ISPRS Journal of Photogrammetry and Remote Sensing 131 (2017) 134–146

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Toward combining thematic information with hierarchical multiscale segmentations using tree Markov random field model

Xueliang Zhang, Pengfeng Xiao *, Xuezhi Feng

Department of Geographic Information Science, School of Geographic and Oceanographic Sciences, Nanjing University, Nanjing 210023, China
Collaborative Innovation Center of South China Sea Studies, Nanjing 210023, China
Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China

ARTICLE INFO

Article history:
Received 7 April 2017
Received in revised form 17 June 2017
Accepted 10 August 2017

Keywords:
High-spatial resolution remote sensing image
Multiscale segmentation
Scale selection
Hierarchical context
Geographic object-based image analysis

ABSTRACT

It has been a common idea to produce multiscale segmentations to represent the various geographic objects in high-spatial resolution remote sensing (HR) images. However, it remains a great challenge to automatically select the proper segmentation scale(s) just according to the image information. In this study, we propose a novel way of information fusion at object level by combining hierarchical multiscale segmentations with existed thematic information produced by classification or recognition. The tree Markov random field (T-MRF) model is designed for the multiscale combination framework, through which the object type is determined as close as the existed thematic information. At the same time, the object boundary is jointly determined by the thematic labels and the multiscale segments through the minimization of the energy function. The benefits of the proposed T-MRF combination model include: (1) reducing the dependence of segmentation scale selection when utilizing multiscale segmentations; (2) exploring the hierarchical context naturally imbedded in the multiscale segmentations. The HR images in both urban and rural areas are used in the experiments to show the effectiveness of the proposed combination framework on these two aspects.

1. Introduction

High-spatial resolution remote sensing (HR) images provide specific details of the Earth's surface, while the improved spatial resolution also results in the increase of the intra-class variability (Bruzzone and Carlin, 2006; Myint et al., 2011), bringing the challenge of accurately mapping geographic objects.

To deal with HR images, the geographic object-based image analysis (GEOBIA) has been a widely used method (Hay and Castilla, 2008; Blaschke et al., 2014), in which the HR image is first segmented into homogeneous regions viewed as image objects and the object-based analysis is then performed. A lot of studies have demonstrated the superiority of GEOBIA over pixel-based classification because of the transition of analysis unit from pixel to segmented object (Thomas et al., 2003; Yu et al., 2006; Myint et al., 2011).

Since image segmentation is performed on the spatial similarity or discontinuity of image features to generate homogeneous regions, it is difficult to generate a single segmentation to represent all the geographic objects in an HR image because of the gap between homogeneous segments and heterogeneous geographic objects. A possible solution to this gap is the generation of multiscale segmentations by setting different stopping rules or scale parameters in the segmentation procedure (Baatz and Schäpe, 2000; Benz et al., 2004; Arbelaez et al., 2011; Zhang et al., 2015), in which fine-scale segments are suited for small objects and coarse-scale segments are representing large objects. Another benefit of the multiscale segmentations is to express the hierarchical context from object part, object, object group, and finally to the whole scene as the segments are getting coarser (Lim et al., 2009; Munoz et al., 2010; Ren et al., 2012). This hierarchical context between segmentation scales naturally propagates long-range interactions that may be difficult to capture with pairwise interactions at a certain scale (Munoz et al., 2010). That is because each object at a generic level in the multiscale segmentations is hierarchically related to those at the higher and lower levels (Bruzzone and Carlin, 2006). The multiscale segmentation solution makes the scale parameter playing a key role in segmentation procedure and brings the challenge of optimizing segmentation scale parameter for GEOBIA.

* Corresponding author at: 163 Xianlin Avenue, Nanjing, China.
E-mail address: xiaopf@nju.edu.cn (P. Xiao).

http://dx.doi.org/10.1016/j.isprsjprs.2017.08.003

Originally, the focus of scale parameter optimization is mainly on selecting an optimal scale from a set of segmentation scales by unsupervised evaluation (Espindola et al., 2006; Corcoran et al., 2010; Gao et al., 2011), supervised evaluation (Carleer et al., 2005; Möller et al., 2007; Clinton et al., 2010; Liu et al., 2012; Witharana and Civco, 2014), or object-based classification accuracy (Laliberte and Rango, 2009; Smith, 2010; Gao et al., 2011; Dronova et al., 2012; Yang et al., 2017a,b). Even though segmentation itself is not the end of GEOBIA, the results demonstrate that segmentation scale does affect the effectiveness of the subsequent analysis step in GEOBIA. Usually, neither too fine nor too coarse segmentation scales are preferred. A medium segmentation scale can achieve the best performance in most instances depending on the targets of interest. Furthermore, the negative effect of under-segmentation is generally considered to be worse than that of over-segmentation (Kim et al., 2009; Liu and Xia, 2010; Johnson and Xie, 2013). Even though selecting an optimal scale can help to improve the performance of subsequent classification, the selected scale still contains over-segmented and under-segmented regions because different types of geographic objects may be better represented at different scales. To further overcome this problem, segments at different segmentation scales are selected and combined into a single segmentation to represent the various objects (Chen et al., 2009; Johnson and Xie, 2011; Yang et al., 2017a,b). Another solution could be selecting few meaningful segmentation scales for successive object-based analysis (Dragut et al., 2010, 2014; Yang et al., 2014), which reduces the analysis complexity because fewer segmentation scales are involved.

In addition to performing object-based analysis on a single optimized segmentation scale, it is preferred to perform the analysis directly on multiscale segmentations. A common solution is to employ a multiscale classification approach that classifies different types of objects at the peculiar segmentation scales (Burnett and Blaschke, 2003; Myint et al., 2011; Kim et al., 2011; Dronova et al., 2012), which has shown to achieve higher accuracy than single-scale based classification (Kim et al., 2011). However, it needs to choose appropriate segmentation scale(s) to extract particular objects (Myint et al., 2011), which requires a detailed investigation of segments at each scale, and the procedure can be time-intensive and subjective (Johnson and Xie, 2013). A second multiscale analysis solution is to stack the features of nested multiscale segments into a feature vector for classifying the finest unit in the segmentation hierarchy (Bruzzone and Carlin, 2006; Johnson and Xie, 2013). The improvement of classification accuracy demonstrates the effectiveness of multiscale segmentations on providing spatial context information. However, the hierarchical context within multiscale segmentations is not exploited because the links between segments at different levels are not utilized. The importance of each segmentation scale to specific land cover class is also not clear and the result is still sensitive to the selected multiscale segmentations, especially to the coarse-scale segmentations. A third multiscale analysis solution is to integrate the multiscale segments into the Markov random field (MRF) classification task as a data fusion problem (Moser et al., 2013). However, each segmentation map in the MRF framework is viewed as a distinct information source without taking advantage of the hierarchical context. Furthermore, the effectiveness of segmentation scale on the final classification result is not very clear.

In addition to the above multiscale analysis methods, we can also utilize the segmentation results by combining them with the classification result, achieving the fusion of thematic and spatial information (Tarabalka et al., 2010; Fauvel et al., 2013; Zhang et al., 2013; Konstantinidis et al., 2017). Usually, the combination work is performed on a single segmentation scale. The majority voting strategy is widely adopted for combination. Each region in a segmentation map is assigned to the most frequent class within it. The combination approach retains the accurate image classification while integrating the spatial information provided by segments, which has been proved to be an accurate, simple, and fast technique (Fauvel et al., 2013). However, the combination with a single segmentation is still sensitive to the selected segmentation scale (Tarabalka et al., 2012). Moreover, the segments control the object boundaries as a hard constraint, where the errors in a single segmentation result could directly propagated to the combination result.

Hence, we propose a novel information fusion method to combine multiscale segmentations with thematic information in a tree-MRF (T-MRF) model. The MRF model is defined on a tree representing the hierarchical multiscale segmentations and is thus called as T-MRF. The object boundaries in the combination result is jointly determined by the multiscale segmentations in the T-MRF model rather than purely constrained by a single segmentation. One benefit of the T-MRF model is to reduce the sensitivity of segmentation scales when utilizing multiscale segmentations. Another benefit is to provide an explicit way to explore the hierarchical context within multiscale segmentations by the optimization procedure that involves all the segmentation scales in the tree. The followed Section 2 presents the details of the multiscale combination model based on T-MRF. The datasets and experimental results are shown in Section 3. Section 4 presents the discussions and Section 5 draws the conclusions of this study.

2. Method

2.1. Overview

The basic idea of the proposed multiscale combination model is shown in Fig. 1. The entire workflow of the combination model is presented as Fig. 2. Firstly, a land-cover classification or object recognition map is generated, in which the class labels of each analysis unit are obtained, to represent the thematic information for this model. The focus of this step is usually to distinguish different land cover classes but neglect obtaining accurate object boundaries. Concurrently, hierarchical multiscale segmentation maps are generated by a nested multiscale segmentation method and represented as a segmentation tree. The segmentation step focuses on exploiting the spatial context to determine accurate boundaries while the class labels are usually not known. The thematic information and the segmentation tree are then combined to jointly determine the class labels and boundaries of geographic objects. During the combining procedure, the object labels are determined to be as close as those in the thematic information and at the same time, the object boundaries are jointly determined by the thematic labels and the saliency of the multiscale segments at each location.

2.2. Thematic information generation

The thematic information is represented as the land-cover class labels of each pixel. It could be either the image classification result that covers the entire scene or the object detection result that only involves several objects in the scene. Several strategies, e.g. land-cover classification based on pixels or over-segmented regions, object recognition or detection, information extraction, can be applied to obtain the thematic information. Generally, the quality of the thematic information would positively relate to the accuracy of the multiscale combination result. In this study, we are not focusing on improving the accuracy of the thematic information itself, but aiming at utilizing multiscale segmentations based on the thematic information. To serve as an example of thematic information, the pixel-based classification using the support vector machine (SVM) classifier (Chang and Lin, 2011) is adopted in the

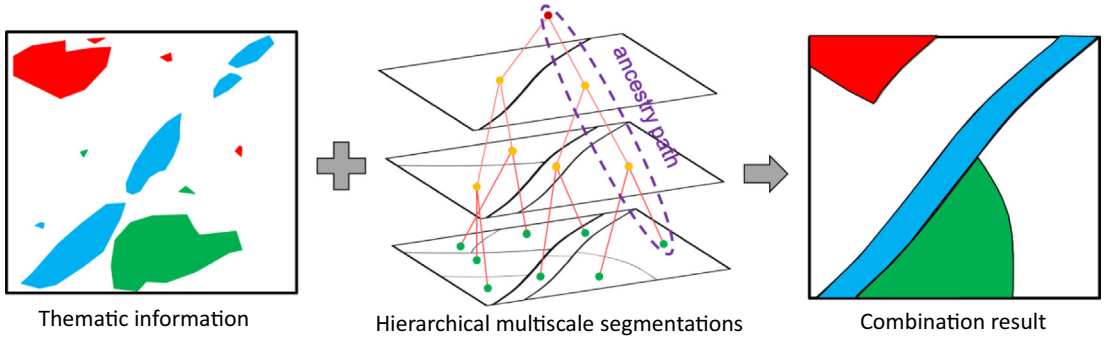

Fig. 1. Schematic diagram of the proposed multiscale combination framework. The multiscale segmentations are represented by a segment tree. The tree nodes are represented by the circles. The red and green circles in the segment tree indicate the root and leaf nodes, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

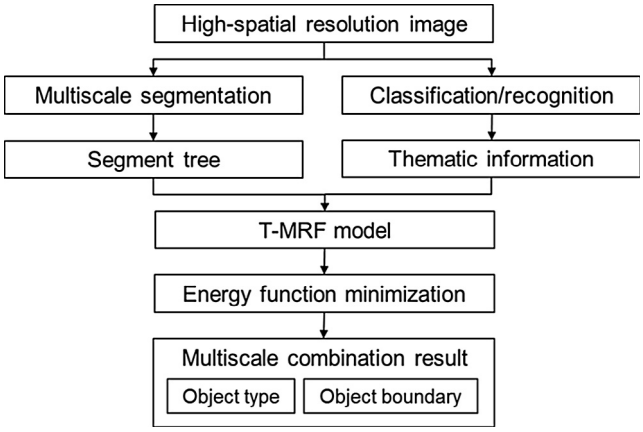

Fig. 2. Flow chart of the proposed multiscale combination framework.

combination model. The classifier could be freely chosen as long as it provides land cover class information. Other sources of thematic information as described above can also be used in the proposed multiscale combination model. The original spectral bands and the textural features of variance, homogeneity, contrast, and entropy from grey-level co-occurrence matrix (GLCM) (Haralick et al., 1973) are formed into the feature vector. The radial basis function kernel is used for SVM. The parameters of C and γ are tuned manually. The format of the classification result could be class probability or class label in the combination model, and the later one is adopted in this study.

2.3. Hierarchical multiscale segmentation

The hierarchical multiscale segmentations are composed of multiple segments from fine to coarse at each location, in which the small objects are supposed to be represented by fine segments at certain segmentation scales and the large objects are represented by coarse segments correspondingly. Furthermore, a fine-scale segment smaller than an real object is supposed to represent a part of the object, while a coarse-scale segment larger than an real object is to represent an object group. A preliminary requirement for the multiscale segments is that the segments at the same location should be nested. Otherwise the object boundaries would be confused when combining or fusing the multiscale segments.

The hierarchical multiscale segmentations are represented by the segment tree model as shown in Fig. 1. The tree nodes at different levels represent segments at different scales. An arc connecting a parent and a child node represents the inclusion relation between segments at adjacent scales. The leaf nodes represent the segments at the finest scale, and the nodes at upper levels represent segments at coarser scales. Finally the root node represents the whole image. An ancestry path in the tree is defined as the path from a leaf node up to the root node, revealing the transition from object part to the whole scene. The hierarchical context of each leaf node is conveyed by the ancestry path, in which a segment is gradually getting coarser and finally reaching the whole image.

Several region-based segmentation methods could be applied to produce the required hierarchical multiscale segmentations, e.g. multi-resolution segmentation method (Baatz and Schäpe, 2000), binary partition tree method (Salembier and Garrido, 2000), mean-shift method (Comaniciu and Meer, 2002), and hierarchical method (Arbelaez et al., 2011). Specially, the binary partition tree method (Salembier and Garrido, 2000) is also suited to the proposed combination method as it naturally produces a segment tree by iteratively recording the new region merged from two adjacent regions as a parent node. However, since the multiscale segmentation strategy is widely used in remote sensing applications, we do not adopt the binary partition tree method but build the segment tree based on multiscale segmentations. Specifically, the hybrid region merging method (Zhang et al., 2014) is used as the segmentation method, which combines the advantages of both the local- and global-oriented region merging methods. To produce nested multiscale segments, the strategy of setting adaptively increased scale parameters (Zhang et al., 2015) is adopted to control the segmentation procedure, where a set of increased scale parameters is automatically generated and the segments at coarser scales are generated by merging segments at next finer scale to make the multiscale segments nested. The regions in each segmentation are represented by the nodes at the same level in the segment tree. Finally the segment tree is constructed by recording the multiscale segmentations.

2.4. T-MRF combination model

The multiscale combination model has two objectives: one is to determine the object label as close as possible to the initial thematic information, and the other is to jointly determine the object boundaries by the thematic labels and the multiscale segments at each location. These objectives are respectively achieved by the unary potential and the pairwise potential in the energy function of the T-MRF model, which is defined on the segmentation tree. All the tree nodes and arcs are involved in the energy function. The T-MRF model proceeds to optimize the labels of all the segments in the tree.

The energy function $\varphi(Y)$ of the T-MRF model is defined as Eq. (1), where Y represents the class labels, the unary potential $\varphi^t(Y_i^l)$ indicates the cost of assigning a label to the ith segment at the lth

level (S_i^l) in the tree, and the pairwise potential $\varphi^s(Y_i^l, Y_p^{l+1})$ indicates the cost of assigning labels to S_i^l and its parent segment S_p^{l+1}. The weight parameter ω is to balance the two potentials. The unary and pairwise potentials are jointly represented by $\varphi(Y_i^l, Y_p^{l+1})$, which is convenient for expression in followed Eqs. (6) and (7).

$$\varphi(Y) = \sum_{i,l} \varphi(Y_i^l, Y_p^{l+1}) = \sum_{i,l} \varphi^t(Y_i^l) + \omega \sum_{i,l} \varphi^s(Y_i^l, Y_p^{l+1}) \tag{1}$$

$$\varphi^t(Y_i^l) = \sum_{(a,b)\in S_i^l} (1 - I(Y_i^l, L(a,b))) \tag{2}$$

The unary potential controls the segment label to be close to the existed thematic information and thus defined as Eq. (2), where $I(\cdot)$ is an identification function, whose value is 1 if the existed thematic label of a pixel (a, b) in S_i^l is same as that of the segment label Y_i^l, otherwise the function value is 0. According to Eq. (2), the unary potential value would be small if the label of the segment is same as that of most pixels within this segment. In the extreme case, when the existed thematic label of all the pixels in a segment is "water", then the unary potential value is 0 if the segment label is assigned as "water" while it is 1 if assigned as other class labels.

$$\varphi^s(Y_i^l, Y_p^{l+1}) = \begin{cases} -|S_i^l|(1-h_i^l), & if\ Y_i^l = Y_p^{l+1} \\ |S_i^l|(1-h_i^l), & if\ Y_i^l \neq Y_p^{l+1} \end{cases} \tag{3}$$

$$h_i^l = \exp\left(-2/|Std_{S_p^{l+1}} - Std_{S_i^l}|\right) \tag{4}$$

The pairwise potential jointly controls the labels of parent-child nodes in the segment tree as defined in Eq. (3). The saliency of a segment is defined in terms of the difference to its parent in the tree as Eq. (4). We suppose that as the growing of the segments within an ancestry path, a suddenly significant change of the homogeneity would indicate the merging of other types of objects. The segment right before this change is viewed as salient. The homogeneity of a segment is defined as the spectral standard deviation (Std) of the pixels within it. By combining Eqs. (3) and (4), if the segment S_i^l is salient, assigning a different label to its parent $Y_i^l \neq Y_p^{l+1}$ is preferred to lower the cost, while if the segment is not salient, it is preferred to assigning the same label to its parent $Y_i^l = Y_p^{l+1}$. This constraint helps reduce the risk of separating a salient segment into different thematic classes. A similar saliency-driven method has been utilized for unsupervised segmentation scale selection (Chen et al., 2009; Drăguţ et al., 2010), where the saliency is viewed as a "hard" constraint by selecting segmentation scale with the highest saliency. While in our study, the saliency serves as a "soft" constraint in the pairwise potential and the final combination result would also consider the existed thematic information in the unary potential. Additionally in Eq. (3), the area of a segment $|S_i^l|$ is included in the pairwise potential, which results in a great influence to the energy function for a large segment when changing its label.

2.5. Energy function minimization

The energy function is minimized to assign the optimal labels to each node in the segment tree. Since the tree structure has no loops, it enables to globally minimize the energy function using the sum-product message passing algorithm (Kschischang et al., 2001; Borenstein and Ullman, 2008) as described below.

According to Eq. (1), the energy function is decomposed into local terms of $\varphi^t(Y_i^l)$ for each node and $\varphi^s(Y_i^l, Y_p^{l+1})$ for each arc in the tree. The computation proceeds by sending messages between neighboring nodes in the segment tree, which is composed of a bottom-up and a top-down phase. It starts from the leaf nodes and iteratively passes to the root during the bottom-up procedure, and each node S_i^l sends a message $m_{S_i^l \to S_p^{l+1}}(Y_i^l)$ to its parent S_p^{l+1}. Successively, the top-down procedure starts proceeding from the root node and recursively to the leaf nodes, in which each node receives a message $m_{S_p^{l+1} \to S_i^l}(Y_i^l)$ from its parent. The optimal label of each node within the energy function is given by minimizing the summation of the bottom-up message it sent to and the top-down message received from its parent, as shown in Eq. (5). Specially, during the top-down passing phase, the segments are optimally labeled starting from the root node and progressing to the finest leaf nodes according to Eq. (5). The leaf nodes thus receive all the hierarchical context information from its upper nodes in the ancestry path and are output as the final combination result.

$$m(Y_i^l) = m_{S_i^l \to S_p^{l+1}}(Y_i^l) + m_{S_p^{l+1} \to S_i^l}(Y_i^l) \tag{5}$$

$$m_{S_i^l \to S_p^{l+1}}(Y_i^l) = \sum_{j=1}^{J} \min\left\{ m_{(S_j^{l-1}) \to S_i^l}(Y_j^{l-1}) + \varphi(Y_j^{l-1}, Y_i^l),\ Y_j^{l-1} \in (1,2,\ldots,n_c) \right\} \tag{6}$$

$$m_{S_p^{l+1} \to S_i^l}(Y_i^l) = \min \left\{ \begin{array}{l} \sum_{j=1, j\neq i}^{J} \min\left\{ m_{S_j^l \to S_p^{l+1}}(Y_j^l) + \varphi(Y_j^l, Y_p^{l+1}),\ Y_j^l \in (1,2,\ldots,n_c) \right\} \\ + \min\left\{ m_{S_p^{l+2} \to S_p^{l+1}}(Y_p^{l+1}) + \varphi(Y_p^{l+1}, Y_p^{l+2}),\ Y_p^{l+2} \in (1,2,\ldots,n_c) \right\} \\ + \varphi(Y_i^l, Y_p^{l+1}),\ Y_p^{l+1} \in (1,2,\ldots,n_c) \end{array} \right\} \tag{7}$$

How to calculate the messages in Eq. (5) is explained as followed. The message passing paths of calculating these messages are presented in Fig. 3. Supposing that there are n_c thematic classes in a specific combination task, a message consists of n_c values with each value for a thematic class label. The bottom-up message from a node S_i^l to its parent S_p^{l+1} is locally computed from the messages arriving from all its child nodes $\{S_j^{l-1}\}$, as shown in Eq. (6). The top-down message arriving to S_i^l from its parent S_p^{l+1} is also locally calculated as Eq. (7), involving the messages from the parent S_p^{l+1} and all the descendents of S_p^{l+1} except for S_i^l. There are two special cases to calculate messages for the leaf and the root node.

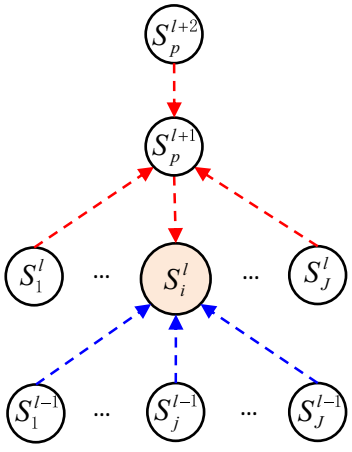

Fig. 3. Message passing paths for calculating the bottom-up (marked as blue) and the top-down (marked as red) messages for the segment S_i^l. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

To calculate bottom-up message for a leaf node, its child nodes are lacked. To calculate the messages for the root node, its parent node is viewed as itself.

3. Experiments

The experiments are designed to show the effectiveness of the multiscale combination model on reducing the dependency of segmentation scale selection and the benefit of utilizing the hierarchical context within multiscale segmentations. The sensitivity of segmentation scales is firstly analyzed in terms of the coarseness, interval, and number of segmentation scales. The combination weight in the energy function is then evaluated to show the trade-off between the existed thematic information and the multiscale segmentation during the combination procedure. Finally, the multiscale combination result is compared with single-scale combination result to highlight the effectiveness of utilizing the hierarchical context.

To implement the proposed combination model, the steps of multiscale segmentation and multiscale combination are fulfilled by the C# programming language. The codes are available upon request. Since the thematic information is task dependent and could come from various sources, we do not specially write codes to generate thematic information. It is obtained by the SVM classifier imbedded in the commercial software ENVI in this study.

3.1. Datasets

A WorldView-2 image and an IKONOS-2 image are used in the experiments. To clearly show the results, two subsets from each of the image are selected and shown in Fig. 4. The details of the images are presented in Table 1. The spatial resolution of the subset from the WorldView-2 image is 2 m, represents the rural landscape with villages surrounded by large-area croplands with windbreaks. The main land cover classes include buildings, other impervious area, cropland, tree, and water. The spatial resolution of the subset from the IKONOS-2 image is 1 m after pansharpening, representing the complex urban landscape with buildings of various sizes, shapes, and roof materials. The main land cover classes in this image include building, other impervious area, tree, and grass. Since the building shadows are clearly presented in the image, the shadow is viewed as a separate class in the result.

As a preparation for the multiscale combination, the SVM classification is first applied to each image as described in Section 2.2 to obtain the thematic information. There are 80 training pixels for each land cover class in the WorldView-2 image. Since the IKONOS-2 image is of greater complexity, more training pixels are involved in the classification procedure than that of the WorldView-2 image. Specifically, there are 2800 training pixels for each class of building and other impervious area, and 400 training pixels for each of the other three classes. The multiscale segmentations are produced by the hybrid region merging method described in Section 2.3. To produce multiple segmentation scales for the successive analysis, the finest segmentation scale is firstly generated with apparent over-segmentation. The number of segments at the next coarser scale is approximately half of the former scale. The coarsest scale only contains tens of segments, which is apparently under-segmented in terms of objects and represents object groups. Totally, there are 14 and 13 segmentation scales produced for the WorldView-2 and IKONOS-2 image, respectively. The serial number (SN) and the corresponding segment number (N) of each segmentation scale are presented in Table 2. To quantitatively evaluate the combination results, the reference areas are manually delineated for each land cover class in the images and shown in Fig. 4. Totally, there are 58 and 110 reference objects for the WorldView-2 and IKONOS-2 image, covering 15.9% and 13.9% of the whole image area, respectively. The measures of overall accuracy (OA) and Kappa coefficient in terms of all land cover classes are calculated to reveal the quality of combination results.

3.2. Sensitivity of segmentation scales

The segment tree is composed of multiscale segmentations. The sensitivity analysis in this part is to indicate how to construct an effective segment tree for the multiscale combination model. The combination weight ω in the energy function of Eq. (1) is set as the default value of 1 for the analysis.

We first analyze the influence of the fine-scale and coarse-scale segmentations in the tree to the combination results. To show the fine-scale influence, different finest segmentation scales are set given the same coarsest segmentation scale. For example, the SN of the coarsest scale ($SN_{coarsest}$) is given as 14 for the WorldView-2 image, and the SN of the finest scale SN_{finest} is set from 1 to 13. Then the segmentation scales between $SN_{coarsest}$ and SN_{finest} are all integrated in the segment tree. A similar strategy is used to show

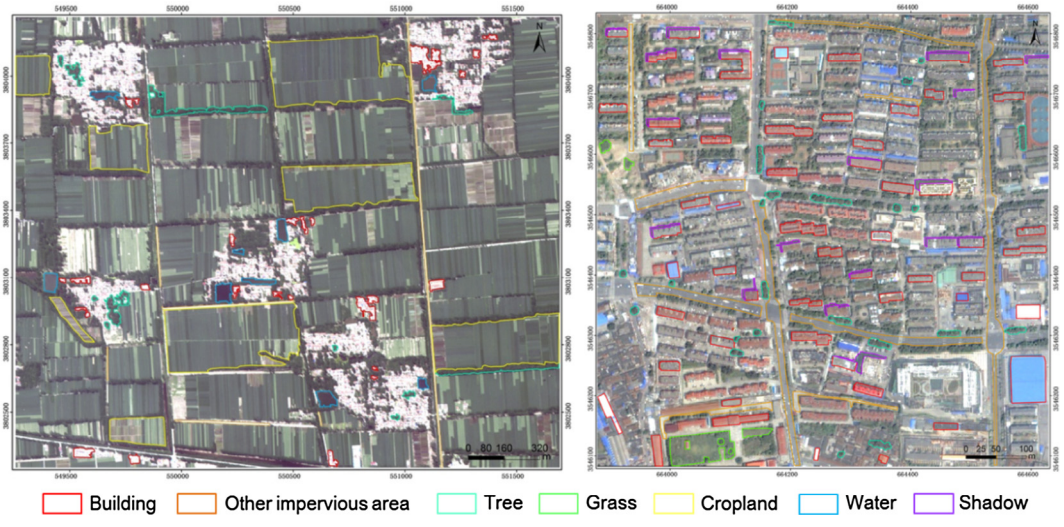

Fig. 4. Original experimental images (left: WorldView-2, right: IKONOS-2) and the reference objects. The images are shown with true color. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Details of the experimental HR images.

Satellite	Acquisition date	Size (pixel)	Spatial resolution	Spectral bands	Landscape	Location
WorldView-2	May 20, 2010	1217 × 1006	2 m	B, G, R, NIR	Rural	Xuzhou, China
IKONOS-2	June 18, 2009	750 × 750	1 m	B, G, R, NIR	Urban	Nanjing, China

Table 2
Produced hierarchical multiscale segmentations, where *SN* and *N* represent the serial number and the number of segments of each segmentation scale, respectively.

WorldView-2 image				IKONOS-2 image			
SN	*N*	*SN*	*N*	*SN*	*N*	*SN*	*N*
1	30,476	8	729	1	27,788	8	477
2	18,905	9	436	2	14,468	9	244
3	11,827	10	290	3	8457	10	139
4	7853	11	173	4	5008	11	86
5	4882	12	91	5	3044	12	36
6	2851	13	55	6	1650	13	25
7	1461	14	28	7	865		

the coarse-scale influence by setting different $SN_{coarsest}$ given the SN_{finest} as 1. The accuracies of the combination results based on the different segment trees are then evaluated. The results for the two test images are shown in Fig. 5, where each column shows the accuracy change by setting different $SN_{coarsest}$ and each row shows the accuracy change by setting different SN_{finest}. We can see from Fig. 5 that the accuracy change in each column is small, indicating the low sensitivity of the coarse-scale segmentations. While the accuracy presents apparent decrease in the first several rows when SN_{finest} is greater than a certain value, indicating the great sensitivity of the fine-scale segmentations if they are too coarse to represent the geographic objects.

To clearly show the accuracy change in Fig. 5, the accuracy values in the first row and the first column are presented as curves in Fig. 6. Given SN_{finest} as 1, the accuracy is slowly increased at first and then keeps steady as $SN_{coarsest}$ is getting large, which indicates that the coarse-scale segmentations in the segment tree should be coarse enough and too coarse segmentation scale would not greatly do harm to the combination results. However, given $SN_{coarsest}$, the accuracy first keeps steady and then decreases rapidly as SN_{finest} is increasing. Specifically, the smallest SN_{finest} leading to rapid decrease is 10 and 7 for the WorldView-2 and the IKONOS-2 image, respectively. From visual analysis we can find that the segmentation with *SN* greater than 10 for the WorldView-2 image and greater than 7 for the IKONOS-2 image appears apparent under-segmentation, as shown in Fig. 7. This indicates that the fine-scale segmentations in the segment tree should be fine enough and avoid being under-segmented.

Once the sensitivity of the SN_{finest} and $SN_{coarsest}$ has been understood, what is the influence of the internal segmentation scales to the combination results? To answer this question, the sensitivity of the scale intervals and the number of segmentation scales in the segment tree is then analyzed. To show the influence of the scale interval, a set of segment trees is produced for each scale interval with different number of segmentation scales. For example, to form the set of segment trees for the WorldView-2 image at scale interval of 3, the segment trees could be composed of segmentation scales {{1,4,7,10}, {2,5,8,11}, ..., {5,8,11,14}} when the number of segmentation scales is 4. The maximum and the medium accuracy values of the combination results from the set of segment trees for each scale interval are calculated and shown in Fig. 8. Generally, the influence of scale interval to the multiscale combination results is not great. The *OA* change range is restricted to 0.005 and 0.01 for the WorldView-2 and the IKONOS-2 image, respectively. Within the small change range of *OA*, it is preferred not to set the scale interval too large, which could decrease the combination accuracy, especially for the IKONOS-2 image in urban area.

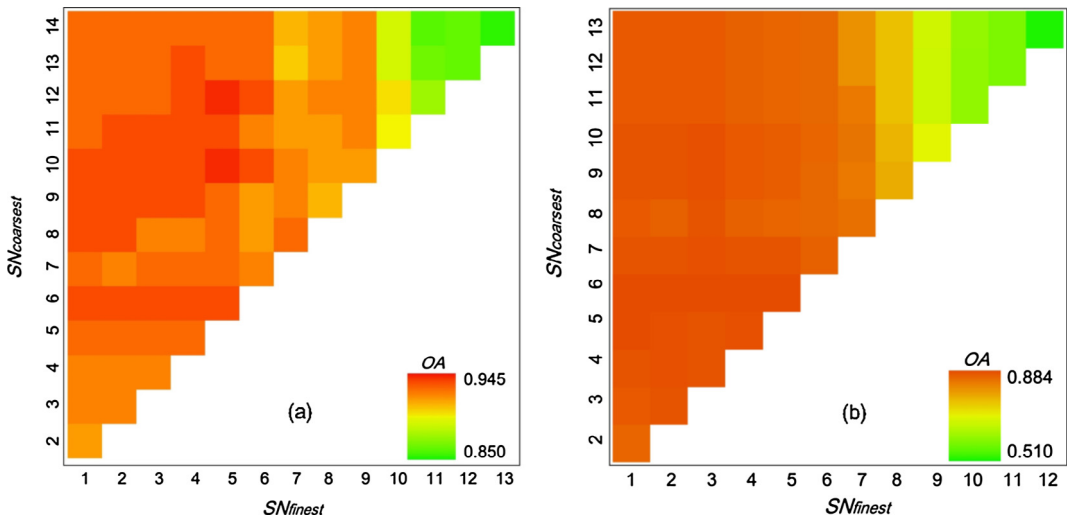

Fig. 5. Sensitivity of the coarseness of the segmentation scales for the proposed multiscale combination framework. (a) The WorldView-2 image, (b) the IKONOS-2 image.

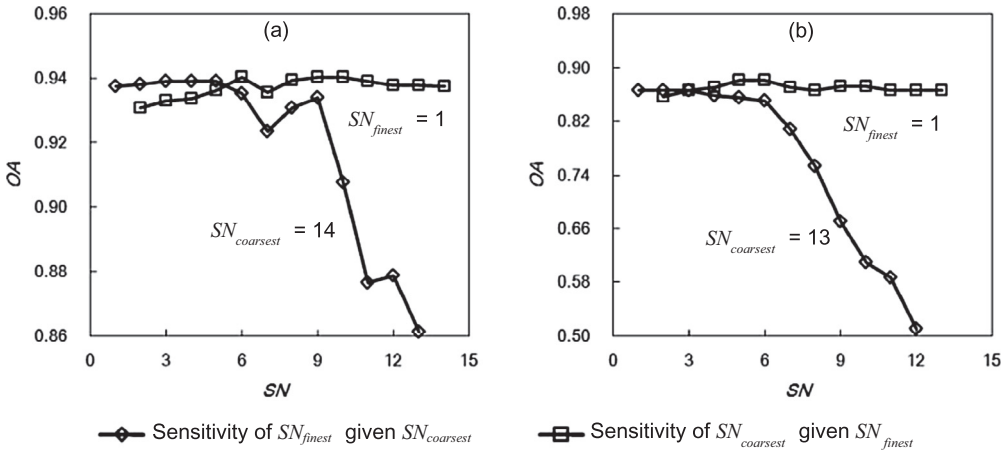

Fig. 6. Change of combination accuracy by setting different finest (SN_{finest}) and coarsest ($SN_{coarsest}$) segmentation scales for (a) the WorldView-2 image, (b) the IKONOS-2 image.

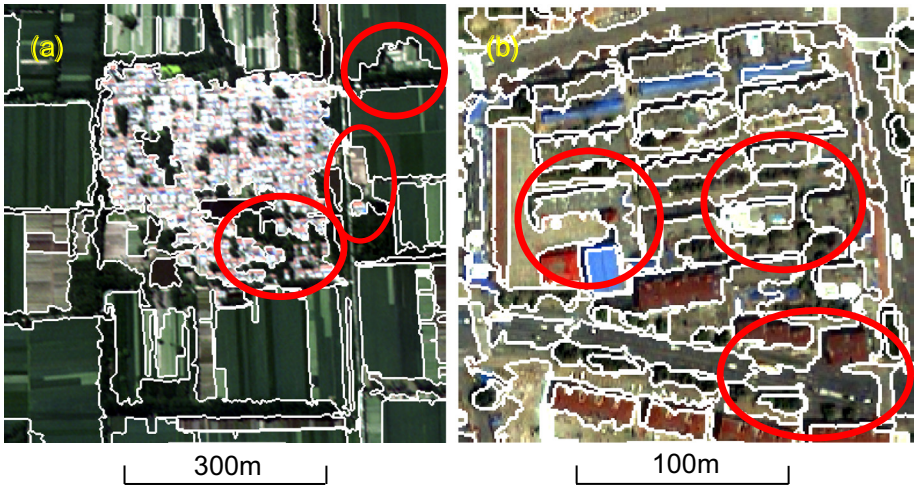

Fig. 7. Examples of under-segmentation in fine-scale segmentations that results in decrease of combination accuracy. The under-segmented regions are marked within the red areas. (a) Subset of the WorldView-2 image, $SN = 10$, $N = 290$; (b) subset of the IKONOS-2 image, $SN = 7$, $N = 865$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

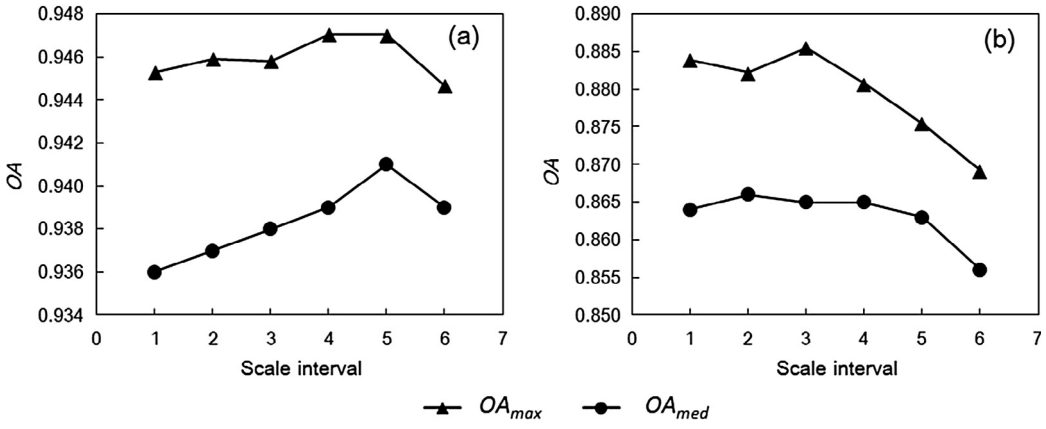

Fig. 8. Sensitivity of the scale interval for the proposed multiscale combination framework. (a) The WorldView-2 image, (b) the IKONOS-2 image.

To analyze the sensitivity of the number of segmentation scales, a set of segment trees is generated for each number of segmentation scales (*ns*) by selecting all possible *ns* segmentations from the produced segmentations in Table 2. For example, in order to generate the set of segment trees for the WorldView image given $ns = 2$, the number of trees in the set is $C_{14}^{2} = 91$. The maximum and the medium *OA* values are calculated for the segment tree set of each *ns* and presented in Fig. 9. Generally, the influence of

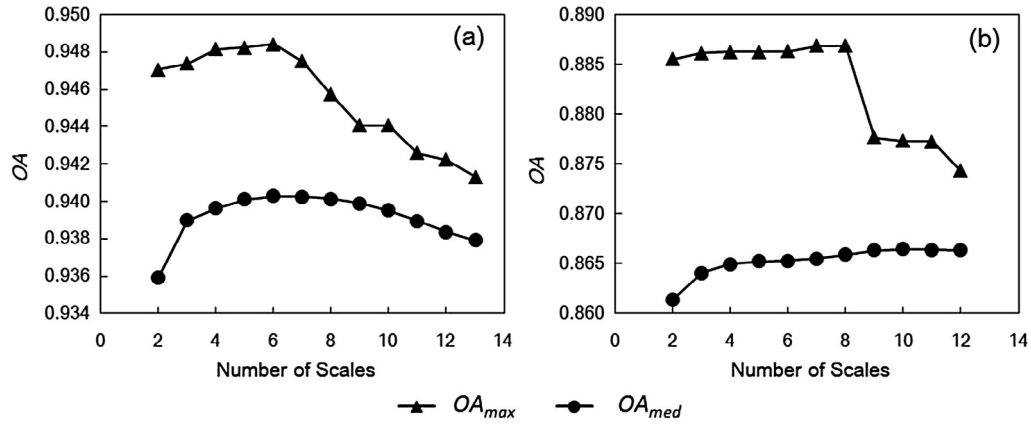

Fig. 9. Sensitivity of the number of segmentation scales for the proposed multiscale combination framework. (a) The WorldView-2 image, (b) the IKONOS-2 image.

scale number is also not great as the change range of accuracy is within 0.01 and 0.005 for the maximum and the medium *OA* values, respectively. In terms of the medium *OA* value, a small number of segmentation scales in the segment tree could result in relatively low accuracy. While in terms of the maximum *OA* value, the *OA* value slowly increases at first and then shows a relatively rapid decrease after reaching the highest value, which shows that it is not assured to achieving higher combination accuracy by integrating more segmentation scales.

3.3. Influence of combination weight

The combination weight ω in the energy function of Eq. (1) serves as a trade-off between the existed thematic information and the multiscale segmentations. Theoretically, a larger weight would result in a stronger effectiveness of the context in the multiscale segmentations, while a smaller weight would bring a weaker effectiveness of the context.

Given the scale interval as 1, a set of segmentation trees are generated with different numbers of segmentation scales. Different combination weights are applied to the set of trees and the maximum *OA* values for each combination weight are presented in Fig. 10. Within the given range of ω, the change of *OA* values can reach 0.015, indicating a relatively greater influence to the multiscale combination results than that of the scale interval (Fig. 8) and the scale number (Fig. 9). From Fig. 10, we can see that a medium ω value is preferred as both the small and large ω values could decrease the combination accuracy. To further show how the combination result is influenced by the context in multiscale segmentations, the change of *OA* values as the coarsening of the coarsest segmentation scale in the segment tree is presented in Fig. 11. It can be clearly seen that the accuracy differences caused by ω values are gradually enlarged as the increase of $SN_{coarsest}$. This shows the increase of sensitivity of the combination weight as the segmentation scale is getting coarse, revealing that if coarser segmentation scales are integrated in the segment tree, the combination weight needs to be set more carefully than that for finer cases.

A subset of the WorldView-2 image is zoomed into show the combination results by setting different ω values (Fig. 12). To further explain the influence of the context in segmentations, four multiscale segmentations in the same region are also presented in Fig. 12. As the ω value increases, the local details, e.g. isolated small objects, are merged into its adjacent regions resulting from the enhancement of the context in multiscale segmentations. If the ω value is too large (e.g. $\omega = 4$), the strong context influence, especially the under-segmented regions in the coarse segmentations, makes the combination result over-smoothed. While if the ω value is too small, the context would not contribute to the combination result. Hence, a proper ω value would be an important parameter to be set for the proposed multiscale combination procedure.

3.4. Comparisons

The multiscale combination framework is proposed to utilize the context information in multiscale segmentations, which consists of the spatial context by each segment and the hierarchical context by the multi-level segments in the segment tree. To highlight the effectiveness of the hierarchical context, the single-scale combination method is adopted for comparison, which combines

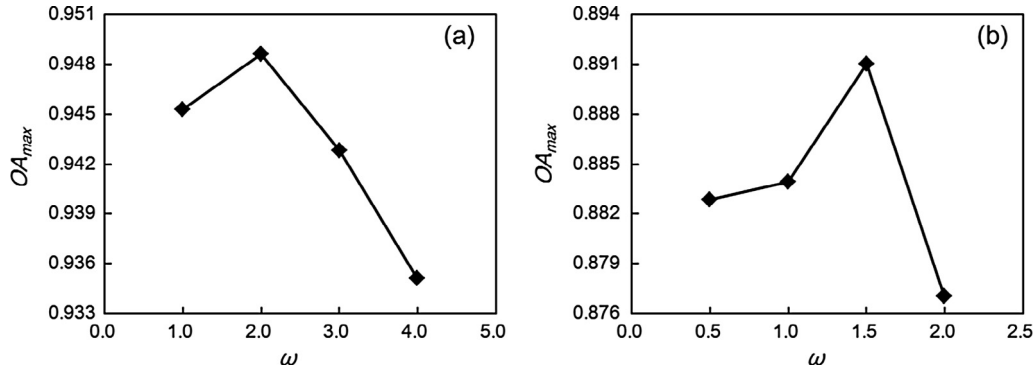

Fig. 10. Influence of the combination weight (ω) in the T-MRF energy function to the accuracy of the combination result. (a) The WorldView-2 image, (b) the IKONOS-2 image.

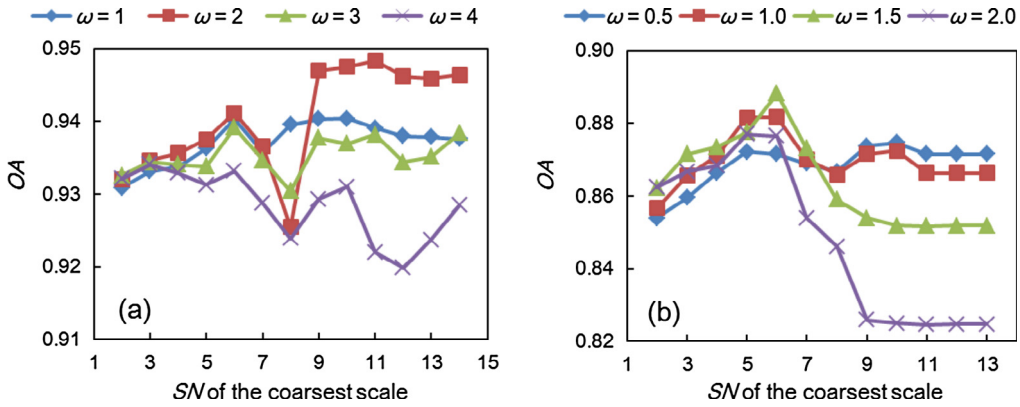

Fig. 11. Changes of *OA* with increased *SN* of the coarsest scale by setting different combination weight (ω) values, (a) the WorldView-2 image, (b) the IKONOS-2 image.

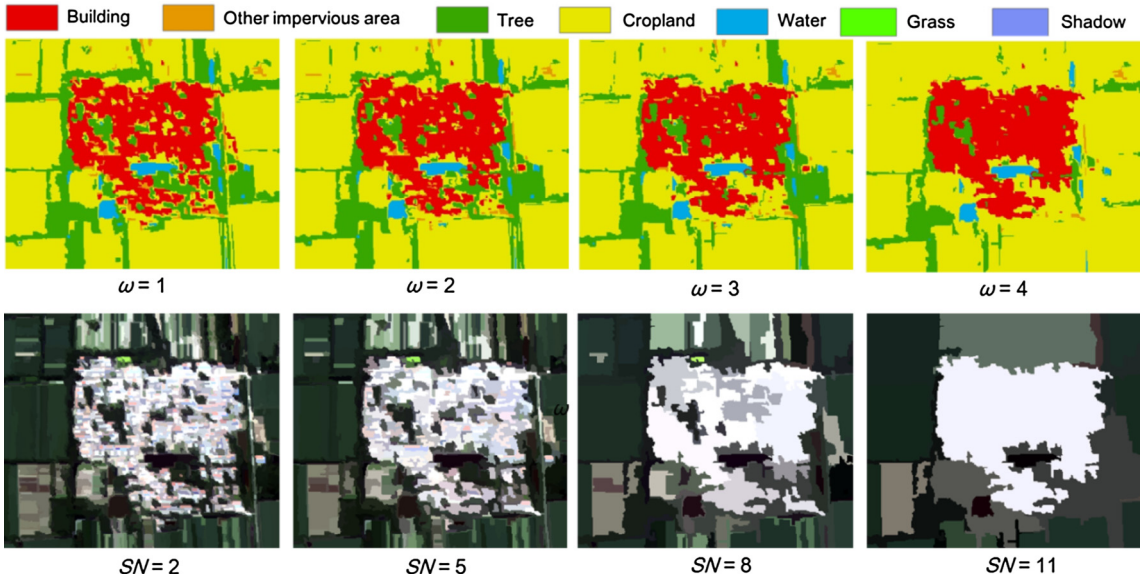

Fig. 12. Multiscale combination results of a subset of WorldView-2 image by setting different combination weight values. The first row presents the combination results and the second row shows the multiscale segmentation results in the combination procedure.

the thematic information with a single segmentation to refine the existed thematic information by exploring the spatial context in the single segmentation (Fauvel et al., 2013). The majority voting strategy is used for the single-scale combination.

The accuracy changes as the coarsening of segmentation scale are presented in Fig. 13. The single-scale combination result is highly dependent on the segmentation scale as its accuracy shows a slightly increase at first and then decreases rapidly as the segmentation is getting coarser. However, given the finest segmentation scale, the multiscale combination result shows no significant changes by adding coarser segmentation results in the segment tree. This further shows the benefit of utilizing hierarchical context

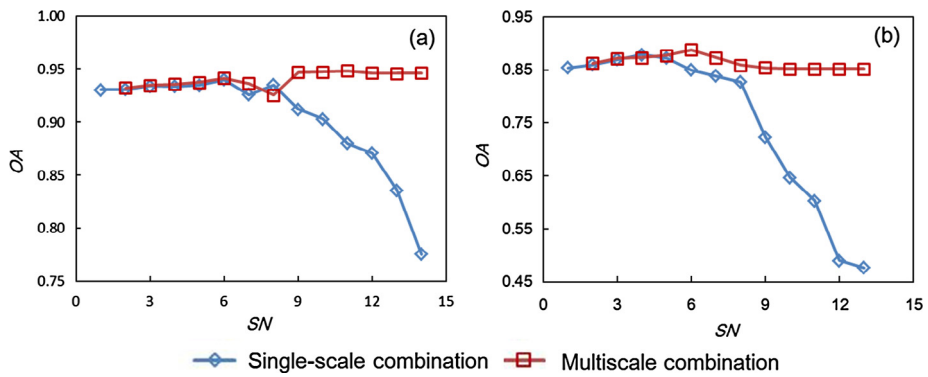

Fig. 13. Changes of the accuracy for the single-scale and the multiscale combination results with the increase of scale parameter, (a) the WorldView-2 image, (b) the IKONOS-2 image.

Table 3
Accuracies of the results with different combination methods.

Image	*OA*		
	Pixel-based classification	Single-scale combination	Multiscale combination
WorldView-2	0.907	0.940	0.949
IKONOS-2	0.802	0.879	0.891

to reduce the sensitivity on segmentation scales for the multiscale combination procedure.

It can also be seen in Fig. 13 that the highest accuracies between the single-scale and multiscale combination results are not very different. The highest accuracies for the two combination strategies are presented in Table 3. The pixel-based classification serves as the existed thematic information for the combination procedure. The great accuracy difference between the pixel-based

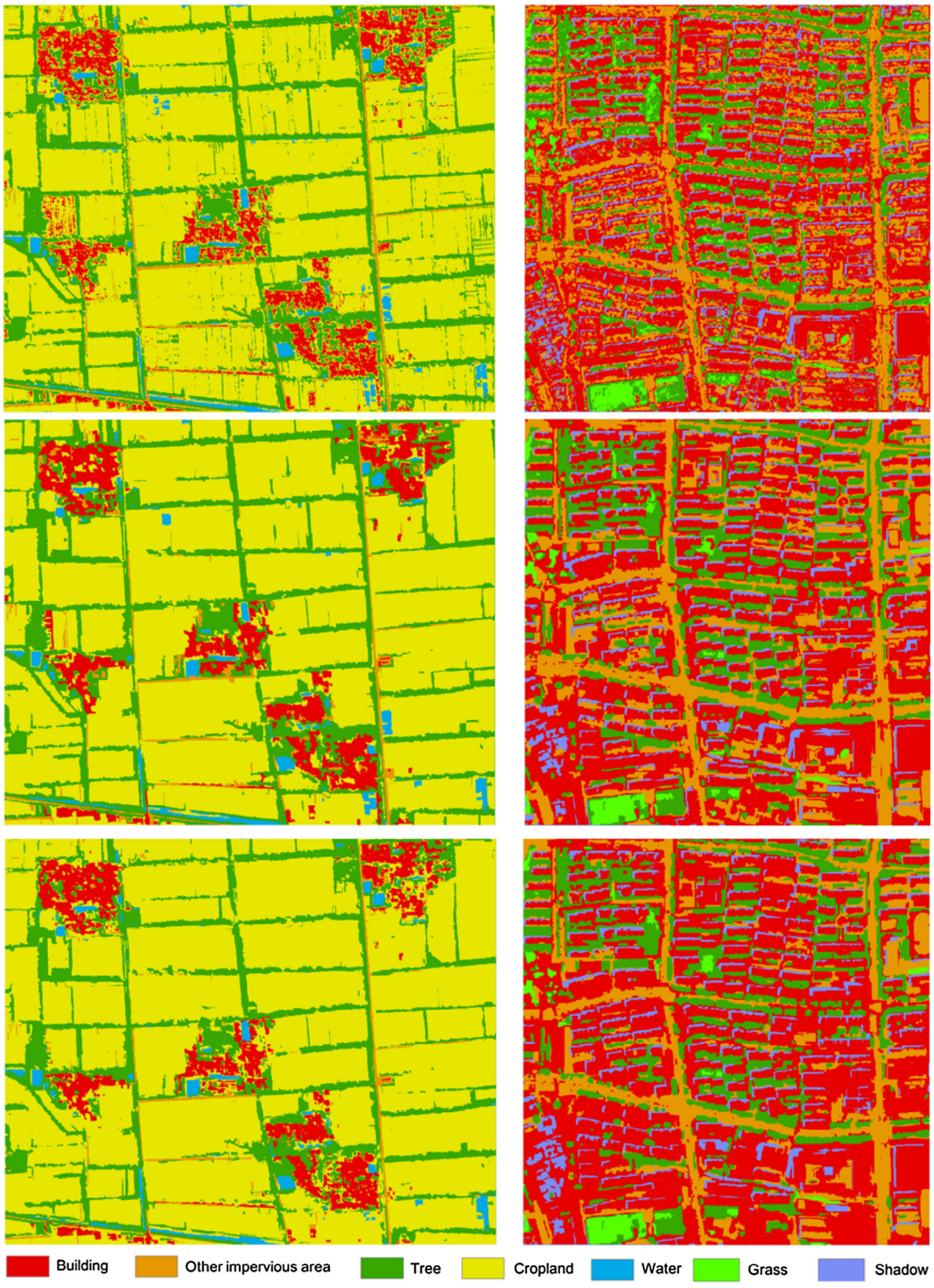

Fig. 14. Comparison of results with different combination methods. The pixel-based classification result, the single-scale combination result, and the multiscale combination result are shown in the first to third rows, respectively. Left: the WorldView-2 image, right: the IKONOS-2 image.

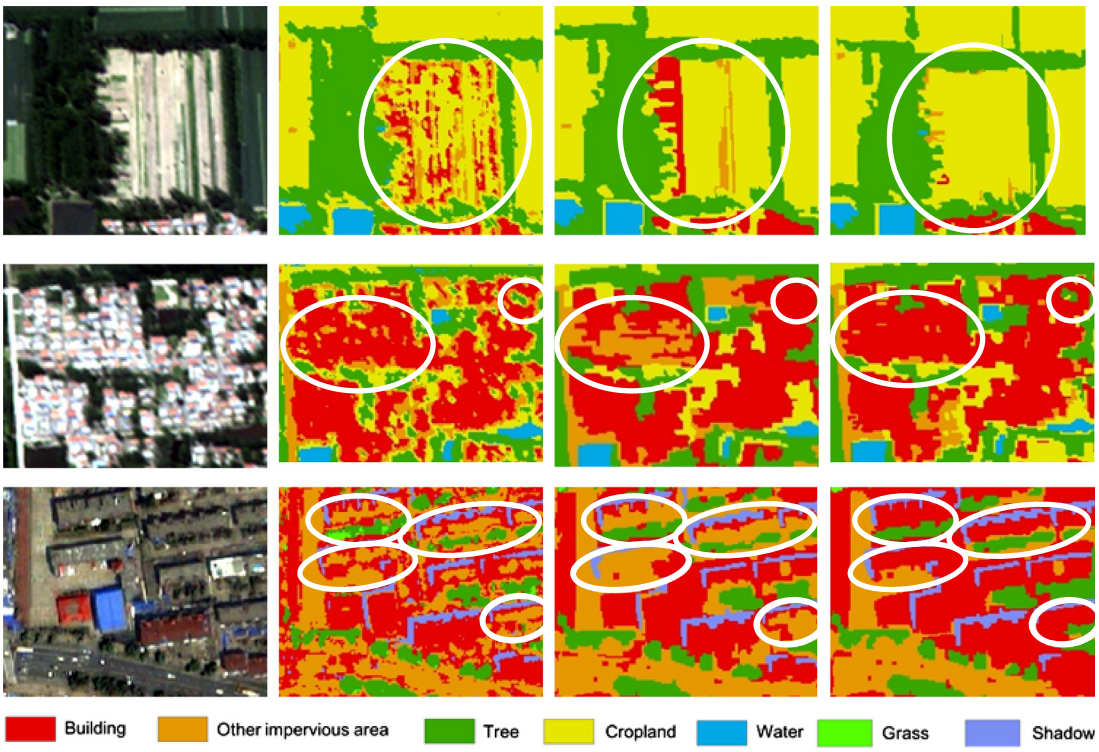

Fig. 15. Subsets of Fig. 14 are zoomed into show the difference. The original image, pixel-based classification result, single-scale combination result, and multiscale combination result are presented from left to right. The above two rows are subsets of the WorldView-2 image and the bottom row is a subset of the IKONOS-2 image.

classification and the single-scale combination indicates the effectiveness of the spatial context in the segmentation result. The accuracy difference between the single-scale and the multiscale combination results indicates the effectiveness of the hierarchical context in the multiscale segmentations, which shows weaker influence than that of the spatial context. We can also see that the WorldView-2 image achieves higher accuracy than that of the IKONOS-2 image, which should be resulted from the different land cover distributions. Specifically, the land cover classes of the rural area in the WorldView-2 image have apparently lower intra-class variability than those of the urban area in the IKONOS-2 image, as shown in Fig. 4.

The single-scale and multiscale combination results with the highest accuracy are presented in Fig. 14 to further show the difference. Comparing the pixel-based classification with the single-scale combination results, the "salt-and-pepper" phenomenon is apparently removed in the single-scale combination result. Comparing the single-scale with the multiscale combination results, the multiscale combination result tends to preserve more details and produce complete objects in certain regions. Three subsets from the two images are zoomed in and presented in Fig. 15 to clearly show the differences between the two combination strategies. In the first subset of the WorldView-2 image, the pixel-based classification results in many errors to classify the cropland as building or other impervious area because of the high spectral reflectance. The single-scale combination result produces a building object and an object marked as other impervious area because this cropland area is segmented into several regions in the single segmentation. However, the multiscale combination result produces a complete cropland object as the influence of the coarse segments in the segment tree. In the second subset of the WorldView-2 image, the under-segmentation in the single segmentation leads to the wrong labeling of buildings as other impervious area and the elimination of small tree objects surrounded by buildings. However, the multiscale combination result avoids the wrong labeling and preserves the details of the small tree objects by integrating fine-scale segmentations in the segment tree. In the third subset of the IKONOS-2 image, the single-scale combination result incorrectly labels several building objects as other impervious area because most pixels within the building objects are incorrectly classified as other impervious area in the pixel-based classification result, while the multiscale combination result can amend these wrong classifications by receiving the context information from coarse segments in the segment tree.

4. Discussions

A significant benefit of the proposed multiscale combination model is to reduce the sensitivity of segmentation scales. As demonstrated in the experiment (see Figs. 5 and 6), the addition of very coarse segmentations into the segment tree would not do harm to the combination results. The fine-scale segmentations in the segment tree would not negatively influence the combination results if they are over-segmented. Both the scale interval and the number of segmentation scales would not significantly influence the combination results (Figs. 8 and 9). The low sensitivity of segmentation scales of the proposed multiscale combination method is further demonstrated by comparing with the single-scale combination method (Fig. 13). Usually, the GEOBIA methods, e.g. object-based classification (Kim et al., 2009; Gao et al., 2011), single-scale combination (Zhang et al., 2013; Konstantinidis et al., 2017), and multiscale analysis (Myint et al., 2011; Dronova et al., 2012; Johnson and Xie, 2013), are sensitive to the selected segmentation scale(s). The proposed T-MRF model intrigues a new way to reduce the dependence of scale selection by integrated utilization of multiscale segmentations.

Another benefit of the proposed multiscale combination model is to utilize the hierarchical context within the multiscale segmentations. It has been evaluated that the spatial context in the seg-

mented regions can help remove the "salt-and-pepper" phenomenon to refine classifications by the studies of single-scale combination (Tarabalka et al., 2010; Fauvel et al., 2013). Even though the additional utilization of the hierarchical context does not result in an apparent increase of accuracy (Table 3), it can help to preserve detailed objects by the integration of the fine-scale segmentations in the segment tree and to generate complete objects by the received context from the coarse segments in the segment tree (Figs. 14 and 15). Furthermore, the low sensitivity of segmentation scales of the proposed multiscale combination model could result from the hierarchical context by exploring the saliency of the segments in each ancestry path. Nevertheless, the proposed model provides an explicit way to explore the hierarchical context, which is not illustrated by existed multiscale analysis studies (Bruzzone and Carlin, 2006; Moser et al., 2013). It deserves to further explore the benefit of the hierarchical context in GEOBIA.

The multiscale combination model aims to produce combination result as close as possible to the existed thematic information and at the same time determine the object boundary by utilizing the context information in multiscale segmentations. The weight parameter ω in the T-MRF energy function serves as a trade-off between the two aspects. A medium weight value is suggested to set according to the experimental results to achieve a proper balance (Fig. 12). The results also show that the influence of the combination weight parameter is relatively greater than that of the segmentation scales (Fig. 10), and needs to be carefully set especially at the case that too coarse segments are integrated in the segment tree. However, the interactions between the classification and segmentation results are not apparently presented by analysis of the combination weight, which needs to be further explored in the future to better fusing the thematic and context information.

The computational complexity should be an important issue for the proposed method. Since the segment tree model has no cycles, the global optimization of the T-MRF energy function can be achieved by a bottom-up message passing and followed by a top-down message passing. During the passing procedures, both the bottom-up and the top-down messages are calculated locally as shown in Fig. 2. Hence the computational complexity of the multiscale combination model is linear to the number of nodes in the segment tree. It is described as $O(2\beta N)$, where β denotes the computational complexity of calculating the local messages, N is the number of nodes in the segment tree.

The pixel-based classification is explored as an example to serve as the existed thematic information for the combination model. It is noted that the thematic information is not restricted to the classification results. Other thematic information, e.g. object detection results, thematic index thresholding results, can be utilized in the multiscale combination model for refinement. The benefits of the proposed model on refine other types of thematic information needs to be further demonstrated in the future.

5. Conclusions

A T-MRF model was explored to combine the thematic information and the multiscale segmentations in this study. The combination result is produced by minimizing the T-MRF energy function defined on a segment tree, achieving the trade-off between the existed thematic information and the context information to determine object types and boundaries. The advantage of the proposed model is to utilize the spatial and hierarchical context in multiscale segmentations to refine existed thematic information. The experimental results on HR images in both urban and rural areas demonstrate the benefits of the proposed combination model on reducing the sensitivity of segmentation scales and the explicit utilization of hierarchical context. The future works would focus on improving the multiscale combination model by further exploring the hierarchical context and demonstrating the effectiveness on other types of thematic information.

Acknowledgement

This work was supported by the China Postdoctoral Science Foundation (Grant No. 2016M590438), the National Natural Science Foundation of China (Grant No. 41601366), the Natural Science Foundation of Jiangsu Province (Grant No. BK20160623), and the Fundamental Research Funds for the Central Universities (Grant Nos. 020914380023, 020914380040). The authors would like to thank the reviewers and editors for the constructive comments to improve the paper.

References

Arbelaez, 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, M., 2000. Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation. In: Angewandte Geographische Informations – Verarbeitung XII. Wichmann Verlag, Karlsruhe, pp. 12–23.

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), 239–258.

Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E., et al., 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.

Bruzzone, L., Carlin, L., 2006. A multilevel context-based system for classification of very high spatial resolution images. IEEE Trans. Geosci. Remote Sens. 44 (9), 2587–2600.

Burnett, C., Blaschke, T., 2003. A multi-scale segmentation/object relationship modelling methodology for landscape analysis. Ecol. Model. 168 (3), 233–249.

Carleer, A.P., Debeir, O., Wolff, E., 2005. Assessment of very high spatial resolution satellite image segmentations. Photogrammet. Eng. Remote Sens. 71 (11), 1285–1294.

Chang, C.C., Lin, C.J., 2011. LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. 2 (3). Article 27.

Chen, J., Pan, D., Mao, Z., 2009. Image-object detectable in multiscale analysis on high-resolution remotely sensed imagery. Int. J. Remote Sens. 30 (14), 3585–3602.

Clinton, N., Holt, A., Scarborough, J., Yan, L., Gong, P., 2010. Accuracy assessment measures for object-based image segmentation goodness. Photogrammet. Eng. Remote Sens. 76 (3), 289–299.

Comaniciu, D., Meer, P., 2002. Mean shift: a robust approach toward feature space analysis. IEEE Trans. Pattern Anal. Mach. Intell. 24 (5), 603–619.

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.

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. Geograph. Inform. Sci. 24 (6), 859–871.

Drăguț, L., Csillik, O., Eisank, C., Tiede, D., 2014. Automated parameterisation for multi-scale image segmentation on multiple layers. ISPRS J. Photogrammet. Remote Sens. 88, 119–127.

Dronova, I., Gong, P., Clinton, N.E., Wang, L., Fu, W., Qi, S., Liu, Y., 2012. Landscape analysis of wetland plant functional types: the effects of image segmentation scale, vegetation classes and classification methods. Remote Sens. Environ. 127, 357–369.

Espindola, G.M., Câmara, G., Reis, I.A., Bins, L.S., Monteiro, A.M., 2006. Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation. Int. J. Remote Sens. 27 (14), 3035–3040.

Fauvel, M., Tarabalka, Y., Benediktsson, J.A., Chanussot, J., Tilton, J.C., 2013. Advances in spectral-spatial classification of hyperspectral images. Proc. IEEE 101 (3), 652–675.

Gao, Y., Mas, J.F., Kerle, N., Navarrete Pacheco, J.A., 2011. Optimal region growing segmentation and its effect on classification accuracy. Int. J. Remote Sens. 32 (13), 3747–3763.

Haralick, R.M., Shanmugam, K., Dinstein, I.H., 1973. Textural features for image classification. IEEE Trans. Syst., Man Cybernet. 6, 610–621.

Hay, G.J., Castilla, G., 2008. Geographic object-based image analysis (GEOBIA): a new name for a new discipline. Object-Based Image Anal., 75–89

Johnson, B., Xie, Z., 2011. Unsupervised image segmentation evaluation and refinement using a multi-scale approach. ISPRS J. Photogrammet. Remote Sens. 66 (4), 473–483.

Johnson, B., Xie, Z., 2013. Classifying a high resolution image of an urban area using super-object information. ISPRS J. Photogrammet. Remote Sens. 83, 40–49.

Kim, M., Madden, M., Warner, T., 2009. Forest type mapping using object-specific texture measures from multispectral Ikonos imagery: segmentation quality and image classification issues. Photogrammet. Eng. Remote Sens. 75 (7), 819–830.
Kim, M., Warner, T., Madden, M., Atkinson, D., 2011. Multi-scale GEOBIA with very high spatial resolution digital aerial imagery: scale, texture and image objects. Int. J. Remote Sens. 32 (10), 2825–2850.
Konstantinidis, D., Stathaki, T., Argyriou, V., Grammalidis, N., 2017. Building detection using enhanced HOG–LBP features and region refinement processes. IEEE J. Sel. Top. Appl. Earth Obser. Remote Sens. 10 (3), 888–1005.
Kschischang, F., Frey, B., Loeliger, H., 2001. Factor graphs and the sum-product algorithm. IEEE Trans. Inf. Theory 47, 498–519.
Laliberte, A.S., Rango, A., 2009. Texture and scale in object-based analysis of subdecimeter resolution unmanned aerial vehicle (UAV) imagery. IEEE Trans. Geosci. Remote Sens. 47 (3), 761–770.
Lim, J.J., Arbeláez, P., Gu, C., Malik, J., 2009. Context by region ancestry. In: 12th IEEE International Conference on Computer Vision (ICCV), pp. 1978–1985.
Liu, D., Xia, F., 2010. Assessing object-based classification: advantages and limitations. Remote Sens. Lett. 1 (4), 187–194.
Liu, Y., Bian, L., Meng, Y., Wang, H., Zhang, S., Yang, Y., Shao, X., Wang, B., 2012. Discrepancy measures for selecting optimal combination of parameter values in object-based image analysis. ISPRS J. Photogrammet. Remote Sens. 68, 144–156.
Möller, M., Lymburner, L., Volk, M., 2007. The comparison index: a tool for assessing the accuracy of image segmentation. Int. J. Appl. Earth Obs. Geoinf. 9 (3), 311–321.
Moser, G., Serpico, S.B., Benediktsson, J.A., 2013. Land-cover mapping by Markov modeling of spatial–contextual information in very-high-resolution remote sensing images. Proc. IEEE 101 (3), 631–651.
Munoz, D., Bagnell, J.A., Hebert, M., 2010. Stacked hierarchical labeling. In: European Conference on Computer Vision (ECCV). Springer, Berlin Heidelberg, pp. 57–70.
Myint, S.W., Gober, 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.
Ren, X., Bo, L., Fox, D., 2012. RGB-(D) scene labeling: features and algorithms. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2759–2766.
Salembier, P., Garrido, L., 2000. Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. IEEE Trans. Image Process. 9 (4), 561–576.
Smith, A., 2010. Image segmentation scale parameter optimization and land cover classification using the Random Forest algorithm. J. Spat. Sci. 55 (1), 69–79.
Tarabalka, Y., Chanussot, J., Benediktsson, J.A., 2010. Segmentation and classification of hyperspectral images using watershed transformation. Pattern Recogn. 43 (7), 2367–2379.
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 Obser. Remote Sens. 5 (1), 262–272.
Thomas, N., Hendrix, C., Congalton, R., 2003. A comparison of urban mapping methods using high-resolution digital imagery. Photogrammet. Eng. Remote Sens. 69 (9), 963–972.
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. Photogrammet. Remote Sens. 87, 108–121.
Yang, D., Chen, X., Chen, J., Cao, X., 2017a. Multiscale integration approach for land cover classification based on minimal entropy of posterior probability. IEEE J. Sel. Top. Appl. Earth Obser. Remote Sens. 10 (3), 1105–1116.
Yang, J., He, Y., Caspersen, J., 2017b. Region merging using local spectral angle thresholds: a more accurate method for hybrid segmentation of remote sensing images. Remote Sens. Environ. 190, 137–148.
Yang, J., Li, P., He, Y., 2014. A multi-band approach to unsupervised scale parameter selection for multi-scale image segmentation. ISPRS J. Photogrammet. Remote Sens. 94, 13–24.
Yu, Q., Gong, P., Clinton, N., Biging, G., Kelly, M., Schirokauer, D., 2006. Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. Photogrammet. Eng. Remote Sens. 72, 799–811.
Zhang, X., Xiao, P., Feng, X., 2013. Impervious surface extraction from high-resolution satellite image using pixel-and object-based hybrid analysis. Int. J. Remote Sens. 34 (12), 4449–4465.
Zhang, X., Xiao, P., Feng, X., Wang, J., Wang, Z., 2014. Hybrid region merging method for segmentation of high-resolution remote sensing image. ISPRS J. Photogrammet. Remote Sens. 98, 19–28.
Zhang, X., Feng, X., Xiao, P., 2015. Multi-scale segmentation of high-spatial resolution remote sensing images using adaptively increased scale parameter. Photogrammet. Eng. Remote Sens. 82 (6), 11–20.