# Multi-scale Segmentation of High-Spatial Resolution Remote Sensing Images Using Adaptively Increased Scale Parameter

**Xueliang Zhang, Xuezhi Feng, and Pengfeng Xiao**

## Abstract

*The adaptively increased scale parameter (AISP) strategy is proposed to control multi-scale segmentation based on region growing methods. AISP strategy contains a set of gradually increased scale parameters to produce nested multi-scale segments. Instead of independently assigning the set of scale parameters ahead of segmentation, the contribution of this study is to dynamically determine scale parameters during segmentation procedure, making scale parameters adaptive to specific images and cover meaningful segmentation scales. Furthermore, the effectiveness of gradually increased scale parameters on segmentation accuracy is analyzed, which gives a thorough understanding to local-oriented region growing methods. The experimental results on a set of high-resolution images proved the effectiveness of AISP on controlling multi-scale segmentation. AISP holds the application potential for object-based analysis of high-resolution images.*

## Introduction

High-spatial resolution remote sensing images present geographic objects in detail, which allows an accurate geometrical analysis (Benediktsson *et al.*, 2012). When dealing with high-resolution images, the geographic object-based image analysis (GEOBIA) method has become the principle method as it is less sensitive to spectral variance within objects and can make use of both object features and spatial relations between objects (Hay and Castilla, 2006; Blaschke, 2010; Blaschke *et al.*, 2014). Image segmentation provides objects for GEOBIA by dividing an image into a set of spatially contiguous regions with spectral homogeneity or semantic coherence, on which the GEOBIA performance highly depends.

In high-resolution images, objects emerge representing various sizes, spectral heterogeneities, and shapes. It is difficult to segment various objects as single regions in one segmentation result (Borenstein and Ullman, 2008). Objects can be represented by multi-scale segmentations with large objects at coarse scales and small objects at fine scales (Burnett and Blaschke, 2003; Benz *et al.*, 2004; Bruzzone and Carlin, 2006; Sharon *et al.*, 2006). Furthermore, users may describe the same object at different semantic levels, i.e., 'tree', 'tree group', 'forest', which also requires the multi-scale segmentation of objects. After multi-scale segmentation, the important work is to select meaningful segmentation scale(s) for specific applications (e.g., Carleer *et al.*, 2005; Neubert *et al.*, 2008; Clinton *et al.*, 2010; Drăguț *et al.*, 2010; Zhang *et al.*, 2012; Yang *et al.*, 2014). Hence, multi-scale segmentation can be viewed as providing candidates of meaningful segmentation scales.

The region growing method is a good choice to generate multi-scale segmentations by setting different stopping rules (Dey *et al.*, 2010). It can produce spatially contiguous regions with closed boundaries intrinsically. These regions are viewed as image objects in GEOBIA directly. The stopping rule can be viewed as the scale parameter. On the control of a scale parameter, if more growing iterations are allowed, the coarser-scale segmentation result is produced. On the contrary, if fewer growing iterations are performed, the segmentation result is at finer scale. Region growing is convenient to form the horizontal-topological links within a segmentation scale. However, to form the hierarchical-nested links between different scales, the multi-scale segments should be nested.

Among region growing methods, the hierarchical stepwise optimization (HSWO) (Beaulieu and Goldberg, 1989) and hierarchical segmentation (HSEG) (Tilton *et al.*, 2012) algorithms can produce a segmentation hierarchy, in which the segmentations at coarser scales are merged from regions at finer scales. Hence, the multi-scale segmentations are nested intrinsically. On the other hand, local-oriented region growing methods (Câmara *et al.*, 1996; Baatz and Schäpe, 2000; Castilla *et al.*, 2008; Wang *et al.*, 2010) search the merging pair within local vicinity. Local-oriented methods run fast and can take the local contexts into account. However, local-oriented region growing cannot guarantee to produce nested multi-scale segments just by setting different scale parameters.

To produce nested multi-scale parameters by local-oriented region growing, the segmentation procedure can be controlled by a set of gradually increased scale parameters (e.g., Zhang *et al.*, 2013), where each scale parameter determines a segmentation scale. The merging iterations at a segmentation scale ($k$) are performed on the segments produced at its former segmentation scale ($k - 1$). Then, the multi-scale segments are nested. Usually, the scale parameters are defined as the threshold of similarity measure between adjacent regions. Then, a set of gradually increased scale parameters is independently assigned ahead of segmentation by setting a constant gap between neighboring scale parameters. A question raised is how to appropriately assign the set of scale parameters to capture the variety among different images and cover meaningful segmentation scales.

Xueliang Zhang, Xuezhi Feng, and Pengfeng Xiao are with the Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, the Key Laboratory for Satellite Mapping Technology and Applications of State Administration of Surveying, Mapping and Geoinformation of China, and the Department of Geographic Information Science, Nanjing University, 163 Xianlin Avenue, Nanjing 210023, Jiangsu, China (xiaopf@nju.edu.cn).

Photogrammetric Engineering & Remote Sensing
Vol. 81, No. 6, June 2015, pp. 461–470.
0099-1112/15/461–470

doi: 10.14358/PERS.81.6.461

The features and objects are various among different images, so that the similarity measure varies across images, resulting in the variety of scale parameters. On the other hand, the gap between scale parameters determines the number of scales in the set, which relates to the coverage of meaningful segmentation scale directly. For an image, if the gap is small, it can produce enough segmentation scales as candidates of meaningful segmentation scales, but too many scales make it difficult to select the meaningful scales. By contrast, if the gap is large, it may lead to partial coverage of meaningful scales. Hence, if the set of scale parameters is assigned ahead of the segmentation procedure, it is independent with specific images, which is difficult to capture the variety among different images and to produce appropriate number of scales to cover meaningful segmentation scales.

In this study, we focus on automatically setting a set of gradually increased scale parameters for region growing method, especially for local-oriented region growing method, to produce nested multi-scale segmentations. It is an extension of the step-wise scale parameter (SWSP) strategy by Zhang *et al.* (2013). The contributions of this study mainly include two aspects. First, instead of setting scale parameters independently by SWSP, we propose the adaptively increased scale parameter (AISP) strategy to determine scale parameters dynamically during the segmentation procedure, making the scale parameters adaptive to specific images and segmentation procedures. Second, the influence on segmentation accuracy by applying gradually increased scale parameter to local-oriented region growing is analyzed, which reveals the importance of changing growing regions.

## Methodology

The flow diagram of the proposed multi-scale segmentation method is presented in Figure 1. First, the region adjacency graph (RAG) is built based on initial segmentation. Then, the local-oriented region growing is applied under the control of AISP. Finally, the segment tree model is used to record and export nested multi-scale segmentation results.

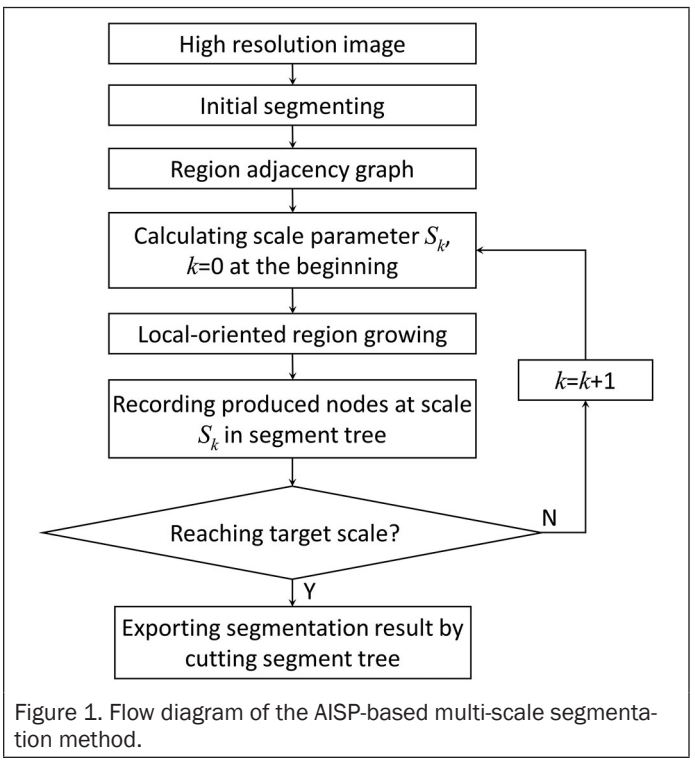



Figure 1. Flow diagram of the AISP-based multi-scale segmentation method.

### Region Growing Based on Graph Model

The local-oriented region growing is performed on the basis of RAG, where regions and adjacency between regions are represented by nodes and arcs, respectively (Trémeau and Colantoni, 2000; Felzenszwalb and Huttenlocher, 2004). RAG is built from initial segmentation, which is over-segmented but has accurate boundaries. The primary region growing method in (Zhang *et al.*, 2013) is used to produce initial segmentation. Other methods, i.e., watershed transform (Vincent and Soille, 1991), mean-shift based method (Comaniciu and Meer, 2002), can also be adopted.

The arc weight is calculated according to the merging criterion (*MC*), indicating the similarity between two adjacent regions. In this study, a smaller arc weight indicates a greater similarity between two adjacent regions. The merging criterion includes four features: the region size ($a$), change of standard variation (*CStd*) and compactness (*CComp*) after a virtual merging, and edge strength (*ES*). Region size is the number of pixels in a region, directly relating to segmentation scale. *CStd* and *CComp* reflect the change of region homogeneity and compactness caused by region merging, which drive to generate homogeneous and compact segments, respectively (Zhang *et al.*, 2013). In terms of *ES*, an adaptive edge penalty function $g(ES)$ is applied (Yu and Clausi, 2008; Zhang *et al.*, 2014). The effectiveness of *ES* is weak at initial merging iterations and getting stronger gradually at latter iterations. Finally, the features are combined to form the merging criterion as below:

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

where $a_1$ and $a_2$ are the size of two adjacent regions, and $\omega$ is the spectral weight with default value 0.5. If $\omega$ is set large, region growing concentrates on generating homogeneous regions. Otherwise, it would be driven to generate compact regions.

The local-mutual best region merging (LMM) (Baatz and Schäpe, 2000) and local best region merging (LBM) (Câmara *et al.*, 1996; Castilla *et al.*, 2008) are selected and applied on RAG. In LMM, two adjacent regions are allowed to be merged if they are the best neighbor of each other. However, in LBM, an adjacent region is merged into the growing region if the adjacent region is the best neighbor of the growing region. After each merging iteration, the features of the new region are recalculated, and the adjacency and arc weights in RAG are updated. The stopping rule is defined as the threshold of arc weight, which also serves as the scale parameter. When all the arc weights in RAG are larger than a given threshold, the region growing procedure stops on the control of this scale parameter.

### Adaptively Increased Scale Parameter

To generate nested multi-scale segmentations by local-oriented region growing, the step-wise scale parameter (SWSP) strategy (Zhang *et al.*, 2013) was proposed. SWSP strategy defines a set of increased scale parameters, $\{S_1, S_2, \ldots, S_k, \ldots, S_{max} \mid S_1 < S_2 < \ldots < S_k < \ldots < S_{max}\}$, in which each scale parameter determines a segmentation scale. Given the target scale parameter $S_k$, the region growing first satisfies the lowest scale parameter $S_1$, and then up to $S_2$, and finally, after the step-wise increasing process, reaches $S_k$. Owing to the gradually increased scale parameters, the regions at coarse scales are merged from the ones at fine scales, resulting in nested multi-scale segments. However, the scale parameters in SWSP are defined prior to segmentation without considering specific image features and merging procedure. To make the scale parameters adaptive to specific images, the following two aspects should be taken into consideration.

1. The distribution of arc weights in different images is various, and the change of arc weights in growing procedure is complicated. The set of scale parameters should capture this variety so that it can be qualified to control the growing procedure for various images.
2. In a given image, different objects are peculiarly presented at several scales, as shown in Figure 2. We consider these scales as meaningful scales, which should be covered by multi-scale segmentation results. To ensure the coverage, the scale parameters should increase smoothly and the increment step parameters should be set properly.

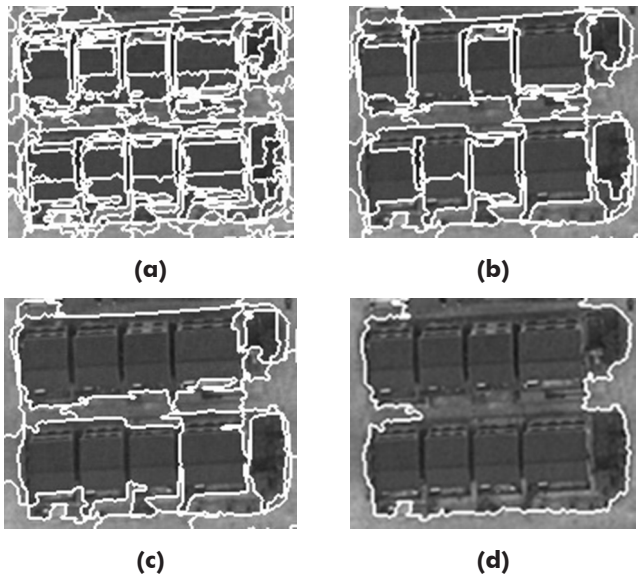



Figure 2. Samples of meaningful segmentation scales. Objects are severely over-segmented in (a) and under-segmented in (d). (b) and (c) are considered as meaningful because the single houses in (b), and the group of houses and trees in (c) are represented as single regions.

Following the above considerations, SWSP is extended to adaptively increased scale parameter (AISP). Instead of assigning scale parameters independently, we calculate the scale parameter based on the mean arc weight ($\bar{w}$) in RAG. In the region growing process, the expansion of regions makes $\bar{w}$ increased according to Equation 1, resulting in increased scale parameters adaptive to specific images and growing procedure.

However, if $\bar{w}$ is assigned as the scale parameter directly, too many regions would be merged at each scale, leading to few scales and partial coverage of meaningful scales. Hence, a normalized factor ($NF_k$) is added to calculate the $k^{\text{th}}$ scale parameter ($S_k$) as below:

$$S_k = \bar{w}/NF_k. \tag{2}$$

The increment of $S_k$ is jointly determined by $\bar{w}$ and $NF_k$. Generally, $\bar{w}$ is increased itself in the merging procedure. $NF_k$ decreases from a large number to 1 at a proper rate to form the continuous increment of $S_k$, which helps to cover meaningful scales entirely. $NF_k$ is involved as below:

$$NF_k = \begin{cases} NF_0, & \text{if } k = 0 \\ NF_{k-1}, & \text{if } k \geq 1 \cap P \geq T_P \\ \beta \cdot NF_{k-1} & \text{if } k \geq 1 \cap P < T_P \end{cases} \tag{3}$$

First, it is assigned as the initial normalized factor ($NF_0$). Then, $NF_k$ ($k \geq 1$) is calculated on the basis of $NF_{k-1}$ at the former segmentation scale. At scale $k-1$, if there are many merging iterations performed, $\bar{w}$ has been significantly changed, then $NF_k$ is set the same as $NF_{k-1}$. On the other hand, if few merging iterations are performed at scale $k-1$, the change of $\bar{w}$ is not significant. In this case, $NF_k$ is calculated by multiplying $NF_{k-1}$ with the decreasing rate ($\beta$, $0 < \beta < 1$). Hence, the number of merging iterations at the former scale determines the normalized factor at current scale. We define the proportion ($P$) of the number of merging iterations to the number of regions at the former scale. If $P$ is less than the threshold ($T_P$), the normalized factor is decreased to increase the scale parameter.

The threshold $T_P$ determines where to change the value of $NF_k$. The influence of $T_P$ is not as significant as the other two parameters of $NF_0$ and $\beta$, which determines how to change the value of $NF_k$. $T_P$ is set as 0.1 in this study because it is small enough to form the smooth change of scale parameters. $NF_0$ should be set large enough so that the scale parameter could be small enough to cover small objects at fine scale. In terms of the parameter $\beta$, a small value would accelerate the decreasing procedure of the normalized factor, leading to partial cover of meaningful scales. However, a large value slows down the decreasing procedure of the normalized factor, producing too many scales. In this study, $NF_0$ and $\beta$ are set as 10 and 0.9, respectively. The effectiveness is analyzed in detail in the experiments to indicate how to set the two parameters properly.

### Multi-Scale Segmentation

On the control of AISP, region growing is implemented on RAG until reaching the target scale by satisfying the smaller scale parameters sequentially. A segmentation scale is indicated by the serial number ($k$). If $k$ is large, segments are at a coarse scale. By contrast, a small $k$ results in a fine-scale segmentation. Since it is difficult to link the scale parameters directly to the semantics of segments, the meaningful scales for a given application still need to be selected from the multi-scale segmentations. To accelerate the selection process, we introduce the segment tree to represent and export multi-scale segmentations. The multi-scale segmentation algorithm is described in Table 1.

Table 1. The Multi-scale Segmentation Algorithm on the Control of AISP Based on the Graph Model and Segment Tree

Input: The region adjacency graph (RAG), the segment tree (ST).

Variables: *nl*–the number of levels in ST, *cs*–the serial number of current scale.

Parameter: The serial number of the target scale *k*.

Output: The segmentation result at scale *k*.

1) If ST is empty, assign the nodes in RAG as leaf nodes of ST, and set *cs* as 1.
   Else if *nl*-1<*k*, set *cs* as *nl*-1.
   Else if *nl*-1≥*k*, go to 3).
2) Repeat when *cs*≤*k*.
   (2-1) Calculate the current scale parameter $S_{cs}$ by Equation 2.
   (2-2) Perform region growing controlled by $S_{cs}$ based on the remaining nodes in RAG at scale *cs*-1.
   (2-3) Assign all the remaining nodes in RAG as the nodes at level *cs*+1 in ST.
   (2-4) Set *cs*=*cs*+1.
3) Output segmentation result by cutting ST at level *k*+1.

The initial segments are the leaf nodes in the tree and the segments at scale $S_k$ are represented as the nodes at level $k+1$ of the tree. As $S_k$ increases gradually, the segments are getting coarser and represented by the nodes at higher level

of the tree. Finally, the root node represents the whole image. Each node records the labels of its parent node and children nodes, building the hierarchical links between multi-scale segments. Then, we can find all the corresponding leaf nodes for each node in the segment tree. After combining the initial segments represented by the corresponding leaf nodes of a tree node, we can obtain the region indicated by the tree node. When exporting a segmentation result at scale $S_k$, we cut the segment tree at level $k + 1$, which is to find all the leaves corresponding to each node at the level. Once the segment tree has been built, it is much more efficient to produce multi-scale segmentations by cutting the tree at different levels than by repeating the similar region growing procedure in many times.

Finally, if the target scale $k$ is coarse enough and the meaningful scales are included in the segment tree, then users can select the proper scale(s) by cutting the tree for specific applications. However, if the nodes at the coarsest level, except for the root node, are still too fine to represent certain objects, a larger scale parameter needs to be set to produce coarser segments. In this case, region growing is not performed based on the initial segmentation but on the graph nodes related to the coarsest level of the segment tree, which avoids repeating the same growing procedure from the initial segmentation to the coarsest level.

## Experiments

### Data and Evaluation Methods

A QuickBird-2 scene acquired on 02 March 2008 in Hangzhou, China and a WorldView-2 scene acquired on 20 May 2010 in Xuzhou, China are used in the experiment, as shown in Figure 3. Three subsets of the QuickBird scene and one subset of the WorldView scene are selected as samples to show the effectiveness of AISP. The test images are called as T1 to T4, respectively.

The spatial resolution of T1 and T2 is sharpened to 0.6 m by the method proposed by Zhang (2002), as shown in Figure 4. The test images T3 and T4 are shown in Figure 9. The spatial resolution is 2.4 m and 2.0 m, respectively. The sizes of T1 to T4 are 658× 504, 538 × 546, 512 × 512, 512 × 512 pixels. The test images T1 and T3 represent an urban landscape, and T2 and T4 indicate a rural landscape.

The supervised evaluation is performed on T1 and T2 to show the segmentation accuracy. The references are produced by manual delineation to separate different geographic objects, as shown in Figure 4. There are 165 and 106 reference objects for T1 and T2, respectively. The supervised evaluation indicators include the bidirectional consistency indicator (*BCE*) (Martin, 2003), symmetric partition distance ($D_{sym}$) (Cardoso and Corte-Real, 2005), and adjusted Rand index (*ARI*) (Hubert and Arabie, 1985).

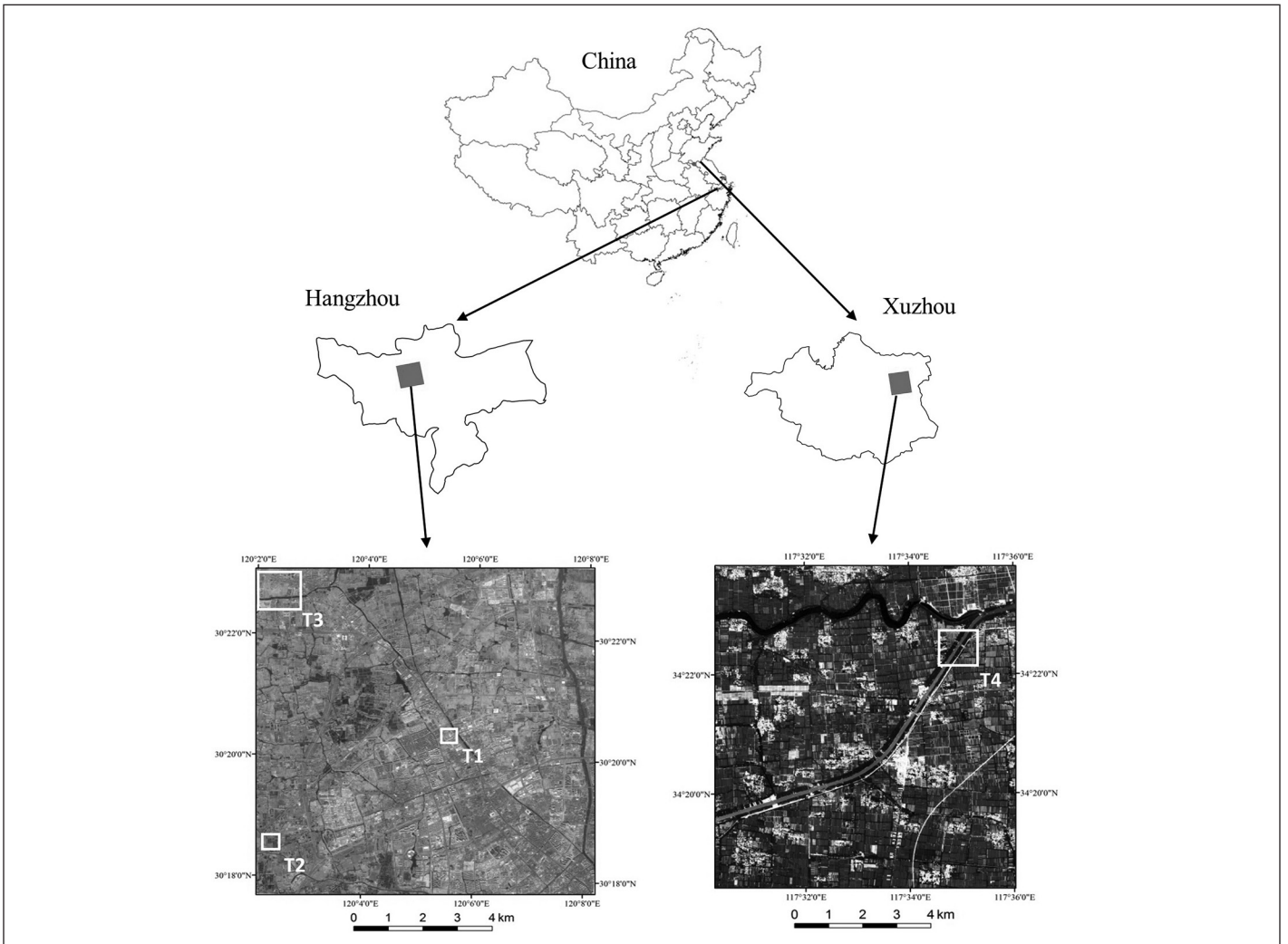
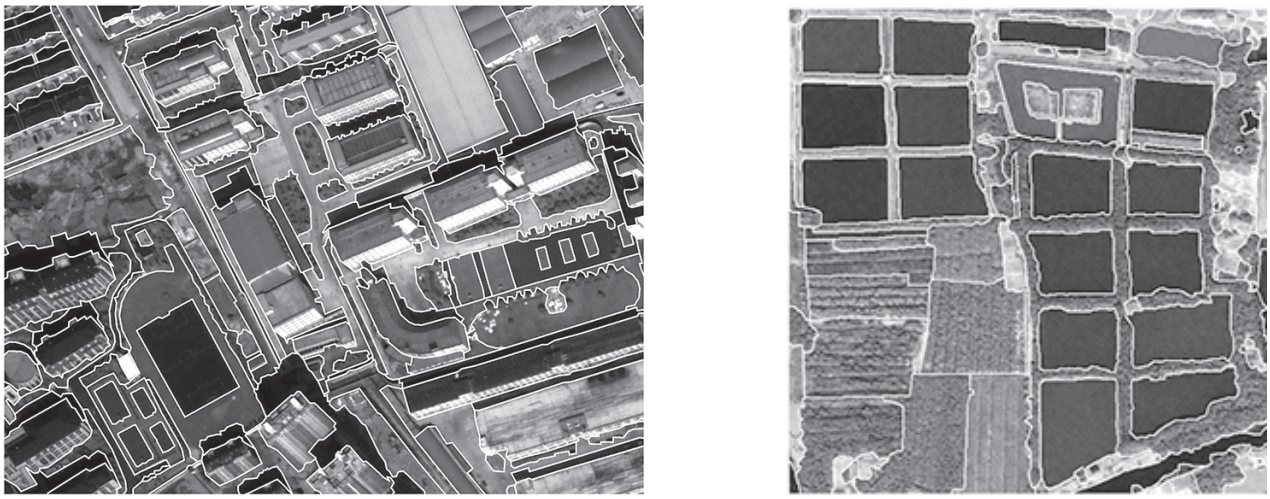



Figure 3. Locations of the test images T1 to T3 in the QuickBird scene in Hangzhou, China and the test image T4 in the WorldView scene in Xuzhou, China.



(a) (b)

Figure 4. The references of the test images T1 (a) and T2 (b). There are 165 and 106 reference objects in T1 and T2, respectively.

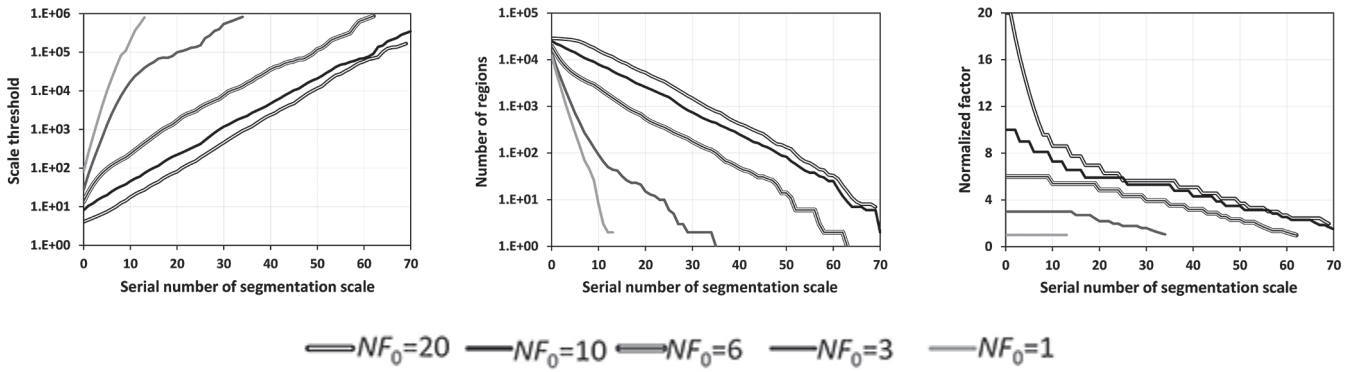



Figure 5. Evolution curves of the (a) scale threshold ($S_k$), (b) number of regions (*NR*) in segmentation, and (c) normalized factor ($NF_k$) as the serial number of segmentation scale ($k$) increases by applying LMM on test image T1. The decreasing rate ($\beta$) is set as 0.9 and the initial normalized factor ($NF_0$) is set differently.

*BCE* measures the consistence between the segmentation result and the reference. At first, the local refinement error (*LRE*) is defined to measure the degree to which the reference ($R$) and the segmentation ($S$) agree at a pixel $p_i$:

$$LRE(S,R,p_i) = \frac{\left| B(S,p_i) \setminus B(R,p_i) \right|}{\left| B(S,p_i) \right|} \tag{4}$$

$$BCE = \frac{1}{N}\sum_{i=1}^{N} \max\left\{LRE(R,S,p_i),\, LRE(S,R,p_i)\right\} \tag{5}$$

where $B(A, p)$ is the segment in the partition $A$ that contains point $p$, and $\cdot\setminus\cdot$ denotes the set difference. For example, $B(R, p)\setminus B(S, p)$ denotes the set of pixels $\{p \mid p \in B(R, p), p \notin B(S, p)\}$. In the over-segmentation case, $LRE(S, R, p_i)$ is large, and $LRE(R, S, p_i)$ is small, and in the under-segmentation case, by contrast, $LRE(S, R, p_i)$ is small, and $LRE(S, R, p_i)$ is large. In order to address the sensitivity to both over- and under-segmentation, *BCE* is defined to adopt the larger local error at each pixel and combine all the $N$ pixels within the image. Moreover, an accuracy map can be generated by presenting the local error at each pixel, as shown in Figure 8 and Figure 9.

$$BCE = \frac{1}{N}\sum_{i=1}^{N} \max\left\{LRE(R,S,p_i),\, LRE(S,R,p_i)\right\} \tag{6}$$

$D_{sym}$ is defined as the minimal proportion of pixels that must be removed from both the reference and the segmentation result so that the remaining pixels are identical. The indicator of *ARI* is the measure of correspondence between the segmented result and the reference; it evaluates the degree of agreement of two partitions based on the comparison of object triples, i.e., how the three distinct objects are delineated by the two partitions. All the three indicators are sensitive to both over- and under-segmentation, and range from 0 to 1. The lower *BCE* and $D_{sym}$ values and the higher *ARI* value indicate the better segmentation quality.

### The Effectiveness of $NF_0$ and β of AISP on Scale Coverage

In this subsection, we focus on analyzing the effectiveness of the initial normalized factor $NF_0$ and the decreasing rate $\beta$ on the involvement of the scale parameters, which indicates how to set $NF_0$ and $\beta$ properly. The number of regions (*NR*) is used to indicate the segmentation scale. A segmentation with more regions is considered at finer scale than that with fewer regions. The experiments have been performed extensively on several images; the test image T1 is selected as a sample to illustrate the findings.

Figure 5a and 5b show that when $NF_0$ is set small, i.e., 1 and 3, the changes of scale threshold ($S_k$) are significant, making *NR* change significantly accordingly. However, when $NF_0$ is set as 6, 10, and 20, the changes of $S_k$ and *NR* are smoother and slower than those with small $NF_0$. Moreover, the changing rate for the three large $NF_0$ is similar with each other. The different changing rates of $S_k$ and *NR* result in different numbers of segmentation scales. Consequently, the coverage of meaningful segmentation scales would be different.

For T1, the segmentations scales with approximately 150 to 1,100 regions are viewed as meaningful candidates because various objects can be represented well within this range through visual analysis. Given a decreasing rate $\beta$, the number of scales within the meaningful range produced by setting different $NF_0$ is shown in Table 2. When $NF_0$ is equal to or larger than 6, the number of meaningful scales is similar for both the LMM and LBM methods, but it is decreased significantly when $NF_0$ is set as 3 and 1, leading to the limited candidate scales.

Combining Figure 5b and 5c, we can see that the meaningful scales are controlled by the normalized factors approximately ranging from 4 to 5.5 when is $NF_0$ 20, 10, and 6. Hence, the large $NF_0$ can capture the sensitive normalized factors and make the scale parameter change smoothly to cover meaningful segmentation scales. On the other hand, when $NF_0$ is 3 or 1, since it is smaller than the sensitive normalized factors, the scale parameter is increased too fast, leading to limited numbers of candidates to cover meaningful scales.

According to the above analysis, we can know that if $NF_0$ is set smaller than the sensitive normalized factors, it cannot produce enough candidates to cover meaningful segmentation scales. The number of meaningful scales is similar when setting different $NF_0$ if they are larger than the sensitive normalized factors. However, the sensitive normalized factors for different images would be various. We set $NF_0$ as 10, which is large enough to cover the sensitive normalized factors after extensive tests.

On the other hand, given a $NF_0$, the number of meaningful candidate scales is also changed when setting different $\beta$. Generally, if $\beta$ is set large, more segmentation scales are produced, but the difference is not very significant according to Table 2. After extensive tests, we choose to set $\beta$ as 0.9 since the evolution of meaningful candidate scales is smooth enough.

Table 2. Number of Segmentation Scales within Meaningful Range for the Test Image T1 When Setting Different Initial Normalized Factors ($NF_0$) and Decreasing Rates ($\beta$) for Local-Mutual Best Region Merging (LMM), and Local Best Region Merging (LBM) Method

| $NF_0$ | Number of scales | | | | | |
|---|---|---|---|---|---|---|
| | $\beta$ =0.8 | | $\beta$ =0.9 | | $\beta$ =0.95 | |
| | LMM | LBM | LMM | LBM | LMM | LBM |
| 20 | 15 | 16 | 19 | 17 | 19 | 21 |
| 10 | 14 | 14 | 18 | 18 | 20 | 19 |
| 6 | 12 | 15 | 17 | 19 | 22 | 22 |
| 3 | 6 | 4 | 6 | 4 | 6 | 4 |
| 1 | 4 | 1 | 4 | 1 | 4 | 1 |

### The Effectiveness of AISP on Segmentation Accuracy

In this subsection, the effectiveness of AISP on segmentation accuracy is analyzed to show how the gradually increased scale parameters could influence the performance of local-oriented region growing. In Table 3, the segmentation accuracy is similar for LMM whether AISP is applied or not, and it is also similar when different $NF_0$ is set. This shows that AISP would not influence the segmentation accuracy significantly for LMM method. For LBM method, the segmentation accuracy is similar when AISP is applied with different $NF_0$. However, AISP can improve the segmentation accuracy for LBM significantly, making LBM achieve similar accuracy as LMM. The segmentation results in Figure 6 further show the difference. The reasons are illustrated in Figure 7.

Table 3. The Accuracy of Segmentations at Similar Scale for Test Image T1 When Setting Different Initial Normalized Factors ($NF_0$). *NR* Represents the Number of Regions in Segmentation. The Lowest Row Represents the Segmentation Result without Using AISP

| $NF_0$ | LMM | | | | LBM | | | |
|---|---|---|---|---|---|---|---|---|
| | ***NR*** | *BCE* | $D_{svm}$ | *ARI* | ***NR*** | *BCE* | $D_{svm}$ | *ARI* |
| 20 | **281** | 0.591 | 0.495 | 0.463 | **250** | 0.582 | 0.476 | 0.494 |
| 10 | **289** | 0.582 | 0.479 | 0.470 | **245** | 0.575 | 0.467 | 0.500 |
| 6 | **275** | 0.595 | 0.500 | 0.462 | **260** | 0.593 | 0.488 | 0.474 |
| 3 | **273** | 0.589 | 0.490 | 0.467 | **230** | 0.587 | 0.489 | 0.494 |
| - | **276** | 0.578 | 0.476 | 0.485 | **245** | 0.646 | 0.557 | 0.420 |

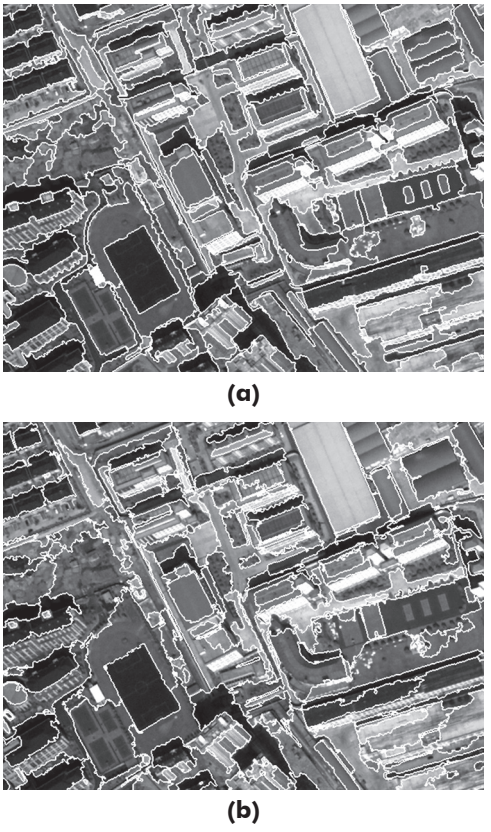



Figure 6. The segmentation results of LBM under the control of (a) AISP or (b)not. The region number of both results is 245.

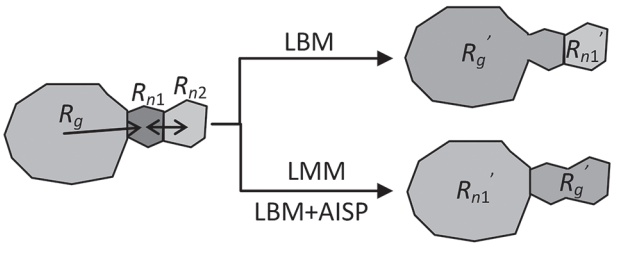



Figure 7. A sample explanation of how AISP can improve the segmentation accuracy for LBM method. The arrows in the left part point to the best neighbor for each region.

Given a growing region, LMM searches its locally mutual best neighbor for merging, which involves not only its neighbors, which is termed first-order neighbors, but also the second-order neighbors that are the neighbors of the first-order neighbors. If there is no mutual best neighbor, the merging procedure stops for the growing region and turns to another one. However, LBM just searches the local best neighbor for merging among the first-order neighbors. The growing region would not change until the minimal adjacent arc weight exceeds the scale threshold. As shown in Figure 7, when the growing region ($R_g$) is getting larger after several merges, its neighbors are still small, including the locally best first-order neighbor $R_{n1}$ and the second-order neighbor $R_{n2}$. Since the size of $R_g$ is large, the weight of arc between $R_g$ and $R_{n1}$ increases according to Equation 1. Then, even though $R_{n1}$ is the best neighbor of $R_g$, the best neighbor of $R_{n1}$ could be $R_{n2}$ because the size of $R_{n2}$ is small. Supposing $R_{n1}$ is also the best neighbor of $R_{n2}$, LMM refuses to merge $R_g$ and $R_{n1}$ and turns to merge the mutual best neighbors $R_{n1}$ and $R_{n2}$. However, LBM would keep merging $R_{n1}$ into $R_g$. On the other hand, when AISP is used to control LBM, since the scale parameter is increased gradually, the merging of $R_g$ and $R_{n1}$ could be prevented by a small-scale parameter during the increment schedule of the scale parameters, and turn to merge $R_{n1}$ and $R_{n2}$ first.

The sample analysis above shows that AISP can prevent LBM from keeping merging adjacent regions into the growing region when the arc weight is getting large, but changing the growing region and merging other region pairs with small arc weight first, which actually enlarges the searching range and enhances the optimization ability for LBM. Hence, AISP can improve the segmentation accuracy of LBM. Since the searching range of LMM is larger than that of LBM, it can change the growing region automatically under the constraint of local-mutual best. Thus, AISP cannot improve the accuracy significantly for LMM, as well as for other region growing strategies that have stronger optimization ability than LMM.

Furthermore, the segmentation accuracy of LMM on the control of AISP is compared with the multi-resolution segmentation (MRS) method in eCognition® Developer 8. The spectral weight of the two methods is set as 0.5. The supervised evaluation is performed on 5 and 4 multi-scale segmentation results for T1 and T2, respectively. The evaluation results are shown in Table 4. The *BCE* and $D_{sym}$ values of LMM+AISP are smaller than those of MRS, and the *ARI* values of LMM+AISP are higher, which shows the slightly better performance of LMM+AISP.

Segmentation results of T1 and T2 are presented in Figure 8 to show the visual difference further. Combined with the accuracy map indicated by *BCE*, the segmentation difference is clearly presented for each object. The object boundaries are accurate for both methods and the difference mainly caused by over- or under-segmenting certain objects. LMM+AISP tends to describe objects as single regions better, which is resulted from the adaptive edge penalty function as analyzed in (Zhang *et al.*, 2014).

### Multi-Scale Segmentation Results

To show the effectiveness of AISP on producing multi-scale segmentations as candidates of meaningful scales, four multi-scale segmentation results of test images T3 and T4 produced by LMM+AISP are selected by visual analysis and presented in Figure 9.

In the fine-scale segmentation of Figure 9a, single farmlands and houses are discriminated. As the segmentation scale is getting coarser, the water areas are segmented as single regions in Figure 9b, and the trees and farmlands are separated as large objects in Figure 9c. Finally, in Figure 9d, the impervious area, farmlands and trees are segmented as a few large regions at the coarse scale. In Figure 9e, the small objects, i.e., houses, farmlands, and alleys, are separated from each other. In Figure 9f, the houses, and farmlands with similar tone are merged into larger objects. In Figure 9g and 9h, the large objects, such as the river, roads, forest, rural residential areas, and super objects of farmland, are represented in the coarse segmentation result. Moreover, Figure 9 shows that the multi-scale segments are nested.

Table 4. Supervised Evaluation Results for Multi-scale Segmentations of Test Images T1 and T2 Produced by LMM+AISP, and MRS; *NR* Represents the Number of Regions

| Test image | LMM+AISP | | | | MRS | | | |
|---|---|---|---|---|---|---|---|---|
| | ***NR*** | *BCE* | $D_{sym}$ | *ARI* | ***NR*** | *BCE* | $D_{sym}$ | *ARI* |
| T1 | **740** | 0.734 | 0.627 | 0.290 | **730** | 0.758 | 0.663 | 0.249 |
| | **519** | 0.677 | 0.566 | 0.354 | **515** | 0.705 | 0.603 | 0.305 |
| | **315** | 0.595 | 0.489 | 0.456 | **319** | 0.617 | 0.517 | 0.433 |
| | **198** | 0.562 | 0.459 | 0.504 | **195** | 0.577 | 0.472 | 0.501 |
| | **133** | 0.547 | 0.423 | 0.536 | **136** | 0.561 | 0.461 | 0.522 |
| T2 | **429** | 0.566 | 0.470 | 0.431 | **445** | 0.681 | 0.586 | 0.315 |
| | **186** | 0.450 | 0.349 | 0.531 | **188** | 0.532 | 0.438 | 0.452 |
| | **107** | 0.426 | 0.329 | 0.560 | **108** | 0.436 | 0.347 | 0.554 |
| | **79** | 0.416 | 0.323 | 0.572 | **79** | 0.424 | 0.349 | 0.553 |

Table 5. Segmentation Time (sec) Difference Caused by AISP; The Size of the Test Images T5 and T6 is 2000×2000 Pixels and 3171 × 3000 Pixels, Respectively

| Test image | AISP | | Without AISP | |
|---|---|---|---|---|
| | LMM | LBM | LMM | LBM |
| T5 | 20.0 | 19.9 | 19.5 | 19.3 |
| T6 | 57.0 | 56.0 | 53.5 | 52 |

### Segmentation Time

The additional computation time when applying AISP to local-mutual best region merging (LMM) and local best region merging (LBM) on two large test images, namely T5 and T6, is shown in Table 5. The segmentation is performed on a desktop computer with the CPU of 3.2 GHz. When AISP is applied, the segmentation time of LMM and LBM is a bit longer. This is because AISP needs to calculate the mean arc weight at each segmentation scale. The additional computational complexity of applying AISP is $O(N_a+B\cdot N_n)$, where $N_a$ and $N_n$ are the number of arcs and nodes in RAG, and $B$ is the number of arcs whose weights are changed at each merging iteration. Furthermore, we calculate the segmentation time of producing multi-scale segmentations by cutting the segment tree. It only takes 0.08 sec for T5 and 0.18 sec for T6. This shows the great priority of using the segment tree to represent multi-scale segments.

## Discussion

The gradually increased scale parameters in AISP are calculated based on the mean arc weight in the graph, which are dynamically determined during the merging procedure and adaptive to various images. AISP avoids the arbitrary setting of scale parameters ahead of segmentation. However, it is noted that AISP is proposed for region growing on the basis that the scale parameter is defined as the threshold of similarity measure between adjacent regions. AISP can be applied to region growing procedure with different merging criteria. However, if other measures, such as region size or number of regions, are defined as the scale parameter, AISP cannot extend to such cases.

To form continuous increment procedure of scale parameters and cover meaningful scales, the normalized factor (*NF*) is added into the scale parameter, which mainly depends on the parameters of initial normalized factor ($NF_0$) and decreasing rate ($\beta$). $NF_0$ determines the coverage of meaningful scales.

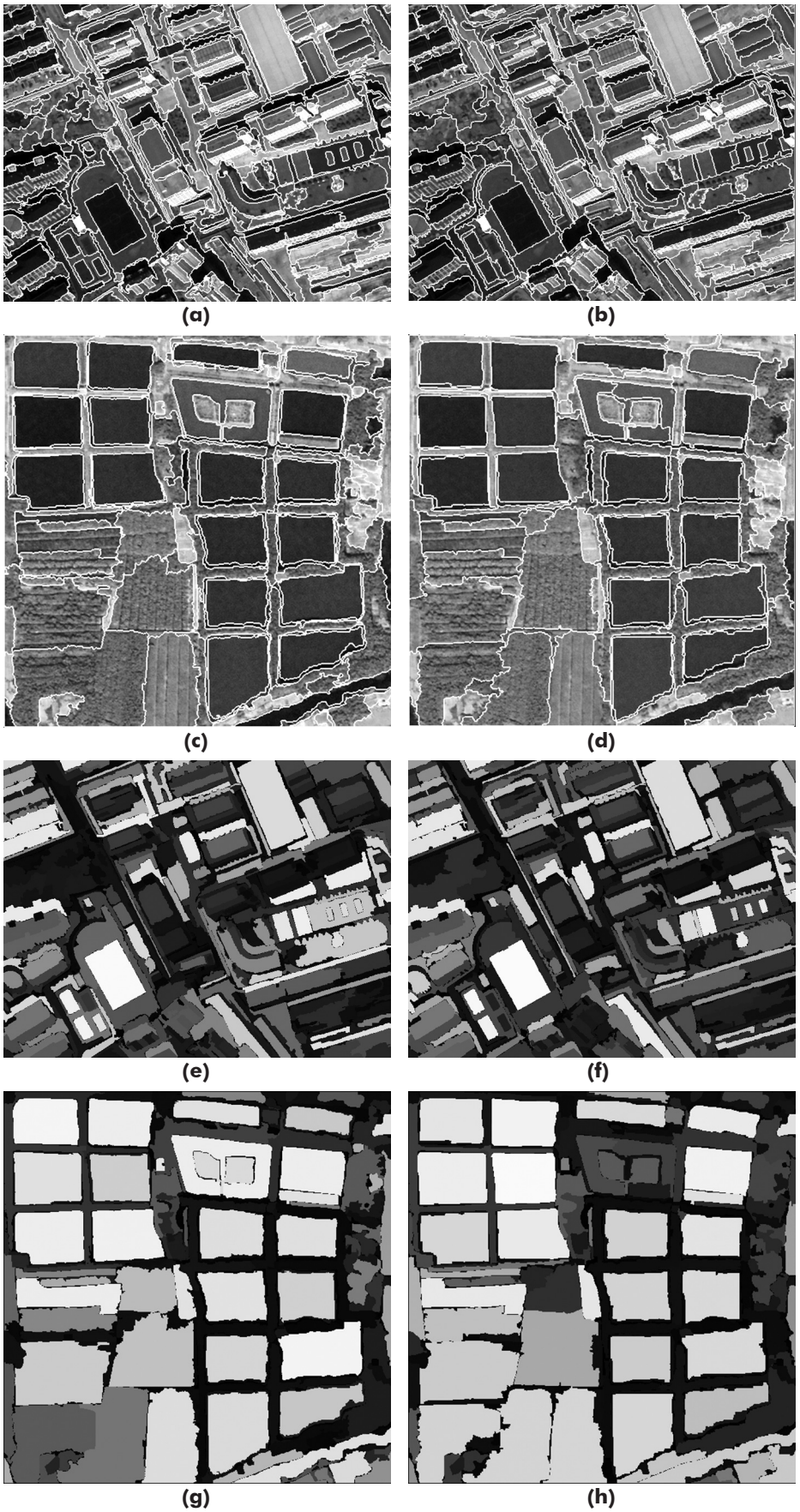



Figure 8. (a) through (d): segmentation results: (a) LMM+AISP, 315 regions, (b) MRS, 319 regions, (c) LMM+AISP, 107 regions, and (d) MRS, 108 regions. (e) through (h) are the accuracy maps indicated by *BCE* measure for (a) through (d), respectively, where the white indicates 0, and the black indicates 1.

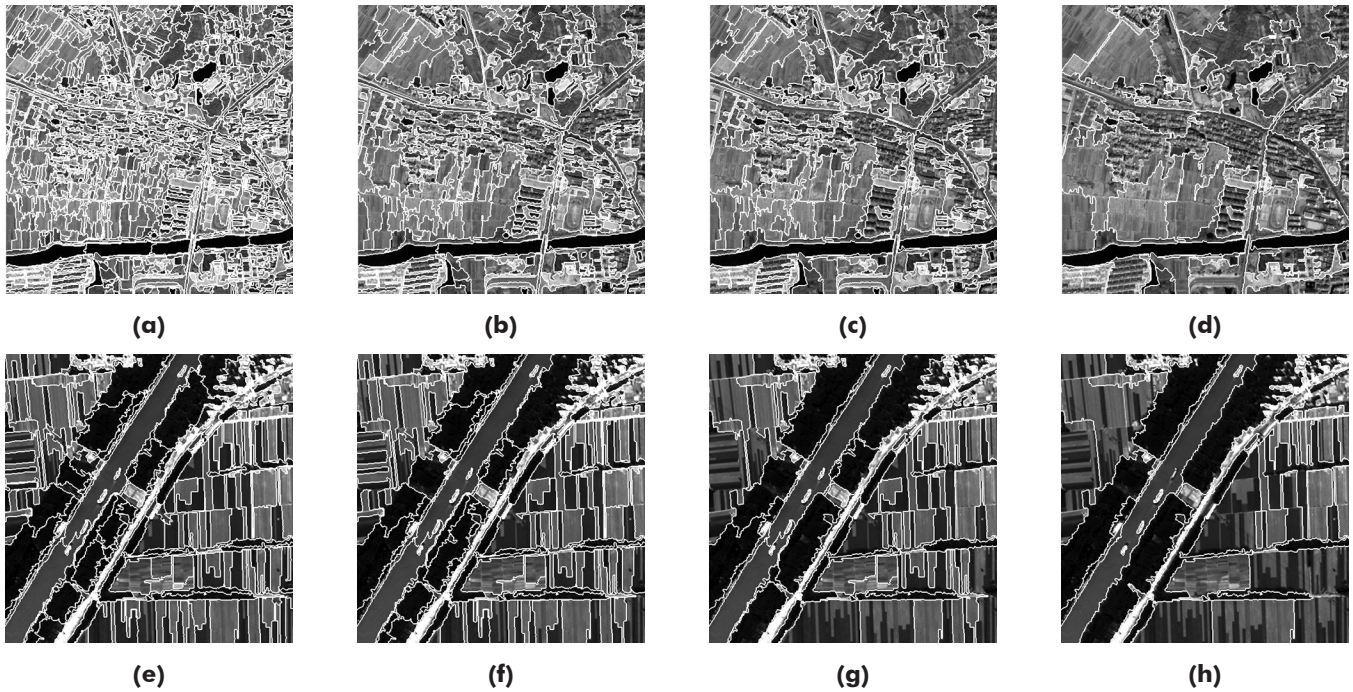



Figure 9. Multi-scale segmentation results of T3 (top row) and T4 (bottom row) using LMM+AISP. The number of regions is 1174, 335, 204, 61 in (a) through (d), and 226, 166, 83, 33 in (e) through (h).

If it is set larger than the sensitive normalized factors, it can make sure to cover the meaningful scales. $\beta$ determines the smoothness and continuity of the increased scale parameters. The experiments show that the increment of scale parameters is not very sensitive to $\beta$, as shown in Table 2. Hence, it is not difficult to set the two parameters according to the analysis.

The effectiveness of AISP on segmentation accuracy for local-oriented region growing is also analyzed to reveal how the gradually increased scale parameters can influence segmentation accuracy. The results show that AISP has no significant influence on the accuracy for LMM, but it can enhance the optimization ability of LBM by changing the growing region and allowing region pairs with large similarity to be merged first. The analysis presents a manner to improve the performance for local-oriented region growing, which is that we should not keep growing a region until it satisfy the stopping rule but change the growing regions after several iterations to merge other region pair with greater similarity first.

AISP controls to produce nested multi-scale segmentations, which serve as candidates of meaningful segmentation scales. In the following step, we should select meaningful scales for given applications. To accelerate scale selection, the segment tree is used to represent the nested multi-scale segments, where the nodes at different levels are hierarchically linked. It is very fast to produce multi-scale segmentations by cutting the segment tree at different levels. In this study, the segmentation scales for evaluation and shown are selected by visual analysis. Owing to the efficiency of cutting segment tree, we suggest to setting the target scale parameter large to build the segment tree. Then, the segment tree records segmentation scales coarse enough to cover meaningful segmentation scales and we can select meaningful segmentation scales by cutting the segment tree. In the future, we would pay attention to choosing meaningful scale(s) for a given application automatically. Then, there is a need to build the correspondence between the scale parameter and the semantic meaning of objects.

## Conclusions

The adaptively increased scale parameter (AISP) strategy has been presented for multi-scale segmentation of high-spatial resolution remote sensing images. AISP is used to control region-growing procedure to produce nested multi-scale segmentations. In this study, we focused on dynamically determining a set of gradually increased scale parameters in AISP, which are adaptive to specific images so that they can capture the variety among different regions. The normalized factor was introduced to calculate scale parameters, which contributes to form continuous upscaling mechanism to cover meaningful scales. The experiments proved the effectiveness of AISP on controlling multi-scale segmentation. Especially, the influence of gradually increased scale parameters on segmentation accuracy was analyzed in experiments, indicating the importance of changing growing regions for improving segmentation performance of local-oriented region growing. The comparison with MRS method showed that LMM + AISP can achieve slightly better performance.

## Acknowledgments

This work was supported by the Science and Technology Program of Zhejiang Province (Grant No. 2014F50022) and the Qinglan Project of Jiangsu Province (Grant No. 201423). The authors thank the editors and anonymous reviewers for their suggestions improving the paper.

## References

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

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

Benediktsson, J.A., J. Chanussot, and W.M. Moon, 2014. Very high-resolution remote sensing: Challenges and opportunities, *Proceedings of the IEEE*, 100(6):1907–1910.

Benz, U.C., P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen, 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., 2010. Object based image analysis for remote sensing, *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1):2–16.

Blaschke, T., G.J. Hay, M. Kelly, S. Lang, P. Hofmann, E. Addink, R.Q. Feitosa, F. van der Meer, H. van der Werff, F. van Coillie, and D. Tiede, 2014. Geographic Object-Based Image Analysis - Towards a new paradigm, *ISPRS Journal of Photogrammetry and Remote Sensing*, 87:180–191.

Borenstein, E., and S. Ullman, 2008. Combined top-down/bottom-up segmentation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(12):2109–2125.

Bruzzone, L., and L. Carlin, 2006. A multilevel context-based system for classification of very high spatial resolution images, *IEEE Transactions on Geoscience and Remote Sensing*, 44(9):2587–2600.

Burnett, C., and T. Blaschke, 2003. A multi-scale segmentation/object relationship modeling methodology for landscape analysis, *Ecological Modelling*, 168(3):233–249.

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

Cardoso, J.S., and L. Corte-Real, 2005. Toward a generic evaluation of image segmentation, *IEEE Transactions on Image Processing*, 14(11):1773–1782.

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

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

Clinton, N., A. Holt, J. Scarborough, Y. Li, and P. Gong, 2010. Accuracy assessment measures for object-based image segmentation goodness, *Photogrammetric Engineering & Remote Sensing*, 76(3):289–299.

Comaniciu, D., and P. Meer, 2002. Mean shift: A robust approach toward feature space analysis, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619.

Dey, V., Y. Zhang, and M. Zhong, 2010. A review on image segmentation techniques with remote sensing perspective, *Proceedings of the ISPRS TC-VII Symposium - 100 Years ISPRS, XXXVIII* (W. Wagner and B. Székely, editors), Vienna, Austria, (7A):31–42.

Drăguţ, L., D. Tiede, and S.R. Levick, 2010. ESP: A tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data, *International Journal of Geographical Information Science*, 24(6):859–871.

Felzenszwalb, P.F., and D.P. Huttenlocher, 2004. Efficient graph-based image segmentation, *International Journal of Computer Vision*, 59(2):167–181.

Hay, G.J., and G. Castilla, 2006. Object-based image analysis: Strengths, weaknesses, opportunities and threats (SWOT), *Proceeding of the 1st International Conference of OBIA*, URL: *http://www.isprs.org/proceedings/XXXVI/4-C42/Papers/01_Opening%20Session/OBIA2006_Hay_Castilla.pdf* (last date accessed: 11 April 2015).

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

Martin, D.R., 2003. *An Empirical Approach to Grouping and Segmentation*, Ph.D. dissertation, Computer Science Division, University of California, Berkeley, California, Report No. UCB/CSD-3-1268.

Neubert, M., H. Herold, and G. Meinel, 2008. Assessing image segmentation quality - Concepts, methods and application, *Object-Based Image Analysis*, Springer Berlin Heidelberg, pp. 769–784.

Sharon, E., M. Galun, D. Sharon, R. Basri, and A. Brandt, 2006. Hierarchy and adaptivity in segmenting visual scenes, *Nature*, 442(17):810–813.

Tilton J.C., Y. Tarabalka, P.M. Montesano, and E. Gofman, 2012. Best merge region-growing segmentation with integrated nonadjacent region object aggregation, *IEEE Transactions on Geoscience and Remote Sensing*, 50(11):4454–4467.

Trémeau, A., and P. Colantoni, 2000. Region adjacency graph applied to color image segmentation, *IEEE Transactions on Image Processing*, 9(4):735–744.

Vincent, L., and P. Soille, 1991. Watershed in digital spaces: An efficient algorithm based on immersion simulations, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6):583–598.

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

Yang, J., P. Li, and Y. He, 2014. A multi-band approach to unsupervised scale parameter selection for multi-scale image segmentation, *ISPRS Journal of Photogrammetry and Remote Sensing*, 94, pp. 13–24.

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

Zhang, X., P. Xiao, and X. Feng, 2012. An unsupervised evaluation method for remotely sensed imagery segmentation, *IEEE Geoscience and Remote Sensing Letters*, 9(2):156–160.

Zhang, X., P. Xiao, X. Song, and J. She, 2013. Boundary-constrained multi-scale segmentation method for remote sensing images, *ISPRS Journal of Photogrammetry and Remote Sensing*, 78:15–25.

Zhang, X., P. Xiao, and X. Feng, 2014. Fast hierarchical segmentation of high-resolution remote sensing image with adaptive edge penalty, *Photogrammetric Engineering & Remote Sensing*, 80(1):71–80.

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 and Remote Sensing*, 34, Part 4.