ISPRS Journal of Photogrammetry and Remote Sensing 98 (2014) 19–28

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Hybrid region merging method for segmentation of high-resolution remote sensing images

Xueliang Zhang [a,b,c], Pengfeng Xiao [a,b,c,*], Xuezhi Feng [a,b,c], Jiangeng Wang [d], Zuo Wang [a,b,c]

[a] Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Nanjing University, China
[b] Key Laboratory for Satellite Mapping Technology and Applications of State Administration of Surveying, Mapping and Geoinformation of China, Nanjing University, China
[c] Department of Geographic Information Science, Nanjing University, China
[d] School of Atmospheric Physics, Nanjing University of Information Science & Technology, China

ARTICLE INFO

Article history:
Received 10 November 2013
Received in revised form 16 September 2014
Accepted 22 September 2014

Keywords:
High-resolution remote sensing
Image segmentation
Region merging
Graph model
Object-based image analysis

ABSTRACT

Image segmentation remains a challenging problem for object-based image analysis. In this paper, a hybrid region merging (HRM) method is proposed to segment high-resolution remote sensing images. HRM integrates the advantages of global-oriented and local-oriented region merging strategies into a unified framework. The globally most-similar pair of regions is used to determine the starting point of a growing region, which provides an elegant way to avoid the problem of starting point assignment and to enhance the optimization ability for local-oriented region merging. During the region growing procedure, the merging iterations are constrained within the local vicinity, so that the segmentation is accelerated and can reflect the local context, as compared with the global-oriented method. A set of high-resolution remote sensing images is used to test the effectiveness of the HRM method, and three region-based remote sensing image segmentation methods are adopted for comparison, including the hierarchical stepwise optimization (HSWO) method, the local-mutual best region merging (LMM) method, and the multiresolution segmentation (MRS) method embedded in eCognition Developer software. Both the supervised evaluation and visual assessment show that HRM performs better than HSWO and LMM by combining both their advantages. The segmentation results of HRM and MRS are visually comparable, but HRM can describe objects as single regions better than MRS, and the supervised and unsupervised evaluation results further prove the superiority of HRM.

1. Introduction

With the thriving development of high spatial resolution sensors, an increasing number of metric and sub-metric resolution remote sensing images are currently available, which allows for accurate geometrical analysis of objects at fine scales. However, the high spatial resolution also raises new challenges. For a given application, even though more relevant information is available, it also comes with an increased amount of irrelevant information alongside (Benediktsson et al., 2012).

Image segmentation remains one of the challenges caused by high spatial resolution. In order to deal with the abundant information in high spatial resolution remote sensing (HR) images, image segmentation can be viewed as a simplification method, which partitions the HR image into a set of regions, providing objects for object-based image analysis (OBIA) (Blaschke, 2010). Compared with pixel-based classification, OBIA is less sensitive to the spectral variance within objects, and can make use of the object features and the spatial relations between objects (Hay and Castilla, 2006). OBIA can achieve higher accuracy than pixel-based classification (Myint et al., 2011), and it is becoming a new and evolving paradigm (Blaschke et al., 2014).

To describe the various objects in HR images, OBIA requires multiscale segmentation (Burnett and Blaschke, 2003; Hay et al., 2003), and it is still struggling with the selection of suitable segmentation scales for successive analysis (Hay et al., 2005). Usually, it is dependent on a trial-and-error strategy by visual assessment (Meinel and Neubert, 2004). In recent years, both the supervised (Carleer et al., 2005; Liu et al., 2012; Witharana and Civco, 2014) and the unsupervised (Drăguţ et al., 2010; Johnson and Xie, 2011; Chen et al., 2012; Zhang et al., 2012) segmentation evaluation methods were introduced to determine the optimal segmentation parameters. In terms of accuracy assessment, the units on which the map is based must be tested on their accuracy

* Corresponding author at: Department of Geographic Information Science, Nanjing University, China. Tel.: +86 25 89680612.
E-mail address: xiaopf@nju.edu.cn (P. Xiao).

http://dx.doi.org/10.1016/j.isprsjprs.2014.09.011
0924-2716/

(Congalton and Green, 2008). Because OBIA is based on the units of image objects, the object-based assessment is required to assess its accuracy (Albrecht, 2008), and the classic random point sampling method may not be sufficient for this purpose. In this paper, we do not focus on researching the accuracy assessment and parameter optimization methods, but propose a novel region-based segmentation method for HR images that provides multiscale segmentation results.

Among various image segmentation methods (Pal and Pal, 1993; Cheng et al., 2001; Yang and Kang, 2009; Dey et al., 2010), the region merging method is an effective method for remote sensing image segmentation. It can make use of the statistics inside a region and the differences between adjacent regions, producing final segments as spatially contiguous regions with closed boundaries, which are viewed as image objects to be analyzed in OBIA. Moreover, it is convenient to produce multiscale segmentations by setting different stopping rules. If few merging iterations are allowed, the mean size of segments is small, resulting in a fine-scale result. On the other hand, if more merging iterations are performed, the segmentation scale is getting coarser. Hence, region merging methods, such as recursive hierarchical segmentation (RHSeg) (Tilton et al., 2012), size-constrained region merging (SCRM) (Castilla et al., 2008), the multiresolution segmentation (MRS) method embedded in the commercial image analysis software eCognition Developer (Baatz and Schäpe, 2000; Benz et al., 2004), and the segmentation method in the object-based analysis system SPRING (Câmara et al., 1996), continue to be the state-of-the-art methods for remote sensing image segmentation.

Region growing (Adams and Bischof, 1994; Shih and Cheng, 2005) can be considered a special case of region merging, where the growing region and the candidate pixel at its boundary are viewed as two regions to be merged. Split-and-merge (Wuest and Zhang, 2009) can also be considered a special case of region merging, in which the splitting stage produces initial segmentation, and the merging stage can make use of different region merging strategies as well. In general, the region merging method is categorized into global- and local-oriented according to the merging strategy. The former searches for the globally most-similar pair of regions to be merged iteratively (Beaulieu and Goldberg, 1989), whereas the latter chooses the merging pair within a local vicinity (Câmara et al., 1996; Baatz and Schäpe, 2000; Wang et al., 2010).

The hierarchical stepwise optimization (HSWO) method (Beaulieu and Goldberg, 1989) has the strongest constraint for optimization problems among all the region merging strategies, and is widely used for segmentation of remote sensing images (Yu and Clausi, 2008; Gaetano et al., 2009; Li et al., 2010b; Carvalho et al., 2010; Zhang et al., 2014; Shui and Zhang, 2014) as well as natural images (Haris et al., 1998; Arbeláez et al., 2011). However, the hierarchical segmentation procedure is time-consuming because only a pair of regions is allowed to be merged per iteration, and it has to search for the merging pair among all pairs of adjacent regions in the entire scene. Hence, the improvement of hierarchical segmentation mainly focuses on two aspects. One is to accelerate the searching procedure by reducing the candidate pairs, or using the heap or priority queue (Beaulieu, 1990; Kurita, 1994; Haris et al., 1998; Gofman, 2006). The other is to allow multiple merges per iteration, such as in the multi-pass approach (Woodcock and Harward, 1992), and the recursive hierarchical segmentation method (Tilton et al., 2012). Moreover, HSWO focuses on achieving global optimization but neglects the influences from local context.

The local-oriented region merging method searches for the merging pair among only the neighbors of the growing region, and hence, it runs faster than the global-oriented method and can be influenced by the local structure. However, its optimization ability is not as strong as that of HSWO. In local-oriented region merging, the starting point of the growing region should be assigned as a pixel or an initial region. However, it is difficult to obtain proper starting points. In the case of seeded region growing, the seeds serve as the starting points, which are determined by several predefined rules (Adams and Bischof, 1994; Shih and Cheng, 2005). In the MRS method (Baatz and Schäpe, 2000), the distributed treatment order is used to determine the starting point by distributing the candidate pairs as far from each other as possible. Recently, the marker-based segmentation method (Tarabalka et al., 2010, 2012) was proposed, in which the marker is determined on the basis of classification result. The marker could also be viewed as the starting point. In any case, a solution to this problem is to choose the starting point in a random sequence.

In this paper, a hybrid region merging (HRM) method is proposed by combining the global- and local-oriented region merging methods. The local-mutual best region merging strategy is embedded in HRM, which accelerates the merging procedure more than the global-oriented method does and allows HRM to be constrained by local context. On the other hand, HRM provides an elegant way to assign the starting point of a growing region by choosing the globally most-similar pair, which also enhances the optimization ability more than the local-oriented method does. The organization of this paper is given as follows. Section 2 describes the HRM method, Section 3 presents the experimental results, and Section 4 presents the discussions and Section 5 concludes the paper.

2. Methodology

2.1. Overview

The flow diagram of HRM is presented in Fig. 1. First, the initial segmentation is performed to produce over-segmented regions. Second, the graph model (Zhang et al., 2014) is built based on initial segments, including the region adjacency graph (RAG) (Trémeau and Colantoni, 2000; Felzenszwalb and Huttenlocher, 2004) and the nearest neighbor graph (NNG) (Haris et al., 1998). Then, hybrid region merging (HRM) is applied on the graph model to produce the segmentation result. The stopping rule of HRM is defined as the threshold of the similarity between adjacent regions.

2.2. Graph model on initial segmentation

Regions and the adjacency between them are represented by nodes and arcs in the graph model, respectively, thus making it

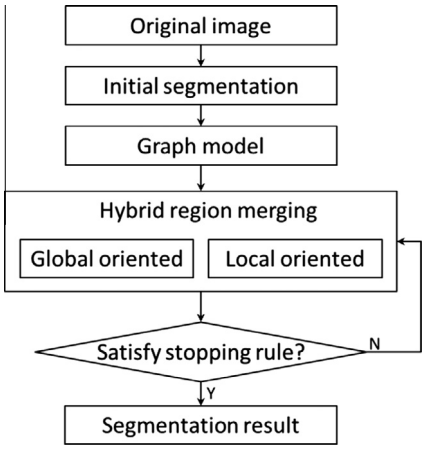

Fig. 1. Flow diagram of the proposed hybrid region merging method.

convenient to perform region merging on the graph model. Furthermore, the stochastic model, such as the Markov random model, can be defined on the graph, and region merging can be performed to lower the energy function of the stochastic model (Sarkar et al., 2000; Yu and Clausi, 2008).

In HRM, the graph model is built on over-segmented initial segments. Several methods, such as region growing, watershed (Vincent and Soille, 1991), and the mean-shift based method (Comaniciu and Meer, 2002), are qualified to produce initial segmentation. Here, we design a primary region growing method to produce initial segmentation. The starting point of a growing region is assigned as a pixel according to the scanning order, which is from top-left to bottom-right of the image. The similarity between the growing region and its adjacent pixel is evaluated by the spectral difference. Then, the stopping rule is defined as the thresholds of spectral difference and region size. If the region size or the spectral difference is larger than the threshold, the growing procedure stops for the current region and turns to next one.

Based on the initial segmentation, the undirected RAG, $G = (V, E)$, is defined, where V is the set of nodes $\{v_i\}$ and E is the set of arcs $\{e_{ij}\}$. A node v_i represents an initial segment, and an arc e_{ij} indicates the adjacency between node v_i and v_j. The weight of e_{ij} represents the similarity between v_i and v_j. Next, the directed NNG is built based on the RAG. The nodes of NNG are the same as those of RAG. Only one directed arc starts from each node, and the arc points to the most-similar adjacent node. Then, the cycle arc in NNG is defined where the arcs of two nodes point to each other coincidently. The globally most-similar pair of regions must belong to the pairs connected by cycle arcs. Supposing there are N nodes in NNG, the number of cycle arcs is $N/2$ in the worst case. Then, the number of candidate pairs is significantly reduced when searching for the globally best pair among cycle arcs. All the cycle arcs are recorded in a priority queue (pq) sorted by the arc weight, where the arc with the smallest weight is at the top of pq. For more details about NNG, please refer to Haris et al. (1998) and Zhang et al. (2014).

The similarity between adjacent regions is calculated and recorded as the arc weight. In order to improve the segmentation performance, various features, such as spectral homogeneity (Baatz and Schäpe, 2000), texture (Trias-Sanz et al., 2008; Ryherd and Woodcock, 1996; Hu et al., 2005), shape (Baatz and Schäpe, 2000; Shackelford and Davis, 2003; Zhang et al., 2005), edge (Yu and Clausi, 2008; Li et al., 2010a), and structure (Pesaresi and Benediktsson, 2001; Akçay and Aksoy, 2008), have been used for remote sensing image segmentation. In this paper, four features are selected to calculate the arc weight, including the region size (a), the change of standard deviation ($CStd$) and compactness ($CComp$) after a virtual merge, and the edge strength (ES). The definitions of the four features are similar to those in Zhang et al. (2013), but the features of $CStd$ and ES are further improved.

The feature of region size is the number of pixels in the region, which is directly related to segmentation scale. The feature of $CComp$ reflects the change of region compactness caused by region merging, which drives to generate compact segments. The definition of $CComp$ is as below:

$$CComp = L/\sqrt{a} - (a_1 L_1/\sqrt{a_1} + a_2 L_2/\sqrt{a_2})/(a_1 + a_2), \quad (1)$$

where L, L_1, and L_2 are the boundary lengths of the newly created region and two original regions, respectively, and a represents the region size.

The feature of $CStd$ reflects the change of region homogeneity caused by region merging, which drives to generate homogeneous segments. To combine the information from multiple spectral bands, $CStd$ is calculated as below:

$$CStd = \sum_{i=1}^{n} CStd_i \cdot std_i / SumStd, \quad (2)$$

where $CStd_i$ and std_i represent the change of standard deviation caused by merging and the standard deviation of band i for the newly created region, n is the number of bands, and $SumStd$ is the sum of standard deviation values over all bands.

The feature of ES is calculated in terms of the common boundary between adjacent regions, which is defined as the spectral difference within the neighborhood of the common boundary (Zhang et al., 2013). Similar to the spirit of graduated increased edge penalty (GIEP) proposed by Yu and Clausi (2008), the effectiveness of ES is increased along with the region merging procedure, where the effectiveness is weak at the initial merging iterations, and it gets stronger with the increase of the region size (Zhang et al., 2014). Then, the feature of ES is normalized as below:

$$g(ES) = \exp(-\varepsilon/ES), \quad (3)$$

where the normalized factor ε is assigned as the square root of the mean edge strength. Because the mean edge strength is increased as the merging procedure progresses, the effectiveness of ES would be enhanced to lower the arc weight.

Finally, the four features are combined to form the merging criterion (MC) as below:

$$MC = (a_1 + a_2)(\omega \cdot CStd + (1 - \omega)CComp)g(ES), \quad (4)$$

where a_1 and a_2 are the size of two regions, ω is the parameter of spectral weight, and its default value is 0.5. If ω is set as a large number, the region merging would concentrate more on generating homogeneous regions and neglect the compactness of the regions. On the other hand, if it is set as a small value, it would be driven to generate compact regions with less consideration of the homogeneity. In particular, when the weighted sum of $CStd$ and $CComp$ is negative in certain cases, the features of region size and edge strength are not integrated in MC. The arc weight is calculated according to MC. The small arc weight indicates that the adjacent regions connected by the arc have greater similarity.

2.3. Hybrid region merging

Both the global-oriented and local-oriented region merging are integrated in HRM. Specifically, HRM is a combination of the hierarchical stepwise optimization (HSWO) (Beaulieu and Goldberg, 1989) and the local-mutual best region merging (LMM) (Baatz and Schäpe, 2000). The stopping rule is defined as the threshold of the arc weight. If the threshold is set to a large value, more merging iterations are allowed, resulting in coarser-scale segmentation. The stopping rule is used to control both the global-oriented and the local-oriented parts in HRM. The starting point of a growing region is determined by the globally most-similar pair of regions. Taking one region in the pair as the growing region, LMM is performed. If the growing region and one of its neighbors are the best neighbors of each other, they are mutually best neighbors. Then, the mutually best neighbors are merged successively. If the mutually best neighbors do not exist or if the arc weight between two mutually best neighbors is larger than the threshold, the local-oriented region merging procedure of the current growing region stops and turns to the next growing region. Finally, the entire hybrid region merging process would stop if the smallest arc weight in the graph is larger than the threshold, producing the segmentation result. The algorithm of HRM based on the graph model is described in Table 1.

Global-oriented optimization in HRM is performed to determine the starting point of the growing region and the ending point of HRM. The globally most-similar pair of regions is viewed as the

Table 1
Algorithm of hybrid region merging (HRM).

Input: the graph model RAG and NNG, the priority queue (pq) recording all the cycle arcs. Parameters: The spectral weight ω and the threshold of arc weight T
Output: the segmented regions
(1) Obtain the arc e_{ij} pointed by the top of pq
(2) If the weight of e_{ij} is smaller than T, view node v_i or v_j as the growing node V_g and go to (3). Else, hybrid region merging is completed, and output all the remaining nodes $\{V_g\}$ as segmented regions
(3) Repeat local-mutual best region merging
(3-1) Find the cycle arc e_c pointed to V_g in NNG
If e_c does not exist or the weight of e_c is larger than T, go to (4)
Else obtain the other node V_n connected by e_c
(3-2) Merge V_n into V_g, and remove V_n from RAG and NNG. Recalculate V_g in RAG
(3-3) Update the adjacency and arc weight in RAG, and the arcs in NNG caused by merging
(4) Update pq by removing the changed cycle arcs and adding the newly emerged cycle arcs, and go to (1) for the next growing region

starting point to grow, which enhances the optimization ability compared with local-oriented region merging. On the other hand, local-oriented region merging is performed to make HRM constrained to local context, which would accelerate the merging procedure compared with global-oriented region merging. Moreover, since the local-oriented region merging is controlled by both the threshold of arc weight and the existence of a mutual best neighbor, when there is no mutually best neighbor for the growing region, the growing procedure stops even though the arc weight is smaller than the threshold. Therefore, the growing process is influenced by the local structure, and the final segmented region is not produced directly by growing from a single starting point but from several starting points successively.

The computational complexity of HRM is lower than that of HSWO, but higher than that of LMM when applied on the same graph model. The computational complexity mainly involves the update of the graph model and that of the priority queue. The complexity of updating the graph model is the same for HRM, HSWO, and LMM, so that the difference mainly results from the update of the priority queue. LMM does not need to record the cycle arcs in a priority queue, so that its computational complexity just depends on the local-oriented update of the graph model. Supposing the number of merging iterations is n, then the complexity of LMM is $O(hn)$, where h represents the complexity of updating the graph model at a merging iteration. Compared with LMM, HSWO has to update the priority queue at each merging iteration. Supposing the queue height is B ($B \leqslant N/2$ according to Subsection 2.2) and the number of times to update the priority queue at a merging iteration is β (Zhang et al., 2014), the computational complexity of HSWO is $O(hn + n\beta \log_2 B)$. On the other hand, HRM just updates the priority queue after several local-oriented merging iterations, so that its computational complexity is $O(hn + m\beta \log_2 B)$, where m is equal to n/k, and k represents the number of local-oriented merging iterations per hybrid merging iteration, as described in step 3 of Table 1. Therefore, the complexity of HRM is $m\beta \log_2 B$ higher than that of LMM and $(n - m)\ \beta \log_2 B$ lower than that of HSWO.

3. Experiment and performance evaluation

3.1. Data and evaluation methods

In the experiment, a set of HR images, as shown in Table 2, is used to test the HRM method, including QuickBird, WorldView, and aerial images. The multi-spectral bands of T1 and T2 are sharpened to 0.6 m by the method proposed by Zhang (2002). The proposed method is compared with other state-of-the-art region merging methods for remote sensing image segmentation according to both supervised and unsupervised evaluation methods and visual assessment. Finally, the segmentation time is presented.

The supervised evaluation indicators include the rightly segmented ratio (*RR*) (Carleer et al., 2005; Zhang et al., 2013), the adjusted Rand index (*ARI*) (Hubert and Arabie, 1985), and the symmetric partition distance (D_{sym}) (Cardoso and Corte-Real, 2005). The indicator of *RR* is calculated based on the number of accurately segmented pixels compared with the reference segmentation, indicating the segmentation accuracy. The indicator of *ARI* is the measure of correspondence between the segmented result and the reference segmentation. The indicator of D_{sym} represents the proportion of pixels that should be removed to make the segmentation result identical to the reference. The higher *RR* or *ARI* value or lower D_{sym} value indicates the better segmentation. In this study, the supervised evaluation is performed in terms of the test image T1. The objects in T1 are very complex. For example, the buildings include homogeneous factories with both different sizes and colors, and other heterogeneous roofs with various colors and textures. Hence, it would be difficult to segment T1 automatically. The reference segmentation of T1 is delineated by a remote sensing expert, who is asked to separate different classes of objects from each other. The produced reference segmentation is shown in Fig. 2, which contains 165 reference segments.

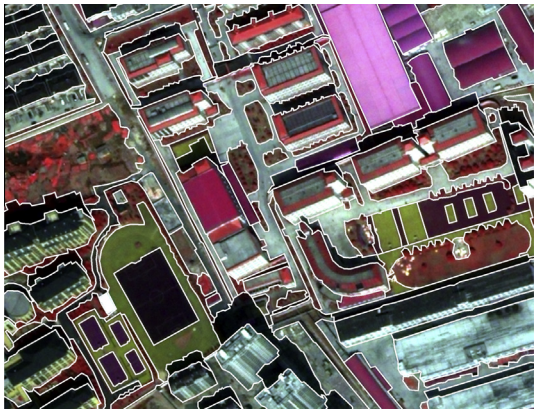

Fig. 2. Reference segmentation of test image T1. There are 165 regions.

Table 2
The list of test images.

Image	Platform	Size (pixel)	Spatial resolution	Shown with color compositing	Landscape
T1	QuickBird	658 × 504	0.6 m	NIR, R, G	Urban area
T2	WorldView	439 × 412	0.5 m	R, G, B	Factory
T3	Aerial	531 × 352	0.5 m	R, G, B	Urban residential area
T4	Aerial	651 × 589	0.5 m	R, G, B	Rural residential area
T5	Aerial	1000 × 800	0.5 m	R, G, B	Rural area
T6	Aerial	701 × 383	0.5 m	R, G, B	Forest area

Table 3
Supervised evaluation results of multiresolution segmentations of test image T1 produced by different methods. *RR*, *ARI* and D_{sym} represent rightly-segmented ratio, adjusted Rand index, and symmetric partition distance, respectively.

Number of regions	HRM			HSWO			LMM			MRS		
	RR	*ARI*	D_{sym}	*RR*	*ARI*	D_{sym}	*RR*	*ARI*	D_{sym}	*RR*	*ARI*	D_{sym}
730	0.914	0.296	0.615	0.911	0.290	0.631	0.910	0.292	0.623	0.915	0.249	0.663
515	0.903	0.397	0.549	0.894	0.353	0.568	0.897	0.354	0.563	0.906	0.305	0.603
360	0.882	0.449	0.496	0.875	0.443	0.509	0.872	0.439	0.516	0.879	0.404	0.518
195	0.804	0.512	0.456	0.797	0.499	0.470	0.791	0.521	0.450	0.799	0.501	0.472

Moreover, three unsupervised evaluation indicators, including the indicator *Z* proposed by Zhang et al. (2012), *SU* by Corcoran et al. (2010), and *E* by Zhang et al. (2004), are adopted to further assess the segmentation accuracy. The indicator *Z* is the sum of two parts. One indicates the homogeneity within regions and the other reflects heterogeneity among regions. The indicator *E* is an entropy-based measure, which is also the sum of two parts. One is the expected region entropy to evaluate the region uniformity, and the other is the layout entropy that is increased with the number of regions. The indicator *SU*, which is the ratio of the separation part to the cohesion part, can take the spatial properties between regions into consideration. The separation part reflects the inter-region heterogeneity by measuring the contrast between adjacent regions, and the cohesion part corresponds to the intra-region homogeneity criterion. Generally, the higher *SU* value or lower *Z* or *E* value indicates the higher segmentation quality.

3.2. Comparison of hybrid region merging with other segmentation methods

Based on the same initial segmentation and graph model in Subsection 2.2, the HRM, HSWO (Beaulieu and Goldberg, 1989), and LMM (Baatz and Schäpe, 2000) methods are applied for comparison. Then, the segmentation difference only results from the merging strategies. When producing the initial segmentation, the threshold of spectral difference and region size is set as 10 and 100 for different test images, respectively. The thresholds would not be the best for all the images, but they can help to produce over-segmented results with high accuracy in the experiment. Moreover, the multiresolution segmentation (MRS) method embedded in the commercial software eCognition Developer is used for comparison. The segmentation method MRS also adopts the local-mutual best region merging strategy (Baatz and Schäpe, 2000), but it uses the distributed treatment order to assign the starting point for the growing region, where the authors defined two conditions of selecting a sequence of starting points and provided a solution of deriving the sequence from a dither matrix produced by a binary counter. However, LMM determines the starting point according to the scanning order in the graph. Compared with the merging criterion of MRS, the feature of edge strength is added, but the shape factor of smoothness is not used in this study. The spectral weight of the merging criterion of MRS and that in this study are both set as 0.5. Because a detailed discussion of the spectral weight is beyond the scope of this paper, we do not show the segmentations produced by setting different spectral weights.

The supervised evaluation results of four multiresolution segmentations of T1 produced by HRM, HSWO, LMM, and MRS are presented in Table 3. The number of regions is used to indicate the segmentation scale, which means that the segmentation with more regions is at the coarser scale, and vice versa. The performance of HSWO and LMM are similar in terms of the three indicators. Comparing HRM with HSWO and LMM, the *RR* value of HRM results is approximately 1% higher than that of HSWO and LMM, and the *ARI* value of HRM results is also higher than that of HSWO and LMM results. Accordingly, the HRM results achieve the lower D_{sym} value than both the HSWO and LMM results. Comparing HSWO and LMM with MRS, the *RR* value of MRS results is about 1% higher than that of HSWO and LMM results. However, the *ARI* and D_{sym} values are significantly higher and lower than those of MRS results, respectively. Furthermore, comparing HRM with MRS, the *RR* value is almost the same, and the difference of the *ARI* and D_{sym} values between the HRM and MRS results is even larger.

The unsupervised evaluation results of the multi-resolution segmentations of T1 are shown in Table 4. The three indicators are not sensitive to the segmentation difference of HRM, HSWO, and LMM. Especially, the *E* and *SU* values for the three methods are almost the same. The performance of HRM is better than MRS according to the three indicators. In terms of all the four segmentation results, the *Z* and *E* values of MRS are approximately 50 and 0.05 higher than those of HRM, respectively, and the *SU* values are approximately 0.01 lower.

The segmentation results of T1 produced by the four methods are presented in Fig. 3. We choose the scale according to visual assessment and the following criteria: (1) the scale is neither too coarse nor too fine, (2) different objects are distinguished from each other, and (3) there are few apparently wrong segments. In Fig. 3, most objects can be separated from each other. The general segmentation patterns of HRM, HSWO, and LMM results are very close to each other. The difference between the HRM and MRS results is also not visually significant.

Then, in Fig. 4, three subsets in the four segmentation results are zoomed in to show the difference. In the first row, where there is mainly a factory with shading roof, only the HRM method can segment it as a single region neglecting the influence of the shading. Compared with HSWO, the difference results from integrating local-oriented merging strategy in HRM, which is influenced by local context. On the other hand, when compared with LMM and MRS results, the difference shows the effectiveness of determining the starting point of the growing region by the globally best pair.

Table 4
Unsupervised evaluation results of multiresolution segmentations of test image T1 produced by different methods.

Number of regions	HRM			HSWO			LMM			MRS		
	Z	*E*	*SU*	*Z*	*E*	*SU*	*Z*	*E*	*SU*	*Z*	*E*	*SU*
730	1668.41	4.45	0.11	1653.25	4.46	0.12	1654.13	4.46	0.11	1720.75	4.51	0.10
515	1723.20	4.36	0.10	1713.36	4.38	0.10	1724.02	4.37	0.10	1779.78	4.43	0.08
360	1835.02	4.28	0.09	1802.74	4.28	0.09	1791.47	4.28	0.09	1861.04	4.32	0.08
195	2108.09	4.11	0.07	2084.12	4.11	0.07	2142.39	4.10	0.07	2146.65	4.16	0.06

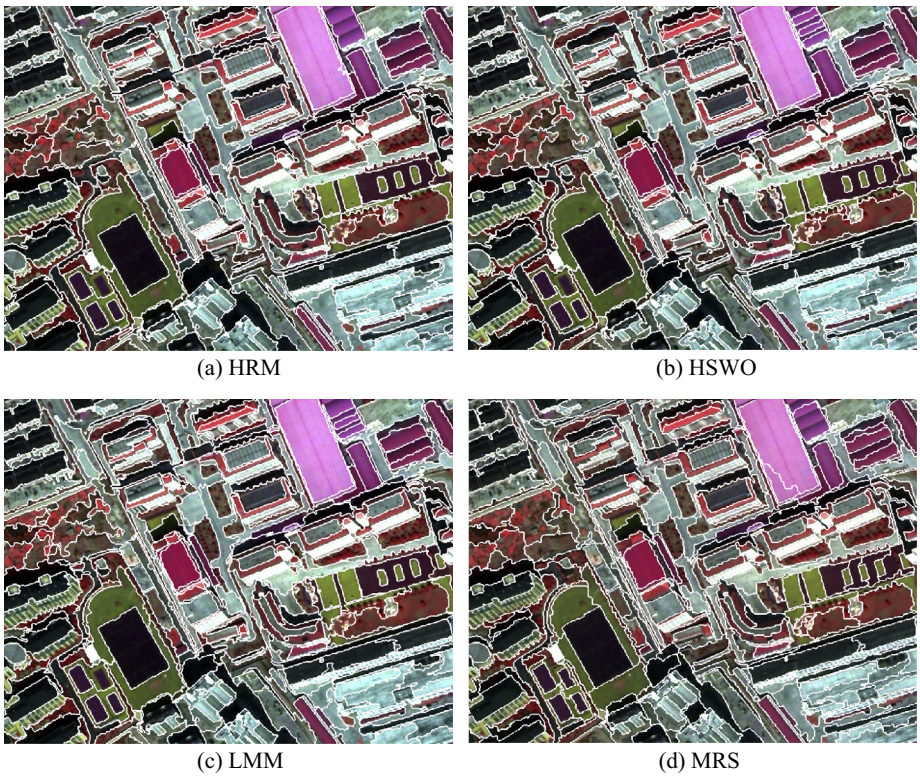

Fig. 3. Segmentation results produced by different methods for test image T1. The number of regions is 360.

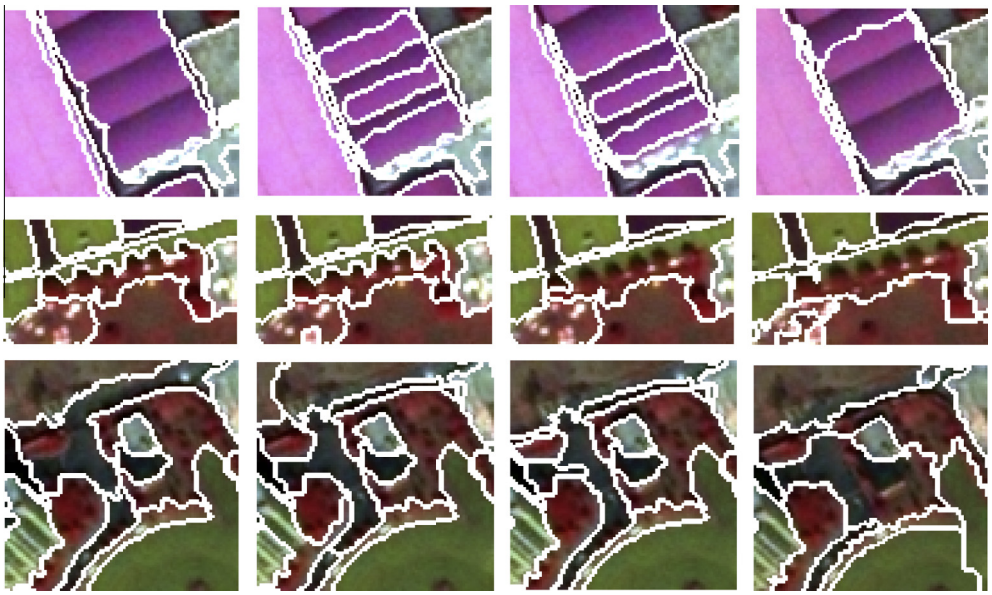

Fig. 4. Subsets of segmentations in Fig. 3, which are in the HRM, HSWO, LMM, and MRS results from left to right.

The distributed treatment order strategy in MRS also works, which results in only two regions. However, it is still not better than HRM in terms of the shading roof. In the second row, the trees are separated from other objects in the entirety by HRM, and the HSWO method still omits one tree in the bottom-right part. Moreover, the trees are wrongly merged with other objects in the LMM and MRS results, but the lower boundary is spatially accurate. The difference shows the effectiveness of the combination of HSWO and LMM into the HRM framework, which makes use of the advantages of both the HSWO and LMM methods. Then, the third row further shows the advantage of the HRM method. The road in this subset is segmented out by HRM, whereas the other three methods cannot distinguish it very well.

Furthermore, the test images T2–T6 with other landscape are used to show the effectiveness of the HRM method. The test images T2–T6 represent the landscape of a factory, an urban residential area, a rural residential area, a rural area and a forest area, respectively. In this part, we select one segmentation scale for each test image, and the scale selection criteria are same as those for T1 in Fig. 3. The segmentation results are shown in Fig. 5 for visual

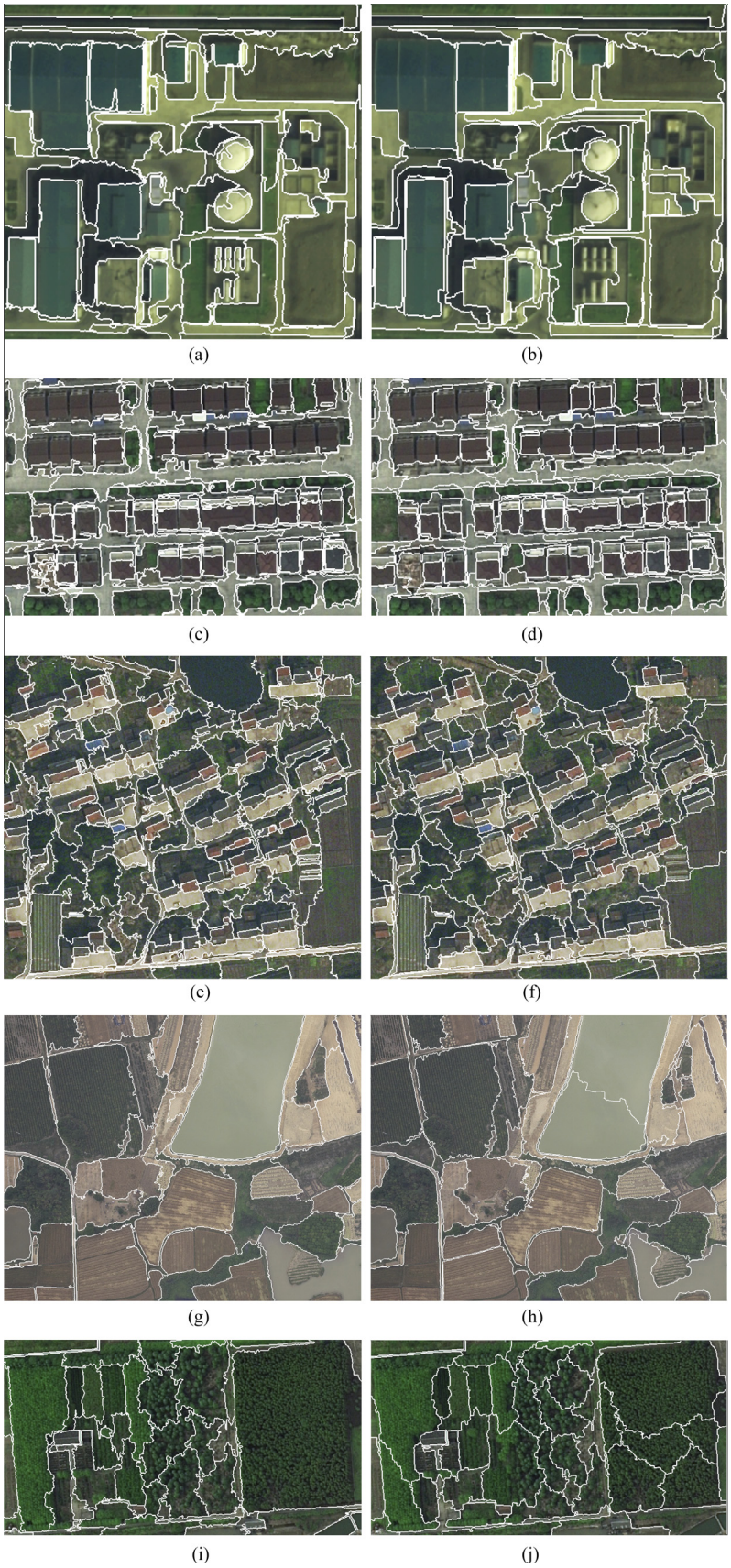

Fig. 5. Segmentation results produced by HRM (left) and MRS (right) for test images T2–T6. The number of regions is 64, 157, 200, 50, and 50 for T2, T3, T4, T5, and T6, respectively.

Table 5
Unsupervised evaluation results of segmentations of test images T2–T6 produced by different methods. The segmentation results of HRM and MRS are shown in Fig. 5.

Test image	HRM			HSWO			LMM			MRS		
	Z	*E*	*SU*	*Z*	*E*	*SU*	*Z*	*E*	*SU*	*Z*	*E*	*SU*
T2	567.59	3.27	0.18	551.66	3.27	0.19	596.04	3.24	0.19	542.87	3.33	0.13
T3	475.74	3.67	0.15	487.78	3.67	0.15	474.03	3.68	0.15	446.23	3.73	0.13
T4	482.38	3.85	0.15	479.96	3.84	0.15	475.27	3.85	0.15	492.33	3.91	0.13
T5	357.71	3.14	0.19	360.88	3.16	0.21	347.00	3.14	0.19	370.14	3.19	0.17
T6	351.41	2.95	0.13	365.33	2.95	0.13	369.70	2.97	0.12	247.56	3.18	0.07

Table 6
Segmentation time of different methods.

Test image	Image size (pixel)	Number of initial segments	Segmentation time (s)			
			HRM	HSWO	LMM	MRS
Aerial image	2000 × 2000	309,035	39.5	57	27.7	29
QuickBird	2453 × 2375	503,795	88.5	140	48.0	45
QuickBird	2781 × 2492	421,552	77.9	124	53.3	58

comparison, where only the HRM and MRS results are presented. The unsupervised evaluation results for HRM, HSWO, LMM and MRS are presented in Table 5. Similar to Table 4, the three unsupervised evaluation indicators are not sensitive to HRM, HSWO, or LMM in Table 5. Comparing HRM and MRS, the indicators *E* and *SU* consistently show that HRM performs better than MRS in terms of all five test images. Generally, the HRM and MRS results are comparable in Fig. 5, and most objects can be distinguished from others. However, MRS tends to produce segments with similar region size, whereas HRM can produce segments representing single objects better with less constraint of the region size. The difference is mainly caused by the adaptive edge strength in HRM. For example, both the large roofs and the small roofs are segmented as single objects in Fig. 5(a) by HRM, whereas the small roofs are merged with each other to form large objects with similar size to other objects in Fig. 5(b) by MRS. On the other hand, the large pool in Fig. 5(g) and the group of trees at the right side of Fig. 5(i) are segmented as single objects by HRM, whereas in Fig. 5(h) and (j), they are separated into several small segments with similar size to other regions.

3.3. Segmentation time

The HRM, HSWO, and LMM are implemented by C# program, and MRS is implemented by eCognition Developer Trial 8.0. The segmentation time performed on a laptop computer with CPU of 2.6 GHz is listed in Table 6. Three large test images are used to calculate the segmentation time. In accordance with the computational complexity analyzed in 2.3, the segmentation time of HRM is larger than that of LMM, but it is significantly reduced compared with that of HSWO. The segmentation time of HRM and HSWO is directly influenced by the number of initial segments since the two methods have to search the globally most-similar adjacent pair within the entire scene. The segmentation time of LMM is similar to that of MRS because the region merging strategy is the same.

4. Discussions

The hybrid region merging (HRM) method combines hierarchical stepwise optimization (HSWO) and local-mutual best region merging (LMM) into a unified framework. Actually, it is a unified framework that can combine HSWO with other local-oriented region merging methods. HRM makes use of the advantages of both HSWO and LMM. Compared with HSWO, HRM is significantly accelerated by allowing several local-oriented merging iterations, and the merging procedure of HRM is constrained by the local context rather than just focusing on achieving the global optimization by HSWO. Compared with LMM, HRM assigns the globally most-similar pair as the starting point for the growing region, which provides a solution to determining the proper starting point for local-oriented region merging. Moreover, it helps to enhance the optimization ability compared with local-oriented region merging.

The stopping rule of HRM is the threshold of the arc weight in the graph model. The HRM method is able to produce multiscale results by setting different thresholds of the arc weight. When the threshold is set larger, more merging iterations are allowed, and the segmented regions are coarser. However, HRM is not certain to produce nested multiscale segments just by setting different thresholds directly because the local-oriented region merging is integrated. Then, the stepwise scale parameter (Zhang et al., 2013) can be directly applied to control the merging procedure to produce nested multiscale segments. In the future, we will focus on expanding HRM to produce multiscale segmentations, and choosing the proper scale for a given application automatically. Then, building the correspondence between the scale parameter and the semantic meaning of objects will be necessary. In this case, the object segmentation strategy (Borenstein and Ullman, 2008) combining the bottom-up and top-down segmentation would be a proper solution.

According to the supervised evaluation results, all the three indicators show that HRM performs better than both HSWO and LMM. Since the merging procedures are applied on the same graph model, it proves the advantage of the combination of global- and local-oriented region merging. The *ARI* and D_{sym} differences between the HRM and MRS results are even larger, which further show the superiority of the combination strategy in HRM method. The visual difference among the HRM, HSWO, and LMM results is not significant, and the differences are mainly reflected from local areas as shown in Fig. 4. This may be the reason that the unsupervised evaluation results for the three methods are similar. Then, the unsupervised evaluation results for test images T2–T6 also show that HRM can achieve better performance than MRS, especially according to the indicators of *E* and *SU*. The visual assessment for T2–T6 show that HRM can produce segments representing single objects better than MRS, and the difference is mainly caused by the adaptive edge strength in HRM (Zhang et al., 2014).

The computational complexity of HRM is higher than that of LMM because HRM has to update the priority queue for the entire image, whereas lower than that of HSWO because HRM does not

need to update the priority queue at each merging iteration but just after several local-oriented merging iterations. Then, the complexity of HRM would depend on the number of local-oriented merging iterations within each hybrid merging iteration. Moreover, since HRM, LMM, and HSWO are performed on the graph model, the computational complexity of the three methods all relates to the number of nodes in the graph, or the number of initial segments. Hence, in Table 6, even though the size of the first QuickBird image is smaller than the second one, it takes more time to segment the first one because it has more initial segments.

5. Conclusions

The hybrid region merging (HRM) method was proposed to segment high-resolution remote sensing images. First, the over-segmented initial segmentation is produced by a region growing method, and then the graph model, including both RAG and NNG, is defined on the basis of the initial segmentation. Finally, HRM is performed on the graph model to produce segmentation results. HRM is a unified framework to combine the global- and local-oriented region merging. Specifically, it is the combination of hierarchical stepwise optimization (HSWO) and local-mutual best region merging (LMM) in this study. In order to show the effectiveness of HRM by making use of the advantages of both HSWO and LMM, a set of high spatial resolution remote sensing images is used to perform the experiments. Three supervised evaluation indicators and three unsupervised evaluation indicators are selected to assess the segmentation quality and to compare HRM with other segmentation methods, including HSWO, LMM and the multi-resolution segmentation (MRS) method embedded in eCognition Developer 8. In terms of the test images, the supervised evaluation results show that HRM can perform better than the other three segmentation methods, and the unsupervised evaluation further prove the superiority of HRM. In the future, we will pay attention to the combination of HSWO with other local-oriented region merging methods and address multiscale segmentation solutions for HRM.

Acknowledgements

This work was supported by the National Basic Research Program of China (Grant No. 2011CB952001), and the Project Funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD). The authors would like to acknowledge the insightful suggestions and comments from the anonymous reviewers.

References

Adams, R., Bischof, L., 1994. Seeded region growing. IEEE Trans. Pattern Anal. Mach. Intell. 16 (6), 641–647.
Akçay, H.G., Aksoy, S., 2008. Automatic detection of geospatial objects using multiple hierarchical segmentations. IEEE Trans. Geosci. Remote Sens. 46 (7), 2097–2111.
Albrecht, F., 2008. Assessing the spatial accuracy of object-based image classifications. In: Geospatial Crossroads @ GI_Forum'08. Proceedings of the Geoinformatics Forum Salzburg. Wichmann Verlag, Heidelberg, pp. 11–20.
Arbeláez, P., Maire, M., Fowlkes, C., Malik, J., 2011. Contour detection and hierarchical image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 33 (5), 898–916.
Baatz, M., Schäpe, A., 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.M., 1990. Versatile and efficient hierarchical clustering for picture segmentation. In IGARSS, 1663.
Beaulieu, J.M., Goldberg, M., 1989. Hierarchy in picture segmentation: a stepwise optimization approach. IEEE Trans. Pattern Anal. Mach. Intell. 11 (2), 150–163.
Benediktsson, J.A., Chanussot, J., Moon, W.M., 2012. Very high-resolution remote sensing: challenges and opportunities. Proc. IEEE 100 (6), 1907–1910.
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 J. Photogramm. Remote Sens. 58 (3–4), 239–258.
Blaschke, T., 2010. Object based image analysis for remote sensing. ISPRS J. Photogramm. Remote Sens. 65 (1), 2–16.
Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E., Feitosa, R.Q., van der Meer, F., van der Werff, H., van Coillie, F., Tiede, D., 2014. Geographic object-based image analysis-towards a new paradigm. ISPRS J. Photogramm. Remote Sens. 87, 180–191.
Borenstein, E., Ullman, S., 2008. Combined top–down/bottom-up segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 30 (12), 2109–2125.
Burnett, C., Blaschke, T., 2003. A multi-scale segmentation/object relationship modelling methodology for landscape analysis. Ecol. Modell. 168 (3), 233–249.
Câmara, G., Souza, R.C.M., Freitas, U.M., Garrido, J., 1996. Spring: integrating remote sensing and GIS by object-oriented data modelling. Comput. Graphics 20 (3), 395–403.
Cardoso, J.S., Corte-Real, L., 2005. Toward a generic evaluation of image segmentation. IEEE Trans. Image Process. 14 (11), 1773–1782.
Carleer, A.P., Debeir, O., Wolff, E., 2005. Assessment of very high spatial resolution satellite image segmentations. Photogramm. Eng. Remote Sens. 71 (11), 1285–1294.
Carvalho, E.A., Ushizima, D.M., Medeiros, F.N.S., Martins, C.I.O., Marques, R.C.P., Oliveira, I.N.S., 2010. SAR imagery segmentation by statistical region growing and hierarchical merging. Digital Signal Process. 20 (5), 1365–1378.
Castilla, G., Hay, G.J., Ruiz, J.R., 2008. Size-constrained region merging (SCRM): an automated delineation tool for assisted photo interpretation. Photogramm. Eng. Remote Sens. 74 (4), 409–419.
Chen, J., Li, J., Pan, D., Zhu, Q., Mao, Z., 2012. Edge-guided multiscale segmentation of satellite multispectral imagery. IEEE Trans. Geosci. Remote Sens. 50 (11), 4513–4520.
Cheng, H.D., Jiang, X.H., Sun, Y., Wang, J.L., 2001. Color image segmentation: advances and prospects. Pattern Recognit. 34 (12), 2259–2281.
Comaniciu, D., Meer, P., 2002. Mean shift: a robust approach toward feature space analysis. IEEE Trans. Pattern Anal. Mach. Intell. 24 (5), 603–619.
Congalton, R.G., Green, K., 2008. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. CRC Press.
Corcoran, P., Winstanley, A., Mooney, P., 2010. Segmentation performance evaluation for object-based remotely sensed image analysis. Int. J. Remote Sens. 31 (3), 617–645.
Dey, V., Zhang, Y., Zhong, M., 2010. A review on image segmentation techniques with remote sensing perspective. In: Wagner, W., Székely, B. (Eds.), ISPRS TC VII Symposium – 100 Years ISPRS, XXXVIII (7A), Vienna, Austria, IAPRS, 2010, pp. 31–42.
Drăguț, L., Tiede, D., Levick, S.R., 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. Int. J. Geogr. Inform. Sci. 24 (6), 859–871.
Felzenszwalb, P.F., Huttenlocher, D.P., 2004. Efficient graph-based image segmentation. Int. J. Comput. Vision 59 (2), 167–181.
Gaetano, R., Scarpa, G., Poggi, G., 2009. Hierarchical texture-based segmentation of multiresolution remote-sensing images. IEEE Trans. Geosci. Remote Sens. 47 (7), 2129–2141.
Gofman, E., 2006. Developing an efficient region growing engine for image segmentation. IEEE Int. Conf. Image Process., 2413–2416.
Haris, K., Efstradiadis, S., Maglaveras, N., Katsaggelos, A., 1998. Hybrid image segmentation using watershed and fast region merging. IEEE Trans. Image Process. 7 (12), 1684–1699.
Hay, G.J., Castilla, G., 2006. Object-based image analysis: strengths, weaknesses, opportunities and threats (SWOT). In: Proceeding the 1st International Conference of OBIA, 2006. <www.commission4.isprs.org/obia06/Papers/01_Opening%20Session/OBIA2006_Hay_Castilla.pdf>.
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 J. Photogramm. Remote Sens. 57 (5), 327–345.
Hay, G.J., Castilla, G., Wulder, M.A., Ruiz, J.R., 2005. An automated object-based approach for the multiscale image segmentation of forest scenes. Int. J. Appl. Earth Observation Geoinform. 7 (4), 339–359.
Hu, X., Tao, C.V., Prenzel, B., 2005. Automatic segmentation of high-resolution satellite imagery by integrating texture, intensity and color features. Photogramm. Eng. Remote Sens. 71 (12), 1399–1406.
Hubert, L., Arabie, P., 1985. Comparing partitions. J. Classif. 2 (1), 193–218.
Johnson, B., Xie, Z., 2011. Unsupervised image segmentation evaluation and refinement using a multi-scale approach. ISPRS J. Photogramm. Remote Sens. 66 (4), 473–483.
Kurita, T., 1994. An efficient agglomerative clustering algorithm for region growing. In: Proceeding of IAPR Workshop on Machine Vision Applications, pp. 210-213.
Li, D., Zhang, G., Wu, Z., Yi, L., 2010a. An edge embedded marker-based watershed algorithm for high spatial resolution remote sensing image segmentation. IEEE Trans. Image Process. 19 (10), 2781–2787.
Li, N., Huo, H., Fang, T., 2010b. A novel texture-preceded segmentation algorithm for high-resolution imagery. IEEE Trans. Geosci. Remote Sens. 48 (7), 2818–2829.
Liu, Y., Biana, L., Menga, Y., Wanga, H., Zhanga, S., Yanga, Y., Shaoa, X., Wang, B., 2012. Discrepancy measures for selecting optimal combination of parameter values in object-based image analysis. ISPRS J. Photogramm. Remote Sens. 68, 144–156.

Meinel, G., Neubert, M., 2004. A comparison of segmentation programs for high resolution remote sensing data. Int. Arch. Photogramm. Remote Sens. 35 (Part B), 1097–1105.
Myint, S.W., Goner, P., Brazel, A., Grossman-Clarke, S., Weng, Q., 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. Remote Sens. Environ. 115 (5), 1145–1161.
Pal, N.R., Pal, S.K., 1993. A review on image segmentation techniques. Pattern Recognit. 26 (9), 1277–1294.
Pesaresi, M., Benediktsson, J.A., 2001. A new approach for the morphological segmentation of high-resolution satellite imagery. IEEE Trans. Geosci. Remote Sens. 39 (2), 309–320.
Ryherd, S., Woodcock, C., 1996. Combining spectral and texture data in the segmentation of remotely sensed images. Photogramm. Eng. Remote Sens. 62 (2), 181–194.
Sarkar, A., Biswas, M.K., Sharma, K.M.S., 2000. A simple unsupervised MRF model based image segmentation approach. IEEE Trans. Image Process. 3 (5), 801–812.
Shackelford, A.K., Davis, C.H., 2003. A hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas. IEEE Trans. Geosci. Remote Sens. 41 (9), 1920–1932.
Shih, F.Y., Cheng, S., 2005. Automatic seeded region growing for color image segmentation. Image Vision Comput. 23 (10), 877–886.
Shui, P.L., Zhang, Z.J., 2014. Fast SAR image segmentation via merging cost with relative common boundary length penalty. IEEE Trans. Geosci. Remote Sens.. http://dx.doi.org/10.1109/TGRS.2013.2296561.
Tarabalka, Y., Benediktsson, J.A., Chanussot, J., Tilton, J.C., 2010. Multiple spectral–spatial classification approach for hyperspectral data. IEEE Trans. Geosci. Remote Sens. 48 (11), 4122–4132.
Tarabalka, Y., Tilton, J.C., Benediktsson, J.A., Chanussot, J., 2012. A marker-based approach for the automated selection of a single segmentation from a hierarchical set of image segmentations. IEEE J. Sel. Top. Appl. Earth Observations Remote Sens. 5 (1), 262–272.
Tilton, J.C., Tarabalka, Y., Montesano, P.M., Gofman, E., 2012. Best merge region-growing segmentation with integrated nonadjacent region object aggregation. IEEE Trans. Geosci. Remote Sens. 50 (11), 4454–4467.
Trémeau, A., Colantoni, P., 2000. Region adjacency graph applied to color image segmentation. IEEE Trans. Image Process. 9 (4), 735–744.
Trias-Sanz, R., Stamon, G., Louchet, J., 2008. Using colour, texture, and hierarchical segmentation for high-resolution remote sensing. ISPRS J. Photogramm. Remote Sens. 63 (2), 156–168.
Vincent, L., Soille, P., 1991. Watershed in digital spaces: an efficient algorithm based on immersion simulations. IEEE Trans. Pattern Anal. Mach. Intell. 13 (6), 583–598.
Wang, Z., Jensen, J.R., Im, J., 2010. An automatic region-based image segmentation algorithm for remote sensing applications. Environ. Modell. Softw. 25 (10), 1149–1165.
Witharana, C., Civco, D.L., 2014. Optimizing multi-resolution segmentation scale using empirical methods: exploring the sensitivity of the supervised discrepancy measure Euclidean distance 2 (ED2). ISPRS J. Photogramm. Remote Sens. 87, 108–121.
Woodcock, C.E., Harward, V.J., 1992. Nested-hierarchical scene models and image segmentation. Int. J. Remote Sens. 13 (16), 3167–3187.
Wuest, B., Zhang, Y., 2009. Region based segmentation of QuickBird multispectral imagery through band ratios and fuzzy comparison. ISPRS J. Photogramm. Remote Sens. 64 (1), 55–64.
Yang, Q.Q., Kang, W.X., 2009. General research on image segmentation algorithms. Int. J. Image, Graphics Signal Process. 1 (1), 1–8.
Yu, Q., Clausi, D.A., 2008. IRGS: image segmentation using edge penalties and region growing. IEEE Trans. Pattern Anal. Mach. Intell. 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. Int. Arch. Photogramm. Remote Sens. 34 (Part 4).
Zhang, H., Fritts, J.E., Goldman, S.A., 2004. An entropy-based objective evaluation method for image segmentation. Proc. SPIE-Storage Retrieval Methods Appl. Multimedia.
Zhang, L., Huang, X., Huang, B., Li, P., 2005. A pixel shape index coupled with spectral information for classification of high spatial resolution remotely sensed imagery. IEEE Trans. Geosci. Remote Sens. 44 (10), 2950–2961.
Zhang, X., Xiao, P., Feng, X., 2012. An unsupervised evaluation method for remotely sensed imagery segmentation. IEEE Geosci. Remote Sens. Lett. 9 (2), 156–160.
Zhang, X., Xiao, P., Song, X., She, J., 2013. Boundary-constrained multi-scale segmentation method for remote sensing images. ISPRS J. Photogramm. Remote Sens. 78, 15–25.
Zhang, X., Xiao, P., Feng, X., 2014. Fast hierarchical segmentation of high-resolution remote sensing image with adaptive edge penalty. Photogramm. Eng. Remote Sens. 80 (1), 71–80.