ISPRS Journal of Photogrammetry and Remote Sensing 78 (2013) 15–25

Contents lists available at SciVerse ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Boundary-constrained multi-scale segmentation method for remote sensing images

Xueliang Zhang [a,b], Pengfeng Xiao [a,b,*], Xiaoqun Song [c], Jiangfeng She [a,b]

[a] Department of Geographic Information Science, Nanjing University, Nanjing, China
[b] Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Nanjing University, Nanjing, China
[c] Information Center of Jiangsu of Land and Resources, Nanjing, China

ARTICLE INFO

Article history:
Received 15 September 2011
Received in revised form 25 August 2012
Accepted 11 January 2013
Available online 13 February 2013

Keywords:
Remote sensing image segmentation
Boundary-constrained multi-scale segmentation
Region growing
Step-Wise Scale Parameter
Edge strength

ABSTRACT

Image segmentation is the key step of Object-Based Image Analysis (OBIA) in remote sensing. This paper proposes a Boundary-Constrained Multi-Scale Segmentation (BCMS) method. Firstly, adjacent pixels are aggregated to generate initial segmentation according to the local best region growing strategy. Then, the Region Adjacency Graph (RAG) is built based on initial segmentation. Finally, the local mutual best region merging strategy is applied on RAG to produce multi-scale segmentation results. During the region merging process, a Step-Wise Scale Parameter (SWSP) strategy is proposed to produce boundary-constrained multi-scale segmentation results. Moreover, in order to improve the accuracy of object boundaries, the property of edge strength is introduced as a merging criterion. A set of high spatial resolution remote sensing images is used in the experiment, e.g., QuickBird, WorldView, and aerial image, to evaluate the effectiveness of the proposed method. The segmentation results of BCMS are compared with those of the commercial image analysis software eCognition. The experiment shows that BCMS can produce nested multi-scale segmentations with accurate and smooth boundaries, which proves the robustness of the proposed method.

1. Introduction

Commercially available high spatial resolution remote sensing images have been significantly increased in recent years. The obtained images provide a lot of details about surface, which are useful for mapping, environmental monitoring, resource investigation, disaster management, and military intelligence, and so on.

Remote sensing imagery needs to be converted into tangible information which can be utilized in conjunction with other data sets (Blaschke, 2010). Because of the high spatial resolution, the ground object would be composed of several pixels. The traditional pixel-based classification method does not make use of spatial concepts and contextual information, which may be limited (Blaschke and Strobl, 2001). On the other hand, the Object-Based Image Analysis (OBIA), which is devoted to partitioning remote sensing imagery into meaningful image objects, and assessing their characteristics through spatial, spectral and temporal scale, is developed to replicate human interpretation of remote sensing images in automated/semi-automated ways (Hay and Castilla, 2006). Advantages of OBIA are utilizing spatial and geometrical properties as well as topological features, and the close relation between real-world objects and image objects (Benz et al., 2004; Hay and Castilla, 2006).

The most common approach of generating objects is image segmentation, which could produce spatially contiguous and spectrally homogeneous segments composed of groups of pixels. A segment is viewed as an object in OBIA. As the first step of OBIA, segmentation influences the effectiveness of OBIA significantly. However, segmentation is an ill-posed problem, which has no unique solution, and even different human interpreters could not produce the same result (Hay and Castilla, 2006). Since it is hard to determine which segmentation is "correct", the multi-scale segmentation strategy seems to be an appropriate choice. Then, a user can choose the suitable scale among multi-scale segmentation results according to the requirement of application.

As recommended in (Lindeberg, 1994), if aiming at describing the structure of unknown real-world signals, a multi-scale representation of data is of crucial importance. Moreover, since it is strongly recognized that landscapes exhibit distinctive spatial patterns associated with different processes at different scales, the multi-scale approaches are required for modern landscape analysis (Hall et al., 2004; Hay et al., 2003). Object-based methods for multi-scale analysis of landscape structure, such as Fractal Net Evolution Approach (FNEA) (Benz et al., 2004) and Multi-scale Object-Specific Analysis (MOSA) (Hay et al., 2001; Hall et al., 2004), contain multi-scale segmentation methods (Baatz and Schäpe, 2000; Castilla et al., 2008).

Scale is about the magnitude or the level of aggregation on which a certain phenomenon can be described (Benz et al.,

* Corresponding author. Tel.: +86 025 83593677.
E-mail address: xiaopf@gmail.com (P. Xiao).

0924-2716/$ - see front matter © 2013 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS) Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.isprsjprs.2013.01.002

2004). In multi-scale segmentations, the object boundaries should follow the ones at finer scales, and segmentation should be constrained by the object boundaries at coarser scales. Thus, objects at different scales can be linked to construct a hierarchy of an image. The FNEA method is able to produce boundary-constrained multi-scale segmentation results, while the rule is not publicly known.

There are many algorithms for remote sensing image segmentation, and some popular ones are listed in (Marpu et al., 2010). Among various segmentation algorithms, the region growing method is widely used for analyzing remote sensing imagery, such as eCognition (Baatz and Schäpe, 2000), Spring (Câmara et al., 1996), and SCRM (Castilla et al., 2008). In this paper, region growing is regarded as the process of aggregating pixels or sub-regions into segments. It can be categorized into seeded region growing (Adams and Bischof, 1994; Fan et al., 2001; Wang et al., 2010) and non-seeded region growing (Chang and Li, 1994; Moghaddamzadeh and Bourbakis, 1997; Tremeau and Borel, 1997; Deng and Manjunath, 2001).

Seeded region growing starts with assigned seeds, and generates regions by adding pixels into its nearest neighboring seed region (Shih and Cheng, 2005). It suffers from the problems of pixel sorting orders for labeling and automatic seed selection (Fan et al., 2005). Especially for complex high spatial resolution remote sensing images, it would be difficult to assign seeds automatically, or even manually. This strategy seems suitable for extracting certain objects in remote sensing images.

Non-seeded region growing corresponds to the region merging process. Region merging can be categorized into global-oriented and local-oriented according to the merging strategy. The typical global-oriented region merging method is hierarchical merging (Beaulieu and Goldberg, 1989). During the iterated merging process, the globally best fitted adjacent regions are merged at each iteration. Hierarchical merging strategy is of strong constraint for the optimization problem and widely used for image segmentation (Haris et al., 1998; Cheng and Sun, 2000; Lhermitte et al., 2008; Yu and Clausi, 2008; Trias-Sanz et al., 2008). However, the global best searching process is time-consuming, and it may lead to an uneven growth of regions.

The local-oriented region merging method finds suitable pairs of adjacent regions for merging within local domain (Câmara et al., 1996; Baatz and Schäpe, 2000; Wang et al., 2010). It is more efficient than the global-oriented method, and can achieve a uniform growth of regions. The merging process is controlled by the threshold of region homogeneity, or of similarity between adjacent regions. Since small threshold permits fewer merges, the mean region size would change with the threshold. Then, multi-scale segmentations are produced by setting different thresholds. However, for local-oriented region merging, if multi-scale segmentations are generated just by setting different thresholds, object boundaries at different scales cannot be constrained with each other. In this paper, the Step-Wise Scale Parameter (SWSP) strategy is proposed to solve this problem. SWSP uses a set of scale parameters, rather than a single parameter, to control the local-oriented region merging process to generate boundary-constrained multi-scale segmentations.

Region-based methods can provide closed contour, but it tends to sacrifice details in the image to gain samples large enough for the calculation of useful statistics (Muñoz et al., 2003). Hence, there is a trend to integrate region growing and edge detection method to obtain closed and accurate object boundaries (Pavlidis and Liow, 1990; Zhu and Yuille, 1996; Tabb and Ahuja, 1997; Arbeláez et al., 2011; Li et al., 2010). Since it is complex to integrate two different segmentation methods, in this paper, the feature of edge strength is introduced as a merging criterion, which is similar to the work in (Yu and Clausi, 2008).

Hence, the main objective of this paper is to propose the Step-Wise Scale Parameter (SWSP) strategy, which provides an explicit

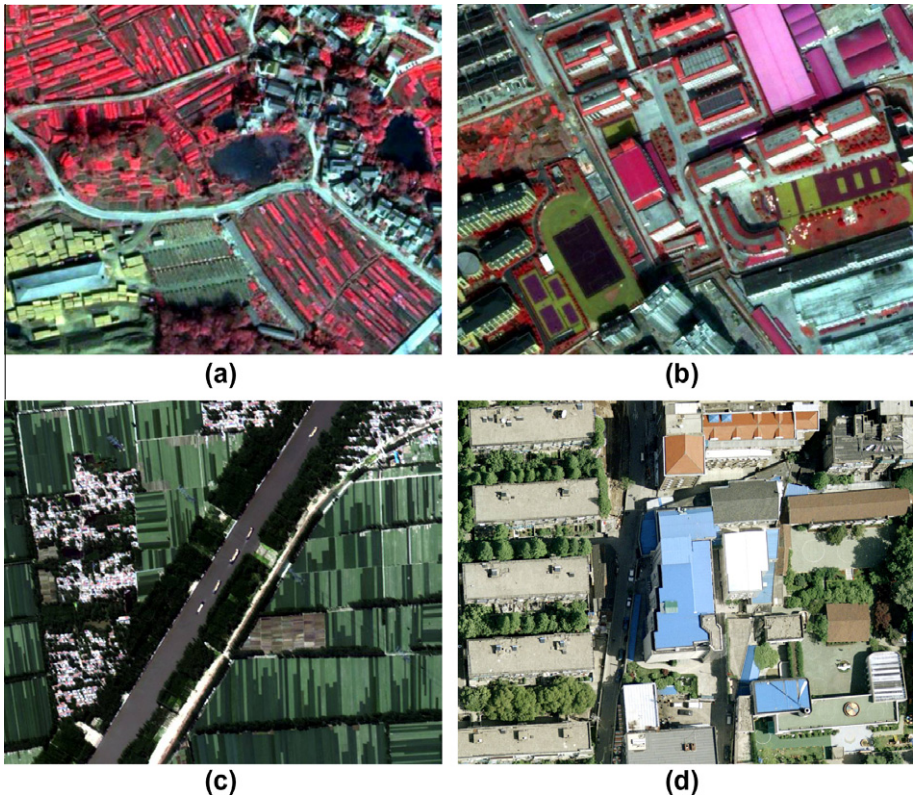

Fig. 1. Original test images, (a) and (b) are QuickBird images, (c) is a WorldView image, (d) is an aerial image. Specific parameters of all the test images are listed in Table 1. (a), (b), (c), and (d) correspond to T-1, T-2, T-3, and T-4, respectively.

Table 1
Parameters of original test images in Fig. 1.

Image	Platform	Size (pixel)	Spatial resolution (m)	Band composition	Band combination	Position
T-1	QuickBird	505 × 404	0.6	B, G, R, NIR	NIR, R, G	Nanjing, China
T-2	QuickBird	658 × 504	0.6	B, G, R, NIR	NIR, R, G	Hangzhou, China
T-3	WorldView	787 × 652	2.0	Coastal, B, G, Yellow, R, Red Edge, NIR1, NIR2	R, G, B	Xuzhou, China
T-4	Aerial plane	1500 × 1200	0.2	B, G, R	R, G, B	Nanjing, China

rule to achieve nested objects at multiple scales. Moreover, the feature of edge strength is adopted to help to improve the accuracy of object boundaries. The study area and data are presented in Section 2. The proposed method is illustrated in Section 3, and followed with experimental results in Section 4. Finally, conclusions and discussions are included in Section 5.

2. Study area and data

Four typical test images are shown in Fig. 1, and the parameters of which are presented in Table 1. Fig. 1a is a subset of a QuickBird scene in Nanjing City, China, which shows the typical rural landscape with irregular settlement, farmland, bending road, and pond. (b) is a subset of a QuickBird scene in Hangzhou City, China, representing a new urban area. The right side of (b) is a part of a factory, and a school is situated at the lower left corner. The spatial resolution of (a) and (b) is sharpened to 0.6 m using the pan-sharping method proposed by Zhang (2002). (c) is a subset of a WorldView scene in Xuzhou City, China, with rural landscape different from (a). There are regular settlements, zonal forest, and rectangular farmlands with varied color in the image. (d) is a part of an aerial image in Nanjing City, China. This image represents typical urban landscape with dense buildings. Regular residential area is located on the left side. Other tall buildings with colorful roofs are situated on the right side. We call the image of Fig. 1a as T-1, (b) as T-2, (c) as T-3, and (d) as T-4 in the following parts.

3. Methods

The proposed Boundary-Constrained Multi-Scale Segmentation (BCMS) method contains three steps: initial region growing, RAG construction, and region merging. The flow diagram is shown in Fig. 2. To begin with, the initial region growing produces over-segmented initial segmentation with high precision. Then, upon the initial segmentation, RAG is built. Finally, the region merging process is applied on RAG to produce multi-scale segmentation results.

3.1. Initial region growing

Before initial growing, the median filter is applied to reduce noise. The value of a pixel is replaced by the median of the spectral values in the neighborhood (Gonzalez and Woods, 2004). The initial region growing process starts from individual pixels, and contains two stages.

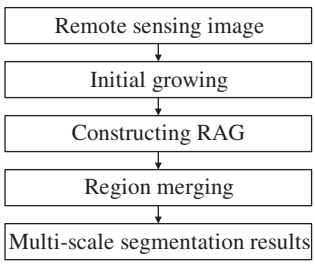

Fig. 2. Flow diagram of the proposed Boundary-Constrained Multi-Scale Segmentation (BCMS) method.

At the first stage, the pixel with high local homogeneity is assigned as initial growing pixel. Based on the coefficient of variation of a 3 × 3 window, the high local homogeneity is defined as:

$$Max(Std_i/M_i|i = 1, 2, \ldots, n) < T_h. \quad (1)$$

where M and Std are the mean value and standard deviation of the pixels in the window, respectively, n is the number of spectral bands, and T_h is the threshold of homogeneity. The local homogeneity constraint strategy helps to make the growing process start from the interior, rather than near the object boundary. If T_h is set small, there will be fewer initial growing pixels, but the selected pixels have greater possibility of being located inside an object. In this paper, T_h is set as 0.05.

Once an initial growing pixel is selected, the scale-constrained local best region growing process is applied to generate a segment. Then, the next initial growing pixel is selected to generate another segment. The select-and-grow process is repeated until all the pixels with high local homogeneity have been evaluated.

The region growing process is to add individual pixels into the seed region. The spectral difference (SD) is used to indicate the similarity between the seed region and its adjacent pixel, as shown in the following equation:

$$SD = \sum_{i=1}^{n} (M_i - V_i)^2 / n. \quad (2)$$

where M is the mean spectral value of the seed region, and V is the spectral value of the adjacent pixel.

The growing process is constrained by scale parameter, which consists of the size (A) and standard deviation (Std) of a region. Std corresponds to the homogeneity of a region. The scale-constrained local best growing process is described as follows:

T-SD	Threshold of spectral difference.
T-A	Threshold of region size.
T-Std	Threshold of standard deviation of a region.
Step 1	Calculate *SD* value for each adjacent pixel of the seed region.
Step 2	Search the adjacent pixel *P* with minimum *SD* value.
Step 3	If the *SD* value of *P* is less than *T-SD*, go to step 4; else, terminate the growing process.
Step 4	Supposing that *P* has been added into the seed region, calculate *A* and *Std* of the seed region.
Step 5	If *A* and *Std* are both less than *T-A* and *T-Std*, respectively, add *P* into the seed region, recalculate the mean value of the seed region, and go to step 1; else, terminate the growing process.

During the growing process, three parameters are used to control the homogeneity of initial segments. The local best adjacent pixel with *SD* value smaller than *T-SD* is just considered as a candidate to be added in the seed region. Then the homogeneity and size of the seed region is evaluated. Only if the region satisfies the scale parameter, the growing action is performed.

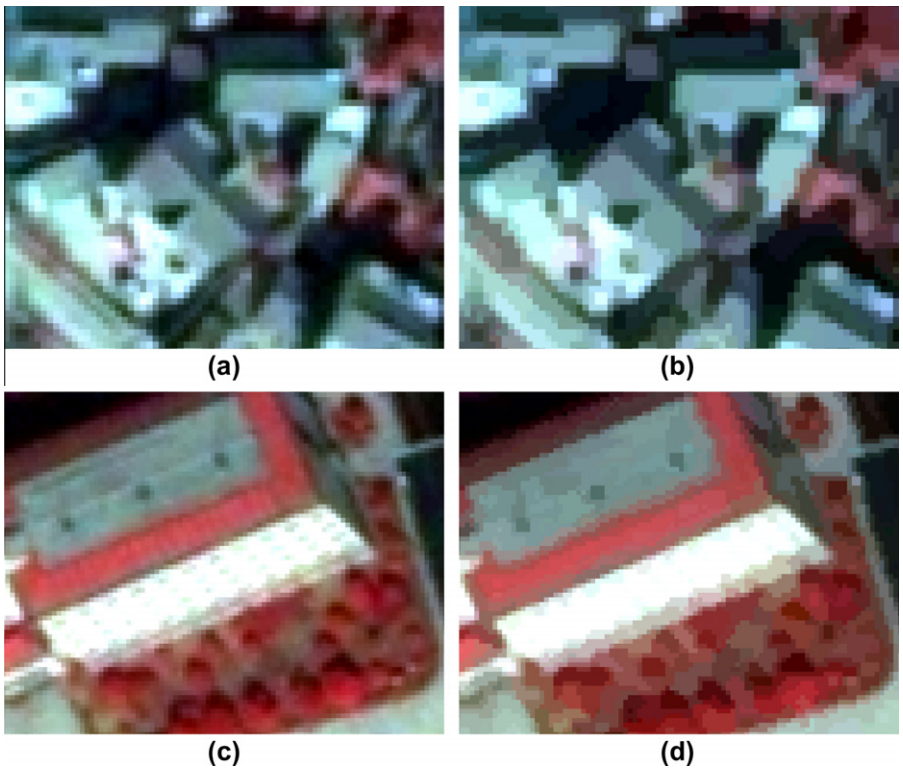

Fig. 3. Sample initial segmentation results. (a) and (b) are the zooms of the same area of T-1. Similarly, (c) and (d) correspond to T-2. (a) and (c) are from the original image, and (b) and (d) are from the initial segmentation results shown with the segment's mean spectral value.

At the second stage, the select-and-grow process is also repeated. But the selecting strategy of initial growing pixels is different with that at the first stage. A pixel which has not been involved in the growing process would be assigned as the initial growing pixel according to the order of traversing the image, which is from top to down, and left to right. After an initial growing pixel is selected, the same scale-constrained local best growing process is applied to generate an initial segment. The second stage stops when the set of initial segments covers the entire image plane.

The initial segmentation could be over-segmented, but should try to avoid wrong segmentation. Hence, the parameters of *T-SD*, *T-Std*, and *T-A* are set as 10, 8, and 40, respectively. If the thresholds are set large, the pixel-based region growing process will take a lot of time, and the number of wrong merges will increase. If the thresholds are set too small, the initial segmentation will contain too many regions, leading to the increase of calculation burden in the successive merging process.

Taking T-1 and T-2 for example, there are 33951 and 26861 regions in the initial segmentation, respectively. Fig. 3a and b are two zooms of the same small area of T-1, among which (a) is a subset of the raw image and (b) is from the initial segmentation shown with the segment's mean spectral value. Similarly, Fig. 3c and d corresponds to two zooms of the same area of T-2. According to visual assessment, even though some adjacent similar pixels are aggregated into initial segment, which helps to reduce the number of elements representing the image, the structure and detailed features of the original image are preserved very well in the initial segmentation.

3.2. Region Adjacency Graph construction

Region Adjacency Graph (RAG) is composed of nodes and arcs. In RAG, a node represents a segment and an arc indicates the adjacency between two nodes (Haris et al., 1998; Sarkar et al., 2000). In this paper, two arrays: the node array (*Node*) and arc array (*Arc*) are built to represent RAG. Raster topology is introduced to describe the adjacency between nodes and arcs. The data structure of *Node* and *Arc* is described as:

```
Data structure of Arc
  struct Arc {int          // labels of two adjacent nodes of an
  AdjNode[2];              arc
   float Weight[];         // arc weight
   float                   // edge features
   EdgeFeatures[];
   },
Data structure of Node
  struct Node {int         // serial number of arcs surrounding a
  SurArc[];                node
   int NodeFeature[];      // features of the corresponding region
   point                   // location of the pixels composed of a
   NodeCompose[];          node
   }
```

The variate of *AdjNode* in *Arc* records the labels of two nodes connected by an arc. *SurArc* in *Node* records the serial number of all the arcs surrounding the node. In order to find the adjacent nodes of the current node in merging process, we can first visit *SurArc* in *Node*, and then find out *AdjNode* in corresponding arcs. The variable of *Weight* in *Arc* is used to record the arc weight, which indicates the merging possibility of two adjacent nodes. *Weight* is calculated according to the merging criterion (defined in part 3.3). *EdgeFeatures* records the features of edge strength (defined in part 3.3) and length. Similarly, the variable of *NodeFeatures* in *Node* records the features of the corresponding region. It contains the region features of size, mean spectral value, and sum of square error in this paper. Actually, other features can be recorded in *EdgeFeatures* and *NodeFeatures* according to the specific requirements of

forming the merging criterion. Finally, the variable of *NodeCompose* in *Node* records the location of all the pixels in the corresponding region.

RAG is constructed upon the initial segmentation. Firstly, assign each initial segment a distinctive label. Then, calculate the features in *Arc* based on the common edge between each pair of adjacent regions. Finally, calculate the features in *Node* based on all the segmented regions.

The successive region merging process is applied only on RAG without involving the original image. Hence, we recommend storing RAG in internal memory or hard disk. Then, different segmentations can be produced without repeating the initial region growing and RAG construction process.

3.3. Multi-scale region merging

The local mutual best merging strategy (Baatz and Schäpe, 2000), which has the strongest optimal constraint among the local-oriented region merging strategies, is adopted. Object A and B are mutual best neighbors if A is B's neighboring object fulfilling the merging criterion best, and the same goes for B.

Since the objective of segmentation is to generate homogeneous regions, the change of homogeneity after a visual merge does make sense to be integrated in the merging criterion. When the homogeneity change is small, there is a greater merging possibility. The feature of standard deviation (*Std*) can indicate the homogeneity of a region. Thus, the change of *Std* (*CStd*) is used to evaluate the change of homogeneity after a visual merge (Baatz and Schäpe, 2000). *CStd* is defined as:

$$CStd = Std - (a_1 Std_1 + a_2 Std_2)/(a_1 + a_2). \quad (3)$$

where *Std*, Std_1, and Std_2 represent the standard deviation of the newly created region and two original neighboring regions, respectively. a_1 and a_2 are the size of two neighbors.

Tiny meaningless regions should be removed in segmentation result. Besides, region size corresponds to the segmentation scale directly. Hence, the feature of region size is integrated in merging criterion.

According to the criteria for good segmentation proposed by Haralick and Shapiro (1985), the boundaries of segments should be simple, and not ragged. Then, the feature of compactness is introduced. Compactness is small for the region with simple outline. Similar to *CStd*, change of compactness (*CComp*) after a visual merge is defined as Eq. (4). If *CComp* is small, the corresponding pair of adjacent regions is more inclined to be merged.

$$CComp = L/\sqrt{a} - (a_1 L_1/\sqrt{a_1} + a_2 L_2/\sqrt{a_2})/(a_1 + a_2). \quad (4)$$

where L, L_1, and L_2 are the outline length of the newly created region and two original regions, respectively. a is the size of the newly created region.

Another criterion for good segmentation is spatial accuracy. Since the features of *CStd* and *CComp* do not impose on local structure information, they cannot ensure to produce accurate region boundaries. Furthermore, *CComp* would sacrifice the boundary accuracy to generate compact objects. Hence, the feature of edge strength is introduced as a merging criterion. Small edge strength indicates greater merging possibility. Edge strength is defined as the mean edge strength in terms of all the points belonging to the edge. The edge strength of a point (*ESP*) is defined as:

$$ESP = \omega|V_{L1} - V_{R1}| + (1-\omega)|V_{L2} - V_{R2}|. \quad (5)$$

where V_{L1} and V_{R1} are the spectral value of the pixel adjacent to the edge pixel on both sides in the direction perpendicular to that of the edge. For a vertical edge, V_{L1} and V_{R1} correspond to the west and east neighbors, respectively. While for a horizontal edge, they correspond to the north and south neighbors, respectively. The definition of V_{L2} and V_{R2} is similar, but the distance of V_{L2} and V_{R2} to edge is one pixel longer than that of V_{L1} and V_{R1}. ω is the weight parameter, and we set it 0.6 to increase the contribution of closer points. Then, the edge strength (*ES*) is defined as:

$$ES = \sum_{i=1}^{l} ESP_i/l. \quad (6)$$

where l is the edge length.

Totally, there are four features: *CStd*, *CComp*, region size, and *ES*, integrated in the merging criterion. Since the definition of *CStd* and *CComp* is similar, the summation strategy is used to combine the two features. The other two features are multiplied with the sum of *CStd* and *CComp*. The feature of region size relates to segmentation scale directly. It is integrated in the merging criterion without normalization. The feature of *ES* just corresponds to local structure, the effect of which should not be too significant. Hence, *ES* is normalized before integration, as shown in the following equation:

$$ES_n = \exp(ES/ES_{max}). \quad (7)$$

where ES_n is the normalized *ES*, ES_{max} is the maximum *ES* in initial segmentation. After normalization, the variation range of ES_n is [1, e]. In this case, the effect of *ES* is limited.

Finally, the merging criterion (*MC*) is defined as below:

$$MC = (a_1 + a_2)(CStd + CComp)ES_n. \quad (8)$$

The arc weight in RAG is calculated according to *MC*. The small arc weight indicates that the corresponding adjacent regions are more inclined to be merged. When two neighboring regions A and B have the same adjacent arc with minimum weight, they are local mutual best neighbors.

The region merging process is controlled by the threshold of arc weight, which serves as the scale parameter. Only if the arc weight between the local mutual best neighbors A and B is smaller than the threshold, the two neighbors are allowed to be merged. If the threshold is set large, more merges are allowed, leading to segmentation with larger mean region size or coarser scale. Hence, multi-scale segmentations are generated through setting different thresholds. But in this case, the multi-scale objects cannot be linked with each other. In other words, the object boundaries do not matching those at higher or lower scales. In order to produce the boundary-constrained multi-scale segmentations, the Step-Wise Scale Parameter (SWSP) strategy is proposed.

SWSP is composed of a set of increased scale parameters $\{S_1, S_2, \ldots, S_k, \ldots, S_{max} \mid S_1 < S_2 < \ldots < S_k < \ldots < S_{max}\}$. When it is necessary to generate the segmentation at scale S_k, the region merging process does not go straightforward to S_k. However, it firstly satisfies the lowest scale parameter S_1, and then up to S_2, and finally, after the step-by-step increasing process of scale parameter, reaches S_k. Thus, the merging process is not controlled by a single but a set of scale parameters.

Since *MC* consists of four different features, the arc weight is unitless, leading to the unitless scale parameter. As region size increases with scale, and *CStd* also increases with scale, the arc weight is increased nonlinearly with scale. Accordingly, SWSP contains a set of scale parameters increased nonlinearly, as shown in the following equation:

$$S_k = 5 + 25(k-1)^2, k = 1, 2, \ldots n. \quad (9)$$

where k is the serial number of scale.

Based on SWSP, the multi-scale region merging process is described as below:

k	serial number of scale, $k = 0$ at first.
S_k	scale parameter.
OutputScale	the serial number of the output scale.
Step 1	If $k = OutputScale$, terminate the process; else, $k = k + 1$.
Step 2	If all the regions have been evaluated at scale k, go to step 1; else, find a region A which is not involved in the merging process and go to step 3.
Step 3	Search the local mutual best neighboring region B of A. If B does not exist, go to step 2; else, go to step 4.
Step 4	If the arc weight between A and B is larger than S_k, go to step 2; else, go to step 5.
Step 5	Add B into A, and recalculate the node and arc array. Go to step 3.

During the region merging process, since the segmentation is generated based on the next lower scale, the object at higher scale is a combination of several small objects at lower scale, resulting in nested objects at different scales.

4. Results

4.1. Effect of features in merging criterion

Segmentation result of test image T-2 at scale 10 is used for the analysis, as shown in Fig. 4. In Fig. 4a, even though the merging criterion contains only two features of *CStd* and region size, different objects are distinguished with each other, showing that the two features can determine the basic structure of segmentation. But the region outlines in (a) are complex and ragged from visual assessment. In Fig. 4b, *ES* is combined with *CStd* and region size in the merging criterion, but the outlines seem more complex than (a). Turning to Fig. 4c, the feature of *CComp* is combined with *CStd* and region size. The region number of (c) is the same as that of (a), which is 698. But the boundaries in (c) are much simpler than (a), showing that *CComp* can help to reduce the boundary complexity significantly.

In Fig. 4d, the merging criterion includes all the four features. Comparing (d) with (c), the fundamental segmentation structure is similar, and the difference mainly exists in the local structure. The same typical part in (c) and (d) is zoomed in and shown in (e) and (f), respectively. The differences within the black rectangles

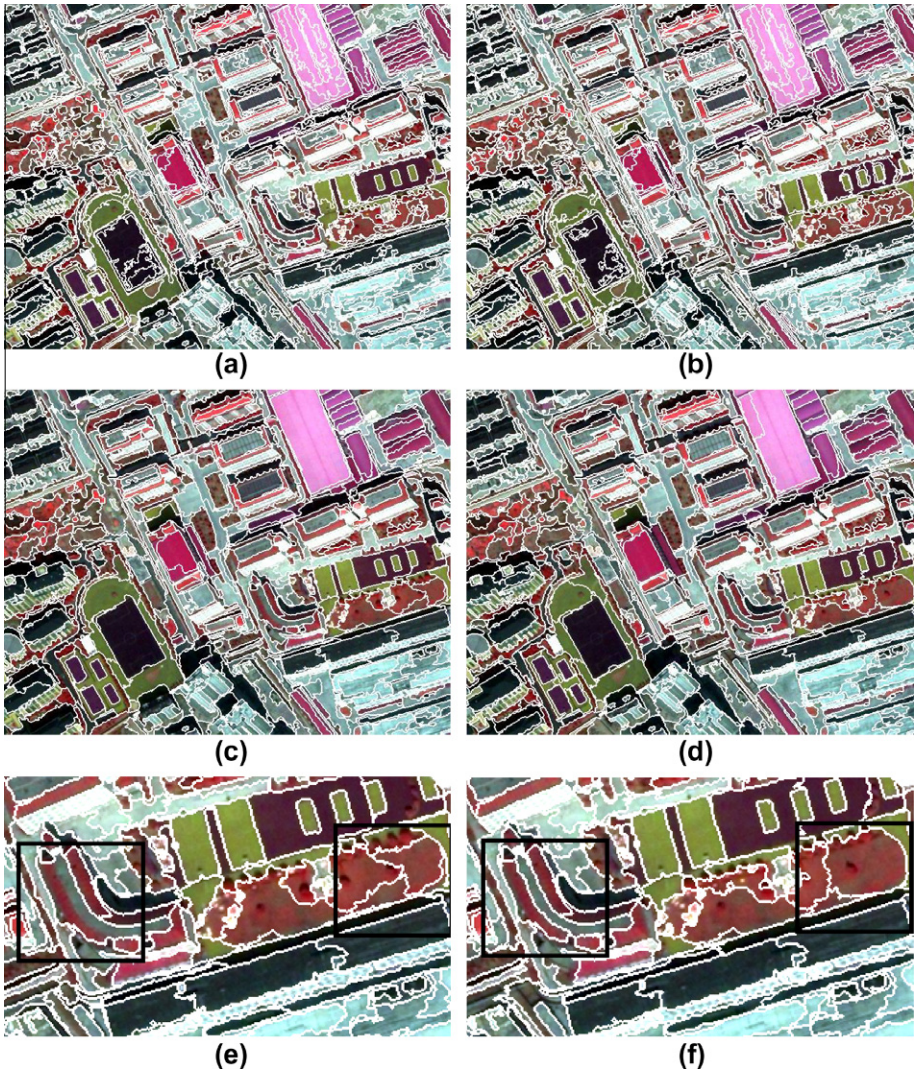

Fig. 4. Analyzing the effect of features in merging criterion, (a–d) are segmentation results of T-2 at scale 10: (a) the features of *CStd* and region size are integrated in the merging criterion, (b) the merging criterion integrates *CStd*, region size, and *ES*, (c) the merging criterion integrates *CStd*, area, and *CComp*, (d) the merging criterion integrates all the four features of *CStd*, area, *CComp*, and *ES*, (e) and (f) are the zooms of the same small part in (c) and (d), respectively.

show that the integration of edge strength in merging criterion can help to improve the accuracy of boundaries without changing the fundamental segmentation pattern. Usually, if the scale parameter is set large, some neighboring regions with great difference would be wrongly merged to gain regions large enough to satisfy the scale parameter. However, the feature of edge strength helps to improve the weight of arc connecting two heterogeneous regions and prevent some wrong merges.

4.2. Multi-scale segmentation results

The test images of T-1, T-3, and T-4 are used to show the boundary-constrained multi-scale segmentation results of BCMS.

Four typical multi-scale segmentation results of T-1 are selected and shown in Fig. 5. The result at scale 10 is shown in (a). There are 614 regions. At this fine scale, even the individual houses in the irregular residential area are differentiated with each other, and the farmlands with different tones are also separated. The result in (b) is at scale 14 with 320 regions. In this result, some neighboring houses with similar tone are merged, and some similar adjacent farmlands are also merged. In (c), the scale is increased to 20, and the region number is reduced to 158. Some similar farmlands are further merged to form larger objects. The homogeneous regions such as shadows, ponds, and the main roads are segmented into independent single objects. Since the scale parameter is increased, some wrong merges emerge, such as the wrong merge of lane and farmland. In the segmentation result at scale 30, there are only 57 regions, as shown in (d). At this coarse scale, the ground objects, such as farmland, impervious surface, water, forest, are segmented into just a few large regions.

The multi-scale segmentation results of T-3 are shown in Fig. 6. In (a), the result at scale 14 with 864 regions is shown. At this scale, even the individual trees in residential area are identified, and the farmlands with different tones are separated. In (b), the result is at scale 20. There are 467 regions. The farmlands with similar tone are segmented as single objects. Some small trees are merged with houses to form larger objects in residential area. In the zonal forest, many tiny regions in (a) are removed. In (c), since the scale is increased to 30, there are only 227 regions. At this scale, some small farmlands with different tones are merged into larger farmland objects. In the residential area, more single trees are merged with houses, and the residential areas at top and upper right are even segmented as single objects. In (d), the scale is 50, and the region number is 142. At this scale, the zonal forests are segmented into a few large objects. Larger farmland and river objects are generated.

In Fig. 7a, the segmentation result of T-4 at scale 10 is listed. There are 3018 regions. At this fine scale, even the cars on road are segmented as single objects. Some complex objects, such as trees and walls are segmented into a lot of small regions. But the smooth objects, such as some roofs and ground areas are segmented into several large regions. This is because that the *CStd* values are quite small for these smooth objects. In (b), the scale is increased to 20. The region number is 971. At this scale, the segmentation pattern of smooth area is similar to (a). But in textural areas, a lot of tiny regions are merged into larger objects. For example, some trees are segmented as single objects. In (c), the segmentation scale is 40. There are 313 regions. Most roofs, ground areas, and shadow areas are segmented as single objects. Some trees are combined into an object, while some are still segmented as single objects. In (d), the scale is 80, and there are only 108 regions. At this scale, no single trees are extracted. The connected trees are segmented as large single objects.

According to visual analysis of the multi-scale segmentation results of T1, T3, and T4, the boundaries of objects at different scales are constrained with each other and the object boundaries have

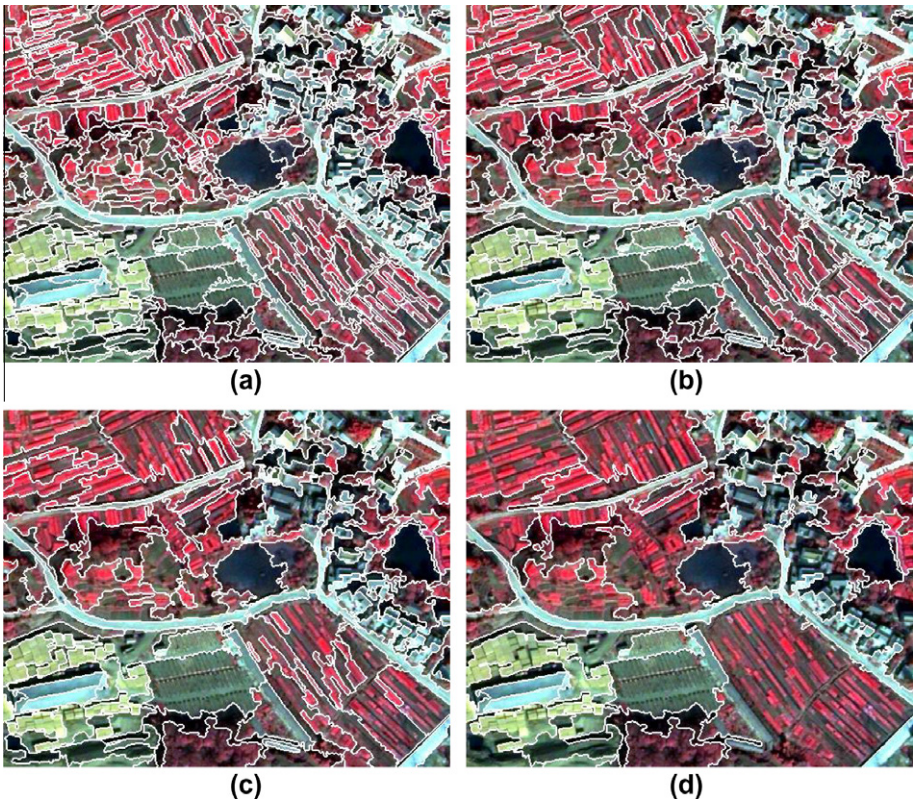

Fig. 5. Multi-scale segmentation results of T-1: (a) scale 10, 614 regions; (b) scale 14, 320 regions; (c) scale 20, 158 regions and (d) scale 30, 57 regions.

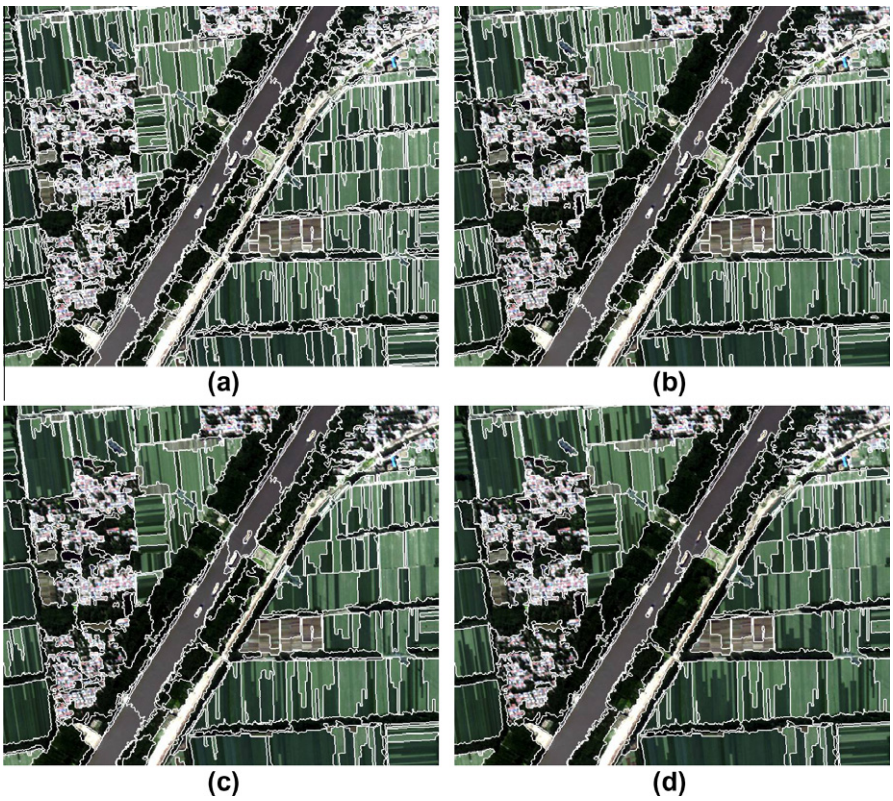

Fig. 6. Multi-scale segmentation results of T-3: (a) scale 14, 864 regions; (b) scale 20, 467 regions; (c) scale 30, 227 regions and (d) scale 50, 142 regions.

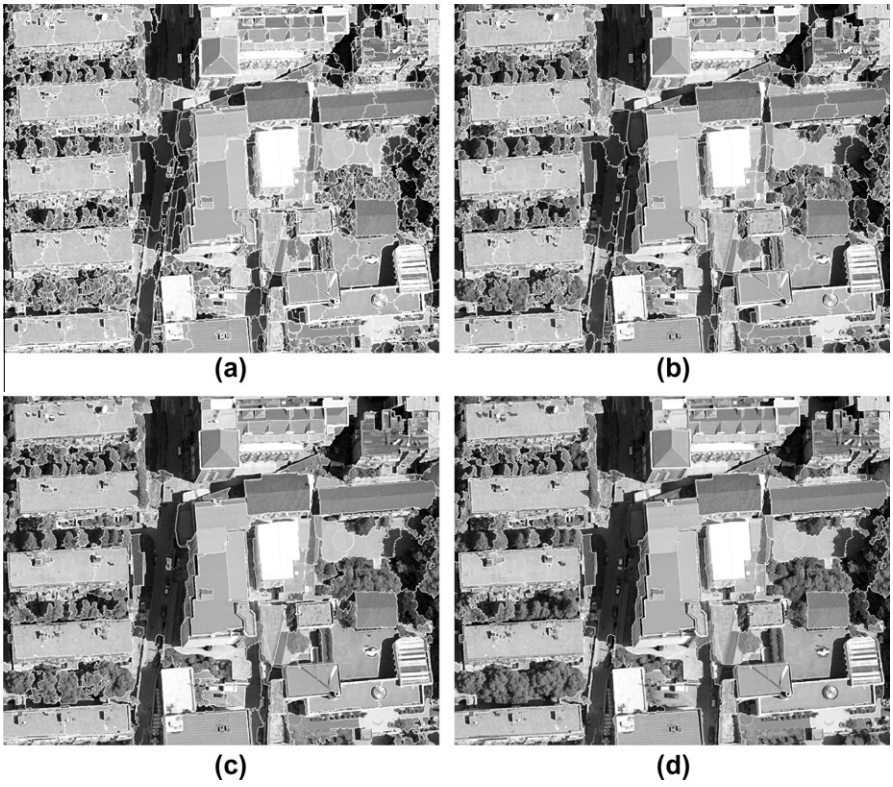

Fig. 7. Multi-scale segmentation results of T-4: (a) scale 10, 3018 regions; (b) scale 20, 971 regions; (c) scale 40, 313 regions and (d) scale 80, 108 regions.

great spatial accuracy. Along with the increase of scale parameter, the region number is decreased, which leads to larger mean region size and accordingly coarser segmentation scale. The multi-scale segmentations from BCMS are able to describe ground objects at different levels of detail. At the fine segmentation scales, small objects, such as trees and houses, are segmented into single regions.

As the scale getting coarser, larger objects, such as forests, farmlands, and residential areas, are segmented into a few large regions or even single regions.

4.3. Comparing BCMS with eCognition

The multi-resolution segmentation method embedded in the commercial software eCognition (Baatz and Schäpe, 2000; Benz et al., 2004) is used to compare with the proposed segmentation method BCMS. A trial version of the software eCognition Developer 8 is used. The parameter of color weight is set as 0.5. Three multi-scale segmentation results of T-2 from BCMS and eCognition are shown in Fig. 8. Region number is used to indicate the segmentation scale. In Fig. 8a, c, and e, there are 730, 435, and 215 regions in the eCognition results, respectively. Accordingly, the corresponding results of BCMS are shown in Fig. 8b, d, and e. There are 772, 459, and 246 regions, respectively.

According to visual assessment, the fundamental segmentation pattern of the two methods is similar. Both BCMS and eCognition can produce multi-scale segmentations with simple and accurate boundaries. Since the introduction of edge strength in BCMS, there are some small differences in local structure between them. Some wrong merges caused by satisfying the large scale parameter are avoided in the BCMS results. The ability of describing objects at different levels of detail is similar between the two methods. Though eCognition is also able to produce the boundary-constrained multi-scale segmentation results, the method is not opened to the public.

In addition, the supervised evaluation method is used for the comparison. The reference segmentation of T-2 (as shown in Fig. 9) is digitized by a remote sensing expert who has not participated in the experiment. Three supervised evaluation methods are adopted. One is the method E proposed by Carleer et al. (2005). E equals to the ratio of the number of mix-segmented pixels to the total number of pixels in the test image. We calculate the Rightly-segmented Ratio (RR) according to the formula $RR = 1 - E$. The other two evaluation methods are the Rand Index (RI) (Rand, 1971) and the Adjusted Rand Index (ARI) (Hubert and Arabie, 1985). The former is a measure of partition correspondence based on the comparison of object pairs, and the latter is based on the comparison of object triples. The segmentation with large RI and

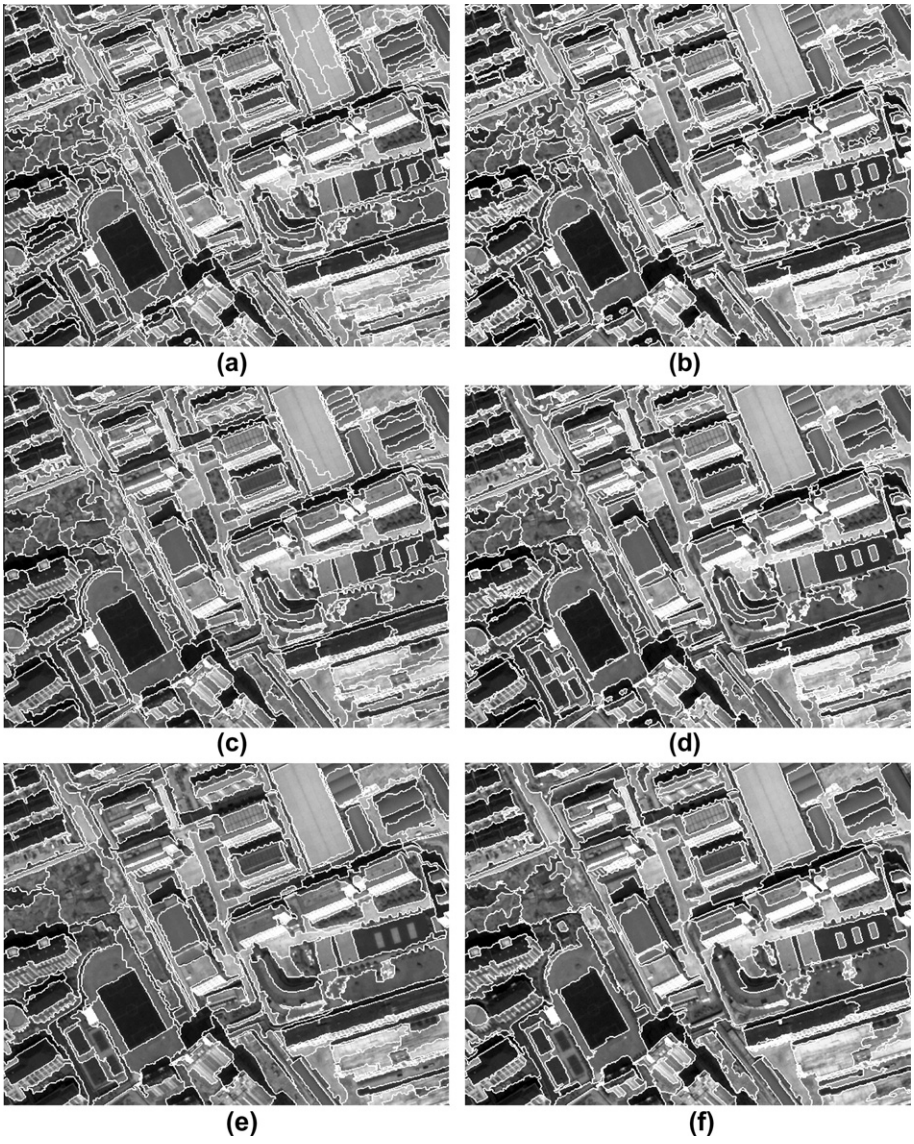

Fig. 8. Multi-scale segmentation results from eCognition and the proposed method BCMS for T-2. (a), (c), and (e) are results from eCognition, (b), (d), and (f) are results from BCMS. (a) scale parameter 35, 730 regions; (b) scale 10, 772 regions; (c) scale parameter 45, 435 regions; (d) scale 14, 459 regions; (e) scale parameter 80, 215 regions and (f) scale 20, 246 regions.

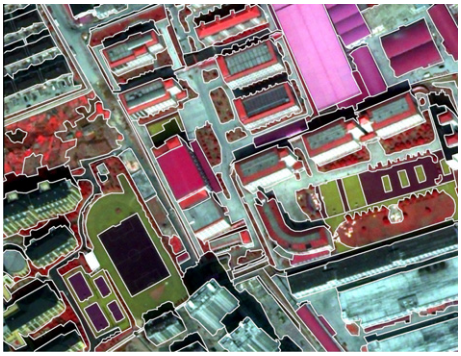

Fig. 9. The reference segmentation image of T-2. There are 186 regions.

Table 2
Supervised evaluation results of the multi-scale segmentations from BCMS and eCognition for the test image T-2. The scale is indicated by the number of regions in segmentation result. RR represents Rightly-segmented Ratio. RI represents Rand Index. ARI represents Adjusted Rand Index.

Scale (region number)	eCognition	BCMS
eCognition/BCMS	RR/RI/ARI	RR/RI/ARI
730/772	0.907/0.987/0.273	0.900/0.988/0.346
435/459	0.895/0.988/0.328	0.868/0.989/0.435
215/246	0.794/0.989/0.506	0.808/0.989/0.503

ARI values is considered to produce greater correspondence to the reference segmentation.

The supervised evaluation results of segmentations in Fig. 8 are presented in Table 2. For the fine-scale segmentations in Fig. 8a and b, RR of the eCognition result is 0.7% higher than that of the BCMS result, but both RI and ARI of the BCMS result are higher. The RI difference is about 0.001, and the ARI difference reaches 0.073. This shows that the number of mix-segmented pixels is similar, but the correspondence of the BCMS result to the reference segmentation image is greater than the eCognition result. At the median scale, which corresponds to Fig. 8c and d, RR of the BCMS result is 2.7% lower than that of the eCognition result. This is mainly owing to the segmentation difference in the natural land on the left side of the image. But RI of the BCMS result is still 0.001 higher than that of the eCognition result, and ARI is even 0.107 higher. For the segmentations at coarse scale, as shown in Fig. 8e and f, RR of the BCMS result is 1.4% higher than that of the eCognition result, but both the RI and ARI values are similar.

Hence, in terms of T2, for segmentations of BCMS and eCognition at fine scales, the number of mix-segmented pixels is similar, but the BCMS result has a greater correspondence to the reference segmentation according to RI and ARI. As the segmentation scale increases, none of the two methods is overwhelmed to the other in terms of the number of mix-segmented pixels, and the difference of segmentation pattern is getting smaller according to the change of RI and ARI.

5. Conclusion and discussion

The Boundary-Constrained Multi-Scale Segmentation (BCMS) method for remote sensing image segmentation is proposed in this paper. Firstly, the initial segmentation is generated through the local best region growing process. Then, the local mutual best region merging process is applied on RAG to produce multi-scale segmentations. The experiments have shown the robustness of BCMS. The main contributions of BCMS can be concluded as twofold:

(1) The Step-Wise Scale Parameter (SWSP) strategy is proposed for the locally mutual best region merging process. Since the merging process is controlled by a set of incremental scale parameters, BCMS is able to produce boundary-constrained multi-scale segmentation results. Then, the multi-scale objects can be linked to construct a hierarchy of an image, which would be quite useful in some application domains, such as cartography, information extraction, and object-based image classification. Actually, SWSP can be used for all the local-oriented region merging methods, allowing them to produce boundary-constrained multi-scale segmentation results.

(2) The feature of edge strength is introduced as a merging criterion to help to produce accurate boundaries. The indexation strategy is adopted for normalization of edge strength and the feature is multiplied with others to form the merging criterion. The experiments prove that the features of standard deviation and region size correspond to the basic segmentation pattern, while the feature of edge strength will influence the local structure of segmentation. Edge strength helps to prevent wrong merges caused by satisfying large scale parameter.

Even though BCMS can produce multi-scale segmentations with simple, accurate boundaries, there are still some problems deserved to be researched in the future.

In the merging criterion, the sum strategy is used to combine *CStd* and *CComp*, but the physical meaning and variation range of the two features are significantly different with each other. Even though the segmentation results are appealing, a better theoretical basis for combining both features is needed to be researched. In addition, the feature of compactness may influence the boundary accuracy if the image contains objects with complex boundaries. In order to reduce the negative effect, the weighted sum strategy can be used for the combination of *CStd* and *CComp*.

The scale parameter is the threshold of arc weight, which is calculated according to the merging criterion. The feature of region size in merging criterion corresponds to the segmentation scale directly, but it cannot take the region homogeneity and boundary accuracy into account. The feature of *CStd* can help to constrain the region homogeneity, but it mainly corresponds to the global segmentation pattern. Once the scale parameter is set large, some wrong merges are appeared to generate large objects satisfying the scale parameter. The feature of edge strength will help to prevent some wrong merges, but the effect is not strong enough and limited in local domain. Especially, as the scale increases, the effect of edge strength is getting weaker. It is deserved to develop a better way to integrate edge strength in the merging criterion, or to find other features which can correspond to both scale and precision quite well.

The scale parameters in SWSP are predefined before the merging process. Thus, the segmentation at the "optimal" scale may not be produced for various remote sensing images. It is deserved to research the adaptive SWSP strategy in the future.

The proposed BCMS segmentation method is integrated in the software of RSEG developed by the authors, which aims at the object-based information extraction from high resolution remote sensing images.

Acknowledgements

This work is supported by Special Fund for Scientific Research of Non-Profit Industry of Ministry of Land and Resources of China (Grant No. 201011015-1), National Basic Research Program of China (Grant No. 2011CB952001) and a project funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions. The authors would like to thank the anonymous referees for their contributing comments.

References

Adams, R., Bischof, L., 1994. Seeded region growing. IEEE Transactions on Pattern Analysis and Machine Intelligence 16 (6), 641–647.

Arbeláez, P., Maire, M., Fowlkes, C., Malik, J., 2011. Contour detection and hierarchical image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 33 (5), 898–916.

Baatz, M., Schäpe, M., 2000. Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation. In: Strobl, J., Blaschke, T., Griesebner, G. (Eds.), Angewandte Geographische Informations-Verarbeitung XII. Wichmann Verlag, Karlsruhe, pp. 12–23.

Beaulieu, J., Goldberg, M., 1989. Hierarchy in picture segmentation: a stepwise optimization approach. IEEE Transactions on Pattern Analysis and Machine Intelligence 11 (2), 150–163.

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., 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.

Blaschke, T., Strobl, J., 2001. What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. GIS-Zeitschrift für Geoinformations systeme 14 (6), 12–17.

Blaschke, T., 2010. Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing 65 (1), 2–16.

Carleer, A.P., Debeir, O., Wolff, E., 2005. Assessment of very high spatial resolution satellite image segmentations. Photogrammetric Engineering and Remote Sensing 71 (11), 1285–1294.

Castilla, G., Hay, G.J., Ruiz, J.R., 2008. Size-constrained region merging (SCRM): an automated delineation tool for assisted photointerpretation. Photogrammetric Engineering and Remote Sensing 74 (4), 409–419.

Câmara, G., Souza, R.C.M., Freitas, U.M., Garrido, J., 1996. Spring: integrating remote sensing and GIS by object-oriented data modeling. Computers and Graphics 20 (3), 395–403.

Chang, Y., Li, X., 1994. Adaptive image region-growing. IEEE Transactions on Image Processing 3 (6), 868–872.

Cheng, H.D., Sun, Y., 2000. A hierarchical approach to color image segmentation using homogeneity. IEEE Transactions on Image Processing 9 (12), 2071–2082.

Deng, Y., Manjunath, B.S., 2001. Unsupervised segmentation of color-texture regions in images and video. IEEE Transactions on Pattern Analysis and Machine Intelligence 23 (8), 800–810.

Fan, J., Yau, D.K.Y., Elmagarmid, A.K., Aref, W.G., 2001. Automatic image segmentation by integrating color-edge extraction and seeded region growing. IEEE Transactions on Image Processing 10 (10), 1454–1466.

Fan, J., Zeng, G., Body, M., Hacid, M., 2005. Seeded region growing: an extensive and comparative study. Pattern Recognition Litters 26 (8), 1139–1156.

Gonzalez, R.C., Woods, R.E., 2004. Digital image processing. Pearson Education.

Hall, O., Hay, G.J., Marceau, D.J., 2004. Multiscale object-specific analysis: Scale problems and multiscale solutions. In: Proc. 12th International Conference on Geoinformatices, University of Gävle, Sweden, 7–9 June, 2004.

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

Haris, K., Efstratiadis, S.N., Maglaveras, N., Katsaggelos, A.K., 1998. Hybrid image segmentation using watersheds and fast region merging. IEEE Transctions on Image Processing 7 (12), 1684–1699.

Hay, G.J., Marceau, D.J., Dube, P., Bouchard, A., 2001. A multiscale framework for landscape analysis: object-specific analysis and upscaling. Landscape Ecology 16 (6), 471–490.

Hay, G.J., Blaschke, T., Marceau, D.J., Bouchard, A., 2003. A comparison of three image-object methods for the multiscale analysis of landscape structure. ISPRS Journal of Photogrammetry and Remote Sensing 57 (5–6), 327–345.

Hay, G.J., Castilla, G., 2006. Object-based image analysis: strengths, weaknesses, opportunities and threats (SWOT). International archives of photogrammetry, Remote Sensing and Spatial Information Sciences, 36.

Hubert, L., Arabie, P., 1985. Comparing partitions. Journal of Classification 2 (1), 193–218.

Lhermitte, S., Verbesselt, J., Jonckheere, I., Nackaerts, K., Aardt, J.A.N., Verstraeten, W.W., Coppin, P., 2008. Hierarchical image segmentation based on similarity of NDVI time series. Remote Sensing of Environment 112 (2), 506–521.

Li, D., Zhang, G., Wu, Z., Yi, L., 2010. An edge embedded marker-based watershed algorithm for high spatial resolution remote sensing image segmentation. IEEE Transactions on Image Processing 19 (10), 2781–2787.

Lindeberg, T., 1994. Scale-space theory: a basic tool for analyzing structures at different scales. Journal of Applied Statistics 21 (2), 225–270.

Marpu, P.R., Neubert, M., Herold, H., Niemeyer, I., 2010. Enhanced evaluation of image segmentation results. Journal of Spatial Science 55 (1), 55–68.

Moghaddamzadeh, A., Bourbakis, N., 1997. A fuzzy region growing approach for segmentation of color images. Pattern Recognition 30 (6), 867–881.

Muñoz, X., Freixenet, J., Cufí, X., Martí, J., 2003. Strategies for image segmentation combining region and boundary information. Pattern Recognition Letters 24 (1–3), 375–392.

Pavlidis, T., Liow, Y., 1990. Integrating region growing and edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 12 (3), 225–233.

Rand, W.M., 1971. Objective criteria for the evaluation of clustering methods. Journal of the American Statistical Association 66 (336), 846–850.

Sarkar, A., Biswas, M.K., Sharma, K.M.S., 2000. A simple unsupervised MRF model based image segmentation approach. IEEE Transactions on Image Processing 9 (5), 801–812.

Shih, F.Y., Cheng, S., 2005. Automatic seeded region growing for color image segmentation. Image and Vision Computing 23 (10), 877–886.

Tabb, M., Ahuja, N., 1997. Multiscale image segmentation by integrated edge and region detection. IEEE Transactions on Image Processing 6 (5), 642–655.

Tremeau, A., Borel, N., 1997. A region growing and merging algorithm to color segmentation. Pattern Recognition 30 (7), 1191–1203.

Trias-Sanz, R., Stamon, G., Louchet, J., 2008. Using colour, texture, and hierarchical segmentation for high-resolution remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing 63 (2), 156–168.

Wang, Z., Jensen, J.R., Im, J., 2010. An automatic region-based image segmentation algorithm for remote sensing applications. Environmental Modelling and Software 25 (10), 1149–1165.

Yu, Q., Clausi, D.A., 2008. IRGS: Image segmentation using edge penalties and region growing. IEEE Transactions on Pattern Analysis and Machine Intelligence 30 (12), 2126–2139.

Zhang, Y., 2002. Problems in the fusion of commercial high-resolution satellite images as well as Landsat 7 images and initial solutions. International Archives of Photogrammetry, Remote Sensing and Spatial. Information Sciences 34 (4).

Zhu, S.C., Yuille, A., 1996. Region competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 18 (9), 884–900.