层次分割方法，选取一幅高分辨率遥感影像进行多层次分割实验，证明其有效性。实验数据为 2011 年 7 月拍摄的北京市航拍数据，裁切的影像大小为 913 像素 × 811 像素，空间分辨率为 0.2 m，包含可见光 3 个波段，原始遥感影像如图 3 所示。

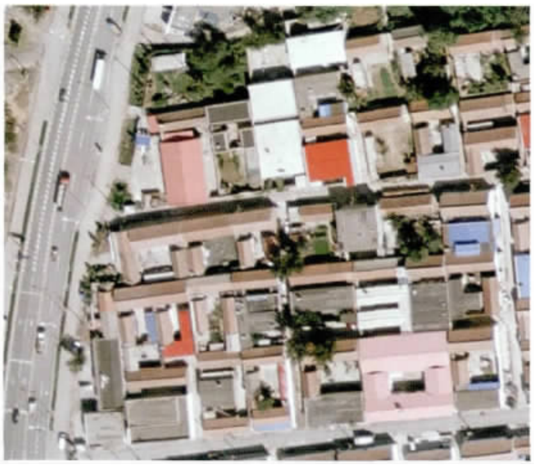
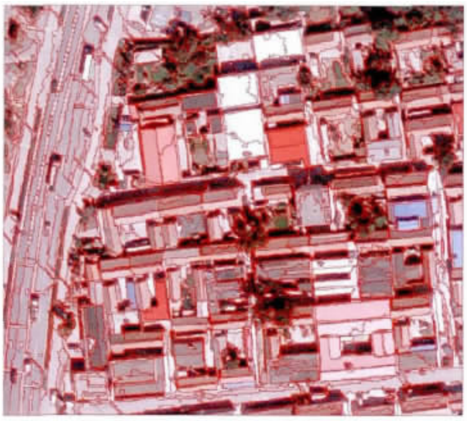


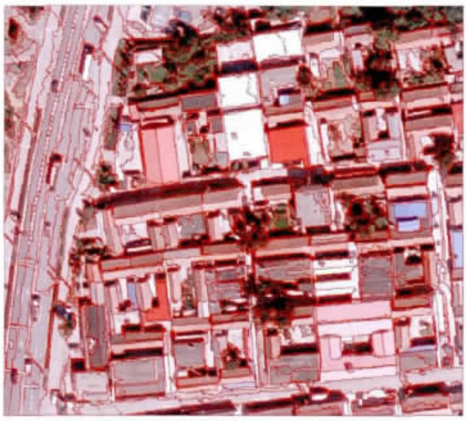
图 3 原始遥感影像

实验设置的各个参数分别为：$S_{\text{start}}=200.0$，$S_{\text{stop}}=382.2$，$w_{\text{stddev}}=0.1$，$w_{\text{shape}}=40.0$，$w_{\text{compact}}=0.5$，$\lambda=1.01$，$\delta=1.3$。实验中参数设置较多，但是实验表明这对于不同的高分辨率遥感影像具有良好的分割效果。在实际使用中，一般地，区域合并控制松弛因子 λ 设置为 1.01，能够满足分割结果精度要求。

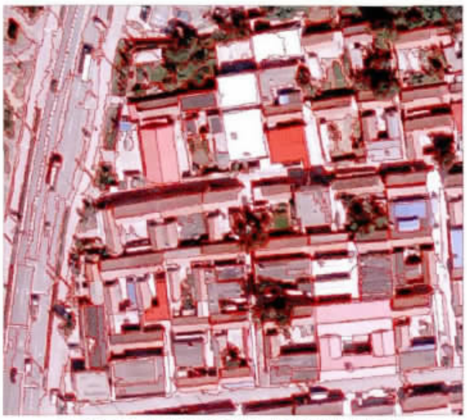
图 4 为原始遥感影像的多层次分割结果，包含 4 个层次的不同尺度分割结果。



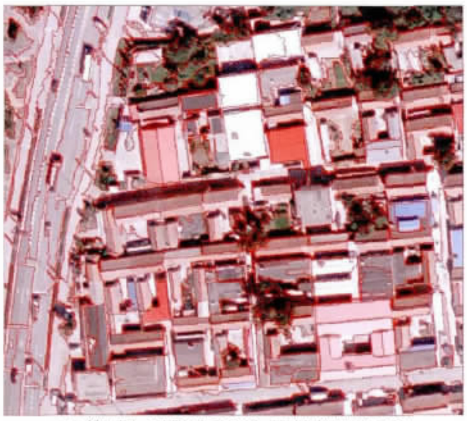
(a) 第1层：尺度为200.0，区域数目为1166



(b) 第2层：尺度为260.0，区域数目为766



(c) 第3层：尺度为338.2，区域数目为500



(d) 第4层：尺度为382.2，区域数目为408

图 4 多层次分割结果

从图4中可以看出地物由细到粗的分割过程，如道路上的白色行车线（虚线）、居民楼顶的小斑块等。从第1层至第4层，逐层进行区域合并，尺度越来越大，区域对象数量则越来越少，小区域对象逐渐被合并至大区域对象。如到第3层时，很多居民楼顶都被合并成一个相对完整的区域对象，使楼顶的区域对象的轮廓边界更加与实际边缘吻合。而到第4层时，道路附属设施的阴影、路面小版块等一些小区域对象被合并至大的区域对象中，此时的分割结果则显得更加整洁。实验结果表明，本文多层次分割方法适合对高分辨率遥感影像进行多尺度分割，形成由细到粗的多层次分割结果集。

接下来结合目视分析和定量分析对比本文分割结果与eCognition（8.7版本）软件的分割结果。比较的基础为分割结果的区域数目相同，选择上面实验第4层的分割结果作为对比，其区域数目为408，通过多次尝试不同尺度参数，使eCognition软件的分割结果区域数目也为408（尺度为90.0，形状权重为0.2，紧凑度权重为0.5），结果如图5所示。

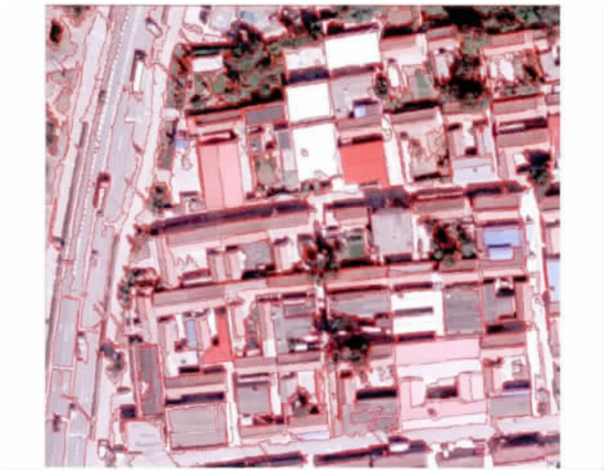
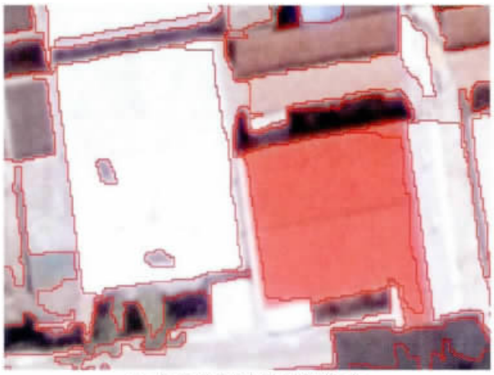
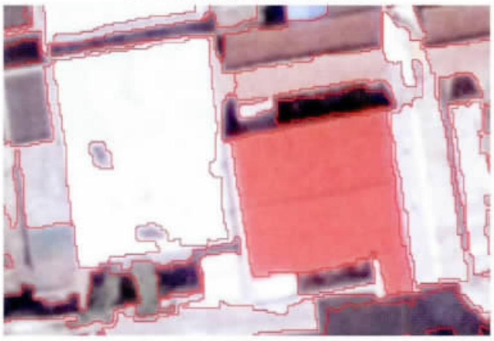


图5　eCognition软件分割结果

通过目视分析，对比本文与eCognition软件的分割结果，整体上二者均能较好地分割不同地物对象，分割结果的区域格局相近，居民楼、道路、树木等稍微大的地物目标基本上都分割出来了。它们的不同在于，在这些稍大地物目标之间的小地物目标以及它们中间的边缘像素过渡位置存在局部差异。通过放大显示，进一步了解二者的这些差异。如图6所示，较大的白色居民楼顶的两个灰色小斑块，本文很好地分割出两个小区域对象，而eCognition软件没有将下侧的小斑块分割为一个单独的区域对象，并且针对整个居民楼顶的边界轮廓，本文的分割结果较为规整，边缘比较吻合。而另外一个较大的红色居民楼顶则分割差异稍大，eCognition软件的分割结果较为规整，边缘像素比较干净整洁。但是，居民楼顶轮廓的边界处均有狭长的区域对象出现，这些狭长区域对象的颜色比中间大区域对象的颜色浅，因此被独立分割出来，本文分割结果尤为明显。这主要是因为二者均采用区域生长法进行区域合并，这种方法在像素（区域）合并时不考虑周围像素（区域），难以找到一致性的地物目标轮廓边界。这也是区域合并法一个有待改进的缺点。



(a) 本文分割结果局部放大



(b) eCognition分割结果局部放大

图6　本文与eCognition分割结果的局部放大对比

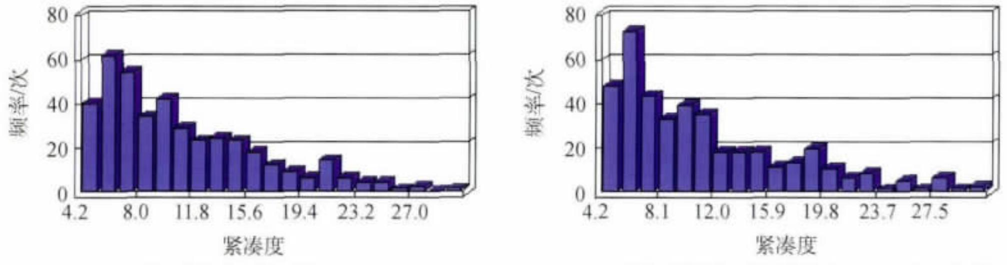
为进一步验证本文多层次分割的精度，下面对二者分割结果区域对象进行形状特征统计及分析，统计它们的紧凑度和光滑度两个形状特征，光滑度和紧凑度数值越小，说明分割质量越好。表2为各特征的统计信息。

表2　本文与eCognition分割结果区域各特征的统计信息

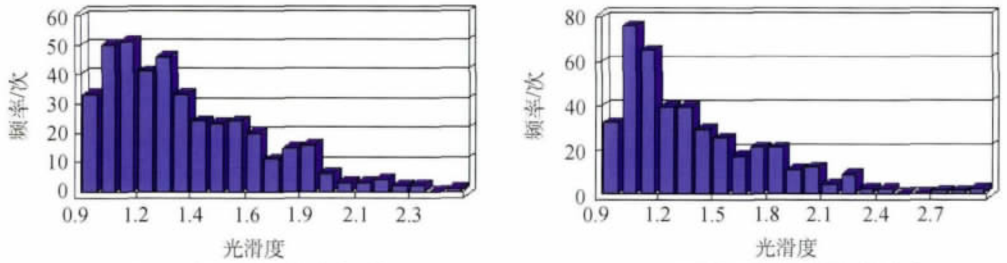
参数	最小值	最大值	平均值	标准差
紧凑度	4.179107	30.786345	11.028516	5.333127
	4.215839	31.384937	11.157507	5.897363
光滑度	0.923593	2.546151	1.360699	0.312535
	0.909091	2.982122	1.394478	0.376648

注：白色底为本文方法分割结果，灰色底为eCognition软件分割结果。

图7为各特征的频率分布图,图7(a)(c)为本文分割结果的统计结果,图7(b)(d)为eCognition软件分割结果的统计结果。



(a) 区域对象紧凑度频率分布(本文方法)　(b) 区域对象紧凑度频率分布(eCognition结果)



(c) 区域对象光滑度频率分布　(d) 区域对象光滑度频率分布

图7　本文与eCognition分割结果区域各特征的频率分布

结合各特征统计信息和频率分布,对比分析如下:(1)本文分割结果区域对象多边形的紧凑度的平均值和标准差均比eCognition软件分割结果区域对象多边形的紧凑度的平均值和标准差小。这两个统计参数均说明本文分割结果区域对象多边形的紧凑度比eCognition软件分割结果区域多边形的紧凑度好,并且差异性较小,比较接近平均值。(2)本文分割结果区域对象多边形的光滑度的平均值和标准差均比eCognition软件分割结果区域多边形的光滑度的平均值和标准差小。这两个统计参数均说明本文分割结果区域对象多边形的光滑度比eCognition软件分割结果区域多边形的光滑度好,并且差异性较小,比较接近平均值。

综上所述,在对一幅遥感影像进行分割时,在分割结果区域数目相同的情况下,通过调试各自的分割参数,整体上二者均能较好地分割不同地物目标,分割结果的区域格局整体相近,局部细节部分则有所差异,但是针对特定地物目标没有严格的好坏之分。而通过定量分析发现,本文分割结果的区域对象多边形的紧凑度和光滑度更好一些,即地物边缘定位较准确,区域对象轮廓边界相对光滑简洁。

5　结　论

本文提出一种引入松弛因子的高分辨率遥感影像自动多层次分割方法。基于区域邻接图最优区域合并原则,利用一个松弛因子调节引导区域对象合并的异质性值的大小,从而通过控制每次递归合并的区域对象个数,提高了整体分割的速度;以区域对象间异质性平均值作为基数,引入另一个松弛因子控制分割过程中层次输出的尺度参数,使整个分割过程自动得到不同尺度的多层次分割结果。实验结果表明,本文提出的分割方法具有较高的分割质量,能够满足进一步影像分析及地物提取的精度要求;本文分割方法的关键是减少了人为因素影响,提高了自动化程度,具有一定的实用价值。并且通过与eCognition软件的分割结果进行对比分析,进一步证明了本文分割方法应用于高分辨率遥感影像具有良好的分割效果。但是,就复杂图像内容的地物目标边界处理和减少狭长区域对象的出现还需要进一步深入研究和实践。

参考文献(References)

Baatz M, Arini N, Schäpe A, Binnig G and Linssen B. 2006. Object-oriented image analysis for high content screening: detailed quantification of cells and sub cellular structures with the cellenger software. Cytometry Part A, 69(7): 652-658

Baatz M and Schäpe A. 1999. Object-oriented and multi-scale image analysis in semantic networks [A]. In: Proc of the 2nd International Symposium on Operationalization of Remote Sensing [C].

Enschede ITC

Beaulieu J M and Goldberg M. 1989. Hierarchy in picture segmentation：a stepwise optimization approach. IEEE Transactions on Pattern Analysis and Machine Intelligence，11（2）：150－163 [DOI：10.1109/34.16711]

Benz U C，Hofmann P，Willhauck G，Lingenfelder I and Heynen M. 2004. Multi-resolution，object-oriented fuzzy analysis of remote sensing data for gis-ready information. ISPRS Journal of Photogrammetry and Remote Sensing，58（3/4）：239－258 [DOI：10.1016/j.isprsjprs.2003.10.002]

Blaschke T. 2010. Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing，65（1）：2－16 [DOI：10.1016/j.isprsjprs.2009.06.004]

Blaschke T，Lang S and Hay G. 2008. Object-Based Image Analysis：Spatial Concepts for Knowledge-Driven Remote Sensing Applications. New York：Springer

Cardelino J，Caselles V，Bertalmío M and Randall G. 2009. A contrario hierarchical image segmentation. In：Proceedings of the 2009 16th IEEE International Conference on Image Processing（ICIP）. Cairo：IEEE：4041－4044 [DOI：10.1109/ICIP.2009.5413723]

Dezsö B，Fekete I，Gera D，Giachetta R and László I. 2012a. Object-based image analysis in remote sensing applications using various segmentation techniques. In：Annales Univ. Sci. Budapest，Sect. Comp，37：103－120

Dezsö B，Giachetta R，László I and Fekete I. 2012b. Experimental study on graph-based image segmentation methods in the classification of satellite images. EARSeL eProceedings，11（1）：12－24

Haralick R M and Shapiro L G. 1985. Image segmentation techniques. Computer Vision，Graphics，and Image Processing，29（1）：100－132

Hay G J and Castilla G. 2006 OBIA. Object-based image analysis：strengths，weaknesses，opportunities and threats（Swot）. The International Archives of the Photogrammetry，Remote Sensing and Spatial Information Sciences

Korting T S，Castejon E F and Fonseca L M G. 2011. Divide and Segment-an Alternative for Parallel Segmentation. In：Procssings XII GEOINFO，November 27－29，2011，Campos do Jordao，Brazil. p 97－104

李德仁，童庆禧，李荣兴，龚健雅，张良培. 2012. **高分辨率对地观测的若干前沿科学问题**. **中国科学：地球科学**，42（6）：805－813

林卉，刘培，杜培军，夏俊士. 2012. **基于改进型统计区域增长的遥感图像分割**. **计算机工程与应用**，48（18）：159－163

吕志勇，张新利，高利鹏，靳菊红. 2012. **基于高分辨率遥感影像数据的** FNEA **分割算法研究与应用分析**. **测绘与空间地理信息**，35（10）：13－16

刘建华，毛政元. 2009. **高空间分辨率遥感影像分割方法研究综述**. **遥感信息**，0（6）：95－101

Marchand-Maillet S and Sharaiha Y M. 1999. Binary Digital Image Processing：A Discrete Approach. San Diego：Academic Press

Marr D. 1982. Vision：A Computational Approach. San Francisco：Freeman & Co.

Murthy Y S S R，Krishna I V M and Gupta S. 2012. Object oriented image analysis for feature extraction from high resolution multispectral imagery. Indian Journal of Electronic and Electrical Engineering，1（1/2）：1－9

Plaza A J and Tilton J C. 2005. Automated selection of results in hierarchical segmentations of remotely sensed hyperspectral images. In：Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. [s. n.]：IEEE：4946－4949 [DOI：10.1109/IGARSS.2005.1526784]

Sarkar A，Biswas M K and Sharma K M S. 2000. A simple unsupervised Mrf model based image segmentation approach. IEEE Transactions on Image Processing，9（5）：801－812 [DOI：10.1109/83.841527]

孙显，付琨，王宏琦. 2011. **高分辨率遥感图像理解**. **北京：科学出版社**

Trias-Sanz R，Stamon G and Louchet J. 2008. Using colour，texture，and hierarchial segmentation for high-resolution remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing，63（2）：156－168 [DOI：10.1016/j.isprsjprs.2007.08.005]

Trimble. 2012. Ecognition Developer 8.7.2 Reference Book. Trimble Germany GmbH，Trappentreustr. 1，D-80339 Munchen，Germany

Tupin F and Roux M. 2005. Markov random field on region adjacency graph for the fusion of sar and optical data in radargrammetric applications. IEEE Transactions on Geoscience and Remote Sensing，43（8）：1920－1928 [DOI：10.1109/TGRS.2005.852080]

Tzotsos A and Argialas D. 2006. Mseg：a generic region-based multiscale image segmentation algorithm for remote sensing imagery. In：Proceedings of ASPRS 2006 Annual Conference. Reno，Nevada：ASPRS：1－13

王爱萍，王树根，吴会征. 2009. **利用分层聚合进行高分辨率遥感影像多尺度分割**. **武汉大学学报（信息科学版）**，34（9）：1055－1058

Wang Z and Bovik A C. 2009. Mean squared error：love it or leave it？A new look at signal fidelity measures. IEEE Signal Processing Magazine，26（1）：98－117 [DOI：10.1109/MSP.2008.930649]

Xia G S，He C and Sun H. 2006. An unsupervised segmentation method using markov random field on region adjacency graph for SAR images. In：Proceedings of CIE'06 International Conference on Radar. Shanghai，China：IEEE：1－4 [DOI：10.1109/ICR.2006.343148]

Zhang H，Fritts J E and Goldman S A. 2008. Image segmentation evaluation：a survey of unsupervised methods. Computer Vision and Image Understanding，110（2）：260－280 [DOI：10.1016/j.cviu.2007.08.003]

章毓晋. 2012. **图像工程（中册）-图像分析（第三版）**. **北京：清华大学出版社**