IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Local scale-guided hierarchical region merging and further over- and under-segmentation processing for hybrid remote sensing image segmentation

Yongji Wang[1], Lili Wu[1], Qingwen Qi[2,3], and Jun Wang[4]

[1] School of Geoscience and Technology, Zhengzhou University, Zhengzhou, 450001 China
[2] State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, 100101 China
[3] University of Chinese Academy of Sciences, Beijing, 100101 China
[4] College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao, 266590 China

Corresponding author: Lili Wu (e-mail: wll_dqkxxy@zzu.edu.cn).

This work was supported by the key scientific research projects of colleges and universities in Henan Province under Grant 22A170021 and the National Natural Science Foundation of China under Grant 41801085.

ABSTRACT With the development of medium- and high-resolution satellites, successfully segmenting differently sized geo-objects remains a challenging issue in the framework of geographic object-based image analysis (GEOBIA). The hybrid image segmentation method is a good alternative to produce good segmentation that best matched the different sizes of geo-objects. However, the existing methods almost use a certain scale or other segmentation parameters (SPs) to control the sizes and shapes of segments. This will lead to two issues: (1) one single scale is impossible to segment every geo-object well due to the land cover complexity within remote-sensing imageries; (2) over- and under-segmented regions still occur in the segmentation results, whatever using any advanced segmentation methods. To solve the above problems, this paper developed a hybrid image segmentation method with local scale-guided hierarchical region merging and further over- and under-segmentation processing. First, the primitive segmentation was produced and then stratified into layers with different land covers. Then, the local scale was calculated for a more objective merging process in the separating layers. Third, the over- and under-segmentation at separating layers was recognized and re-processed for achieving a fine segmentation. To validate the proposed method, it was applied to three test images of gaofen-1 satellite with different land cover types, and ten competing methods were compared. The visual and quantitative results indicated the advantage of our method in segmenting out different sizes of geo-objects, which can effectively reduce the over- and under-segmentation error.

INDEX TERMS geographic object-based image analysis (GEOBIA), hybrid image segmentation, local scale, hierarchical region merging, over- and under-segmentation recognition and re-process.

I. INTRODUCTION

With the increasing availability of detailed geo-object information in medium- and high- resolution remote sensing images, geographic object-based image analysis (GEOBIA) [1, 2] has become a new paradigm in extracting various geo-objects from these data, such as farmland division [3-6], building detection [7, 8], coastal geo-object recognition [9, 10] and change detection [11, 12] . The advantages of the low spectral variation within geo-objects and the full utilization of textural features and shape concepts of geo-objects make GEOBIA yield more accurate image classification than the traditional pixel-based analysis method [13-15]. GEOBIA mainly consists of three processes: image segmentation, feature extraction including spectral, textural, and spatial information, and segment classification [16-19]. Image segmentation is implemented to generate segments and these segments are considered the geo-objects, which are the foundation of subsequent feature

extraction and image classification. Thus, the selection of an appropriate image segmentation method is very critical for generating good segments, which can make the best match with the real geo-objects.

The function of image segmentation is to divide a remote sensing image into spatially heterogeneous and spectrally homogeneous regions [20, 21]. Most existent segmentation methods only consider the boundary information, such as the edge-based method [22-25], or the spatial information, such as the region-based method [26-34]. The edge-based method determines the edge for an image by tracking the pixel values that are discontinuous at different boundary regions [24], but it always tends towards over-segmentation. The region-based method aggregates or merges similar pixels by calculating the similarity between adjacent pixels restricted by some criterion [33], but it is likely to produce an under-segmented result, and is time-consuming with too much iteration calculation.

The hybrid segmentation method employing a split- and -merge strategy is a good alternative to solve the aforementioned problem. Such method is a two-stage technology, which first uses the edge-based method to produce over-segmented results and then applies the region-based method to merge similar segments for the best match with the real geo-objects [35]. The hybrid method takes full use of the boundary information and spatial information. And it searches for the adjacent segment pairs, but not the adjacent pixel pairs, thus running faster than the traditional region-based method and can be affected by the local structure. Recently, some scholars have paid more attention to such methods for improving the segmentation accuracy [29, 32, 35-40].

In the frame of hybrid segmentation, over-segmentation is regarded as a desirable starting and the edge-based method tends towards over-segmentation. Hence, an appropriate merging method is very critical for generating satisfying segmentation results. Following the existent literature, the merging criteria and merging order are the main focus in the merging algorithm design and optimization [29, 30, 34, 41-43]. However, the sizes of segments are controlled by scale or other segmentation parameters (SPs) in these merging algorithms. Different scales will lead to different results, in which small scale tends towards over-segmentation and large scale tends towards under-segmentation. In addition, one single scale is impossible to segment every geo-object well due to the land cover complexity within remote-sensing imageries.

There are two alternatives for solving the above problems. The first approach is the multiple optimal scale (or SP) determination method [31, 44-47]. When the region merging process is initialized by an image with all types of geo-objects, the multiple scales (or SPs) obtained can generate satisfying results that are best matched with differently sized geo-objects, but studying the approach that these results are combined into one ultimate segmentation layer is difficult. Note that different types of land covers have their unique characteristics in geo-object size, shape, and spectral variability. For example, the forest is generally largely sized and obviously spectral variable, and the residential buildings are usually regularly shaped, little sized and lowly spectral changed. It would be a good choice to separate different types of land covers into different layers before the region merging process is executed. First, this will avoid the confusion that different types of land covers are wrongly merged. Second, different optimal scales (or SPs) can participate in controlling the merging process, achieving the multiple scale (or SP) determination. Third, the merging costs are calculated with the adjacent segments with similar characteristics, enhancing the objectivity and effectiveness of the merging criteria.

The second approach is the local scale (or SP) optimization method [36, 48, 49], which adjusts the global scale (or SP) based on local spectral structure. For example, in the recent works of Yang et al. (2017) [36] and Su (2019) [48], they think that homogeneous adjacent segments are more possible to be one geo-object, thus a relatively high scale (or SP) is required for merging one adjacent segment pair with low spectral variance (high homogeneity). But Yang et al. (2017) calculated the homogeneity only using one segment, ignoring the impact of adjacent segments. Su (2019) further improved the approach by considering the homogeneity of the adjacent segments and introducing one parameter ρ to affect the extension rate of local SP. However, Su (2019) neglected that the region merging process just happened in two adjacent segments, not all the image. Thus, the impact of spectral variance change after one pair of adjacent segments are merged is required to be considered to adjust the local scale (or SP) for a more objective merging process.

Moreover, the over- and under-segmented regions still occur in the result, after the operation of the above processing. To further improve the segmentation quality, the over- and under-segmented regions need to be identified and then further processed. Most methods only evaluate the over- and under-segmentation in a global manner [17, 50-53]. In other words, these methods can indicate segmentation quality but not be helpful for further improving the segmentation results. Reasonably, evaluation metrics should remain within the object-based paradigm for effectively identifying over- and under-segmented regions. Recently, only limited corresponding studies were reported [54, 55]. It will be interesting to study a new and simple over- and under-segmentation recognition method from the perspective of segment.

In this paper, a hybrid remote sensing image segmentation method with local scale-guided hierarchical region merging and further over- and under-segmentation processing is developed. The proposed method skillfully integrates multiple optimal scale (or SP) determination and the local scale (or SP) approach, and then recognizes and

processes the over- and under- segmented regions for generating more satisfying segmentation. More specifically, the main contributions of this paper constitute of three aspects:

1) The primitive segmentation produced by the edge-based method is rapidly divided into two layers with different land covers using a simple vegetation index. It would help achieve the multiple scale (or SP) determination and forming the ultimate segmentation layer.

2) The spectral variance change after one pair of adjacent segments are merged is considered to calculate the local scale (or SP) for a more objective merging process in the separating layers.

3) The over- and under-segmented regions at separating layers were recognized using the coefficient of variation (CV), and then re-processed to achieve a fine segmentation. The CV is a comprehensive indicator that can reveal the level of both over- and under-segmentation.

The remaining of this paper is designed as follows. The proposed approach is detailly introduced in Section II, followed by the experiment results in Section III. The discussion and conclusions are given in Section IV and VI, respectively.

II. METHODS

The proposed hybrid remote-sensing image segmentation method can be divided into three parts: (1) the primitive segmentation production and stratification, (2) the local scale-guided region merging at separating layers, and (3) the over- and under-segmented regions recognition and re-process at separating layers. A general flowchart of the proposed approach is shown in Figure 1.

A. PRIMITIVE SEGMENTATION PRODUCTION AND STRATIFICATION

Over-segmentation is regarded as a desirable starting for the hybrid remote sensing image segmentation. The watershed transform is used to generate over-segmented regions (the primitive segmentation). Then, the Normalized Difference Vegetation Index (NDVI), which is a Simple and easily calculated vegetation index, is employed to generate the vegetation map for separating the primitive segmentation into two layers with vegetative segments and non-vegetative segments.

1) PRIMITIVE SEGMENTATION GENERATION

The watershed transformation [56] is designed to detect edges controlled by the spectral distance between adjacent segments. And most scholars applied the method after a gradient image is calculated from a panchromatic image or a single band of one multi-spectral image [57]. However, the edge information contained in multi-spectral bands should be taken full use for detecting more accurate edges with the increasing availability of multi-spectral satellites [36, 58]. In this paper, a gradient weighted by multi-spectral bands are computed and then the watershed transformation is implemented to produce over-segmented results for subsequent merging.

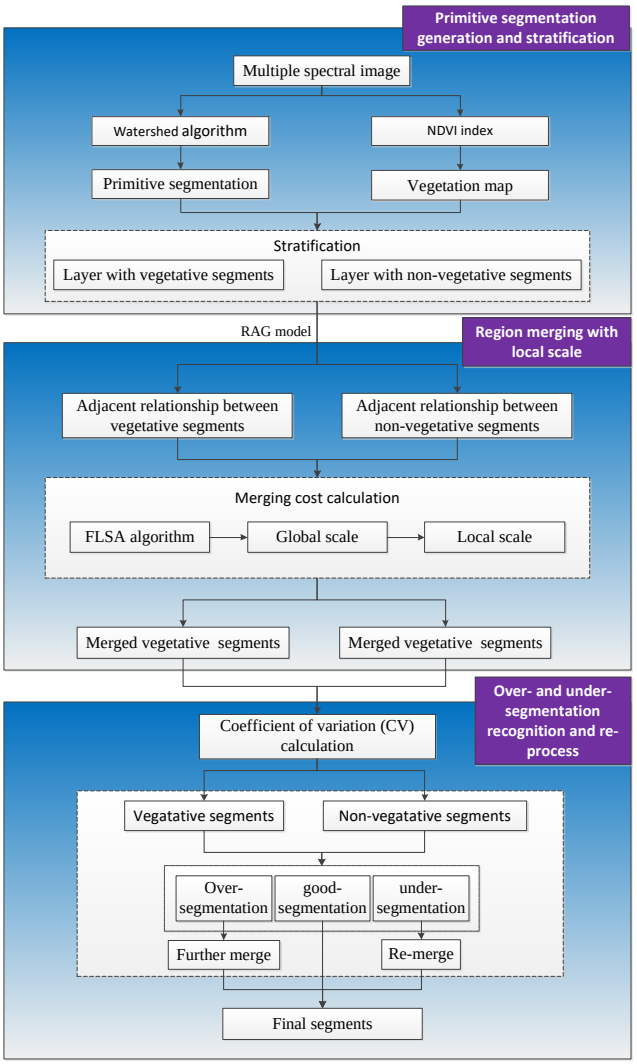


FIGURE 1. The schematic of the proposed hybrid remote sensing image segmentation method with local scale-guided hierarchical region merging and further over- and under-segmentation processing.

2) PRIMITIVE SEGMENTATION STRATIFICATION

The NDVI is a normalized ratio of the NIR and red bands, as follows:

$$\mathrm{NDVI} = \frac{b_{\mathrm{NIR}} - b_{\mathrm{red}}}{b_{\mathrm{NIR}} + b_{\mathrm{red}}} \quad (1)$$

where b_{NIR} and b_{red} are the spectral reflectance of NIR and red bands obtained in a remote sensing image, respectively. The NDVI is a vegetation measure that can reflect the seasonal and inter-annual changes in vegetation growth, and the greater NDVI value indicates higher vegetation coverage.

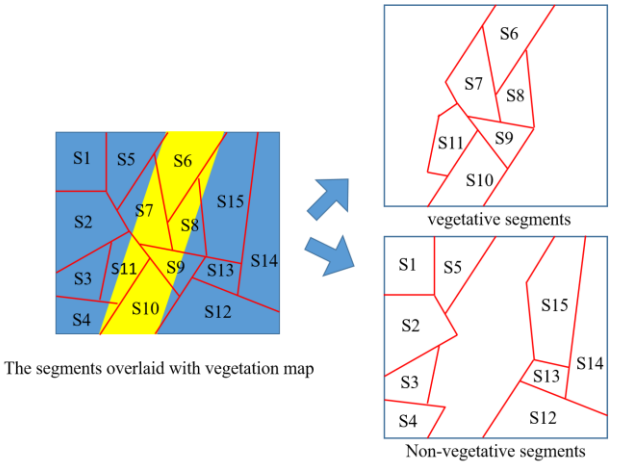


FIGURE 2. Illustration of primitive segmentation stratification. For the left part, the primitive segmentation is overlaid with the vegetation map; for the right part, the top image is the selected vegetative segments and the bottom is the selected non-vegetative segments.

Then, the NDVI is calculated to generate the vegetation map, and the vegetation map is used to separate the primitive segmentation into two layers with vegetative segments and non-vegetative segments. The separating process can be illustrated in Figure 2.

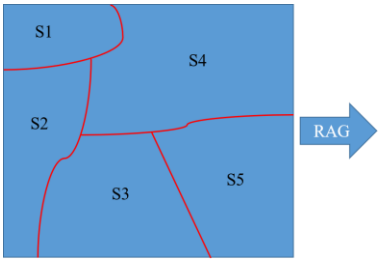
Assuming that S1-S15 are the segments from the primitive segmentation produced by the watershed transformation, and the yellow part in the vegetation map represents vegetation, and the blue part represents the non-vegetation. It is easy to classify one segment when it is completely located in vegetation or non-vegetation part. However, when a segment covers both vegetation and non-vegetation parts, such as S7 and S11, it is a real issue to determine the classification of the segment. The overlapping area metric is used to solve the problem. If the vegetation has the largest overlapping in this segment, the segment will be classified as a vegetative segment. For example, over half areas of S7-S11 are all located in the vegetation part, so they belong to vegetative segments.

B. LOCAL SCALE-GUIDED REGION MERGING

After the operation of primitive segmentation production and stratification, the full lambda-schedule algorithm (FLSA) [27] is used to calculate the merging cost, and the global scale is determined by the parameter optimization method of Johnson et al. (2015) [44]. Then, the local scale is adjusted by considering the spectral variance change after one pair of adjacent segments is merged. Finally, the iterative merging process continues until the stopping rule is satisfied, and the merging results at separating layers are exported.

1) GLOBAL SCALE DETERMINATION

The graph models of RAG are built with separating layers with vegetation segments and non-vegetation segments to represent the adjacent relationship between segments, with



node	arcs
S1	S2
S1	S4
S2	S3
S2	S4
S3	S4
S3	S5
S4	S5

FIGURE 3. Illustration of converting one segmentation into RAG.

node and arcs indicating segments and their adjacent segments [37]. Figure 3 illustrates an example. It is difficult to quantitatively recognize the adjacent segments in the left segmentation result. However, the adjacent segment pairs can be easy to be searched after the RAG is generated, as shown in the right table. Then, the merging cost between adjacent segments is computed using the FLSA, defined as follows:

$$\mathrm{Lambda}_{S_1,S2} = \frac{A_{S_1} \cdot A_{S_2}}{A_{S_1} + A_{S_2}} \cdot \frac{E_{S_1,S_2}}{L_{S_1,S_2}} \tag{2}$$

where S_1 and S_2 are one pair of the adjacent segments, A_{S_1} and A_{S_2} are the areas of S_1 and S_2, which are replaced by the pixel number in this paper. E_{S_1,S_2} is the Euclidean distance between S_1 and S_2, and L_{S_1,S_2} is the common boundary between S_1 and S_2. To take full advantage of the spectral information in multiple bands, this paper calculates the FLSA averaged by multiple bands as the ultimate merging cost. The adjacent segment pairs will be merged into one segment if their merging cost is less than the global scale (t). The cumulative probability analysis adopted in this paper can determine t, as follows:

$$F(x_\beta) = P(x \le t) = \beta \quad (0 < \beta < 1) \tag{3}$$

where x is one of all the merging costs between adjacent segments, P is the cumulative probability in case that $x \le t$, and β is the merging parameter, which controls the merging process. The process of determining t is as follows: First, all the merging costs between adjacent segments are sorted from low value to high value, and then put to a new list with a size of $N \times 1$. N is the number of all the adjacent segment pairs. Then, the location in the new list is searched by calculating the product of N and β, i.e., M, and the merging cost located in (M, 1) in this list is recognized as the t. Specifically, the optimal global scales (i.e., t) at the separating layers are determined using the approach of Johnson et al. (2015), which uses two indexes of the area-weighted variance (WV) and Moran's I and applies the F-measure strategy to combine the two indexes for evaluating the segmentation quality.

2) REGION MERGING WITH LOCAL SCALE

The global scales at the separating layers obtained are not likely to guarantee to delineate every geo-object well. To produce more objective segmentation, this paper develops a local scale determination method, which considers the spectral variance change after one pair of adjacent segments is merged. The proposed approach can be defined as:

$$t_{local} = \frac{v_{S_3}}{v_{S_1+S_2}} \cdot t \tag{4}$$

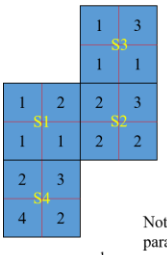
where t is the global scale, v_{S_3} is the variance after the adjacent segments pair of S_1 and S_2 is merged (i.e., S_3); $v_{S_1+S_2}$ is the relative variance of S_1 and S_2, as defined in Eq. (5),

$$v_{S_1+S_2} = \frac{A_{S_1} \cdot v_{S_1} + A_{S_2} \cdot v_{S_2}}{A_{S_1} + A_{S_2}} \tag{5}$$

where v_{S_1} and v_{S_2} is the variance of S_1 and S_2, respectively.

Then an example is presented in Figure 4 to illustrate the proposed approach. The pixel values within segments and the adjacent relationship between segments are shown in the left segment example. The top table presents the variance statistics of every segment and merged adjacent segments. The bottom presents the weight of three pairs of adjacent segments, which is a coefficient that locally adjusts the global scale, such as Eq. (4) in our article and Eq. (1) in Su (2019). S1 and S2 have the lowest variance (both are 0.25), indicating that S1 and S2 are more likely to be part of one geo-object. Both of the proposed approach and Su's approach both put larger weights into the global scale for obtaining a larger local scale. With region merging processing, if homogeneity of segments is decreased, it indicates that over-segmentation is reduced. When one segment is under-segmented, if it is merged with its neighboring segment, the homogeneity will be increased. In this case, the adjacent segments should not be merged and it should be put less weight into the global scale. After the two pairs of adjacent segments (S1-S2 and S1-S4) are merged, the variance increased (S1: 0.25 and S2: 0.25 vs. S1+S2: 0.5, and S1:0.25 and S4: 0.9167 vs. S1+S4: 1.1429), indicating decreased over-segmentation. Then larger weight should be put for larger local scale. The adjacent segments of S2 and S3 have a decreased variance (S2: 0.25 and S3:1 vs. S2+S3:0.6964) after they are merged, indicating that the merging processing should be banned. Then less weight is required for controlling the global scale. The proposed approach puts larger weights in S1-S2 and S1-S4 and less weight in S2-S3, however, the approach of Su (2019) puts similar weights in S1-S4 and S2-S3. This indicates that the spectral variance change after one pair of adjacent segments is merged is also critical for obtaining a more objective local scale estimation.

The process of the region merging with local scale is as follows: All pairs of adjacent segments are corresponding to their respective merging cost, calculated by Eq. (2), and



example

Segment	S1	S2	S3	S4	S1+S2	S1+S4	S2+S3
Variance	0.25	0.25	1	0.9167	0.5	1.1429	0.6964

Variance statistics

Adjacent segments	Weight (proposed approach)	Weight (approach of Su(2019))
S1-S2	2	2.4168
S1-S4	1.9592	1.0357
S2-S3	1.1143	0.9667

Note that the weight is the coefficient that adjusts the global segmentation parameter (Eq.(4) in our article and Eq. (1) in Su (2019)).

Comparison of different approaches

FIGURE 4. Illustration of converting global scale into local scale.

TABLE I
REGION MERGING WITH LOCAL SCALE

Input: original image, vegetative segments/non-vegetative segments, β

Procedure
(1) Create a RAG based on vegetative segments/non-vegetative segments, and calculate the merging cost using Eq. (2);
(2) Calculate the global scale t by Eq. (3), and then compute t_{local} of each pair of adjacent segments using Eq. (4);
(3) Search the segment-pair with lowest merging cost, and merge them if the merging cost is less than t_{local};
(4) Update RAG and then calculate the merging cost and t_{local} with new formed segment-pair;
(5) Repeat (3) and (4) until no segment-pair is merged.
End

Output: merged vegetative segments/ non-vegetative segments

then put to a new list with a size of $N \times 3$. The first two columns store one segment and its neighboring segment, and the third column stores the corresponding merging cost. The segment-pair with the lowest merging cost is first searched and if the merging cost is less than t_{local}, the segment-pair will be merged. Then the RAG is updated by adding the newly formed adjacent relationship and deleting the old adjacent relationship corresponding to the merged two segments. The merging cost of newly formed adjacent segments is also calculated. Finally, the new list is updated. The detailed information can be found in Table I.

Note that with all threshold-based region merging methods, the user must vary the merging parameter for producing segmentation with high quality. Hence, the merging parameter β is varied from 0.1 to 1 in steps of 0.1, and the appropriate β is recognized using the approach of Johnson et al. (2015).

C. OVER- AND UNDER- SEGMENTATION RECOGNITION AND RE-PROCESS

Over- and under-segmentation recognition and re-process is also a critical issue to improve the segmentation quality. After implementing the local scale-guided region merging, the coefficient of variation (CV) is used to estimate the degree of over- and under-segmented regions from the perspective of the object and recognize the over- and under-segmented regions. Then, these regions are further processed. Finally, the segments at separating layers are combined into one ultimate segmentation layer.

1) OVER- AND UNDER- SEGMENTATION RECOGNITION

Different segments have different sizes and ranges of spectral reflectance. To achieve the objective comparison among segments, the CV is applied to calculate the degree of over- and under-segmentation, which can be defined as:

$$CV = \frac{\sqrt{\frac{\sum (X_i - \overline{X})^2}{n}}}{\overline{X}} \quad (6)$$

where $\overline{X}$ is the mean of spectral values within a segment, n is the number of pixels within the segment and X_i is the spectral value of one pixel within the segment. The CV reflects the dispersion degree of spectral values within a segment, i.e., the segment with higher CV tends toward under-segmentation and the segment with lower CV tends toward over-segmentation.

The process of over- and under-segmentation recognition is follows: First, the CV corresponding to each segment is calculated. Second, two CV thresholds are set to recognize the over-segmented regions (lowest 30% of segments) and under-segmented regions (highest 30% of segments) in each segmentation layer (vegetative segments/non-vegetative segments). Third, the separating layer is divided into three layers of over-segmentation, good-segmentation, and under-segmentation.

2) OVER- AND UNDER- SEGMENTATION RE-PROCESS

The recognized over- and under-segmented regions need to be further processed for obtaining more accurate segments. For the over-segmentation layer, the adjacent segments are recognized using the RAG model and merged using the FLSA with local scale; For the under-segmentation layer, the under-segmented regions will be initialized to the original segments obtained by the watershed transformation. Then the adjacent relationship of these original segments is searched by building RAG, and the FLSA with local scale is again implemented to merge the adjacent original segments. Finally, all the separating layers are fused into one ultimate segmentation layer.

D. METHOD VALIDATION

To demonstrate the effectiveness of the proposed method, the visual and quantitative evaluation approaches are adopted in this paper. The total error (TE) and the quality rate (QR) [59] are chosen for quantitative evaluation. The TE is a Euclidean Distance combination of over-segmentation (OS) and under-segmentation (US). The OS and US are both varied from 0 to 1, where OS is equal to zero indicates no over-segmented regions and US is equal to zero indicates no under-segmented regions. However, OS and US are often antagonistic for evaluating one remote sensing image segmentation and cannot indicate the overall segmentation quality. Hence, the TE is used to solve this problem and it varies from 0 to $\sqrt{2}$. The lower TE indicates better overall segmentation quality. The QR is a discrepancy measure in area between a reference polygon and its corresponding segment, and varies from 0 to 1. When segments are over- or under-segmented, the QR value will be large. When QR is equal to 0, indicating the highest segmentation goodness (i.e., no over- and under-segmentation). The more details of TE and QR can be found in Clinton et al. (2010).

A reference polygon may be overlapped with many segments in a segmentation; however, it is required that one recognized geo-object corresponds to a single segment. To determine the corresponding relationship from the reference polygon and its overlapped segments, we develop an overlay measure (OM) to calculate the goodness- of –fit between one reference polygon and its overlapped segments, as follows:

$$OM = \frac{2 \cdot \frac{area(r_i \cap s_j)}{area(r_i)} \cdot \frac{area(r_i \cap s_j)}{area(s_j)}}{\frac{area(r_i \cap s_j)}{area(r_i)} + \frac{area(r_i \cap s_j)}{area(s_j)}} \quad (7)$$

where r_i is one of the reference polygons and s_j is one of the corresponding overlapped segments. Note that OM varies from 0 to 1, and the segment with the highest OM value is chosen to correspond with the reference polygon.

III. RESULTS

A. DATASET

To test the effectiveness of the proposed method, three subsets of multi-spectral high spatial resolution images are acquired from a Chinese remote sensing satellite, gaofen-1, with four bands of 8m resolution (blue, green, red, and near-infrared) and a panchromatic band of 2m resolution. These images are collected on May 8th, 2016, and then the NNDiffuse pan-sharpening method embedded in software ENVI 5.2 is used to generate multi-spectral images of 2m resolution with four bands. Figure 5(a), (b) & (c) show the dataset. The three test images are recognized as T1, T2 and T3, respectively. The sizes of all the test images are 1.1×1.1 km. The testing areas are not very large, but various land cover types in the dataset are capable of demonstrating the proposed method. T1 is a typical urban area containing complex buildings, green belts, and road networks. T2 is a suburban area filled with industrial buildings, water, tree clusters, and large open fields. And T3 is a rural area including a village and differently sized agricultural land. Figure 6(a)-(c) present the corresponding expert-manually digitized reference geo-objects, whose numbers are 106, 124 and 77, respectively. We can observe that differently sized geo-objects have been digitized for all the test images.

B. THE INTERMEDIATE RESULTS OF THE PROPOSED METHOD

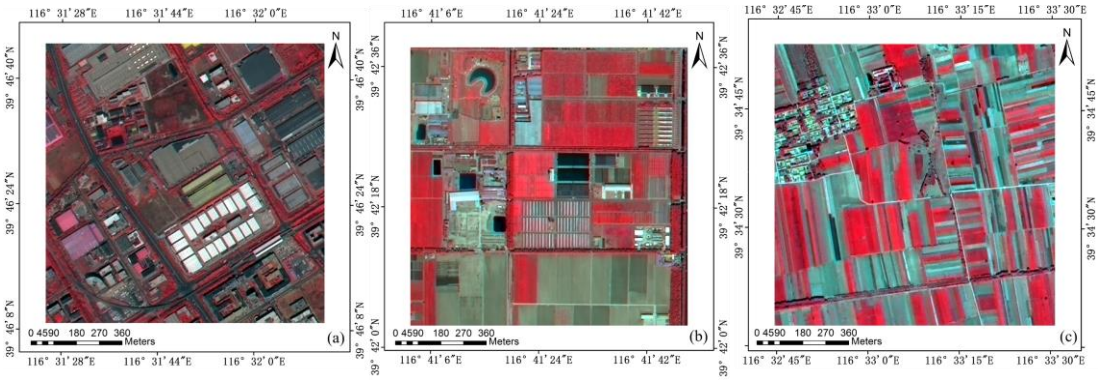


FIGURE 5. **GF-1 images selected for segmentation experiment: (a) T1, urban area, (b) T2, suburban area, and (c) T3, rural area. These images are shown with false color combination by near-infrared, red and green bands.**

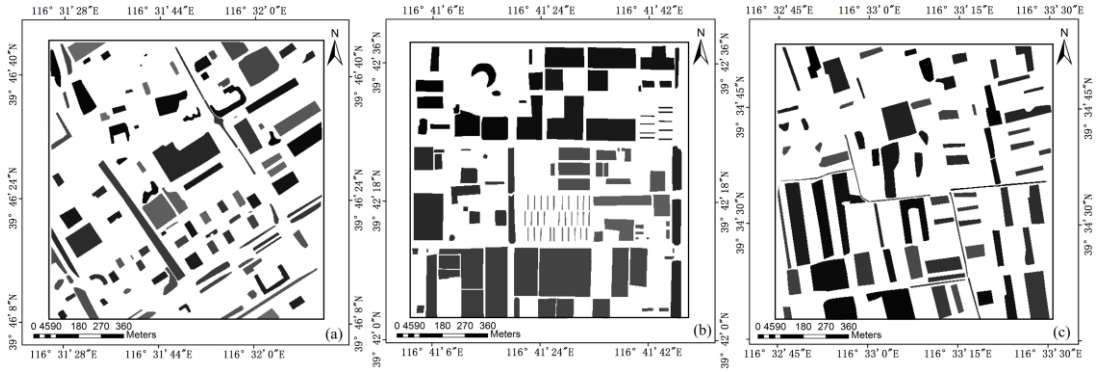


FIGURE 6. **The manually digitized reference geo-objects selected for accuracy evaluation: (a) T1, (b) T2, and (c) T3.**

As described in Section II, the proposed method mainly includes three processes: (1) the primitive segmentation production and stratification, (2) the local scale-guided region merging at separating layers, and (3) the over- and under-segmented regions recognition and re-process at separating layers. To make readers fully understand the whole procedure, some intermediate results are displayed.

For the first process, Figure 7 shows the primitive segmentation results and vegetation maps. We can observe in Figure 7(a) that the initial segmentations of the three test images are seriously over-segmented. Combining the primitive segments and vegetation maps, two layers with vegetative segments and non-vegetative segments are generated, as illustrated in Figure 8. It is observed that the segments of the vegetative part and non-vegetative part have been effectively separated.

For the second process, the local scale-guided region merging is the focus. The local scale is calculated based on the optimal global scale. Figure 9 shows the change of goodness scores of the test images with merging parameter β increasing to determine the optimal global scales. As indicated by the highest goodness scores, 0.7, 0.8 and 0.7 are selected for merging vegetative segments of T1, T2 and T3, and 0.5, 0.7 and 0.6 are selected for merging non-vegetative segments of T1, T2 and T3, respectively. Figure 10 displays the local scale-guided region merging results of the test images at the optimal β. Compared with Figure 8,

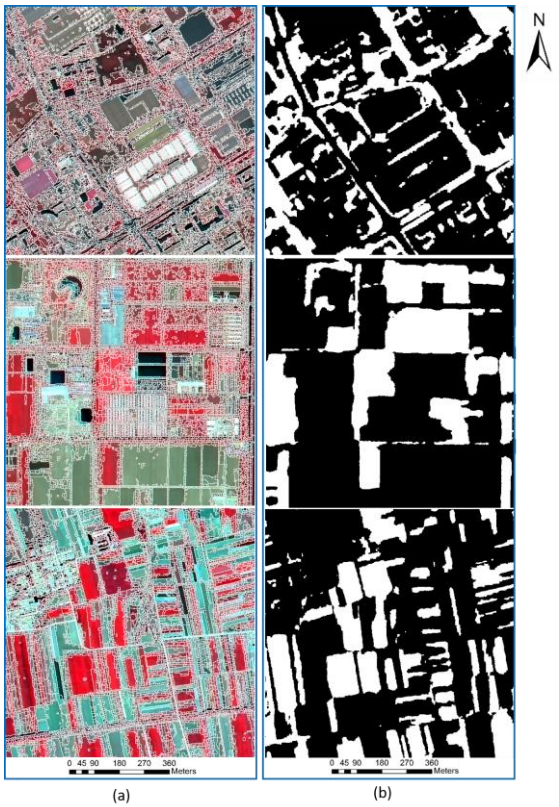


FIGURE 7. **The primitive segmentation results and vegetation maps of three test images: (a) primitive segmentation produced by watershed transformation, and (b) vegetation maps produced by NDVI.**

FIGURE 8. The primitive segmentation stratification results: (a) vegetative segments, and (b) non-vegetative segments.

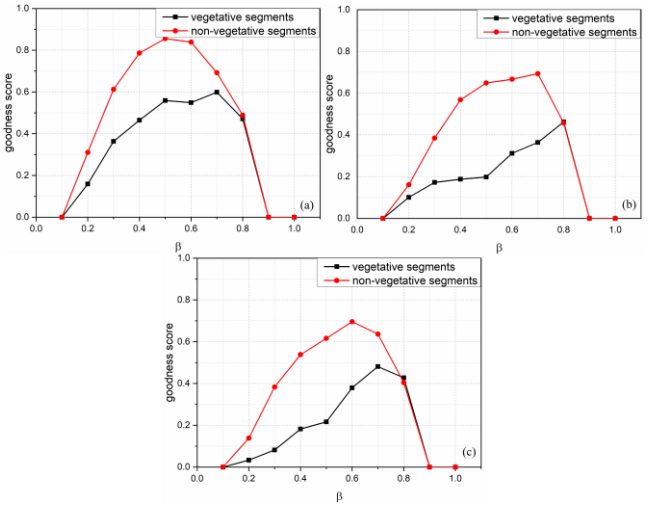


FIGURE 9. The change of goodness scores of the test images with merging parameter β increasing: (a) T1, (b) T2, and (c) T3.

the level of over-segmentation has been greatly reduced.

For the third process, the over- and under-segmented regions are identified and further processed for obtaining more accurate segmentation. After applying CV to recognize the over- and under-segmented regions, the over-segmented regions are further merged using the local scale-guided region merging method, and the under-segmented regions are re-merged after the under-segmented regions are initialized to the original segments obtained by the

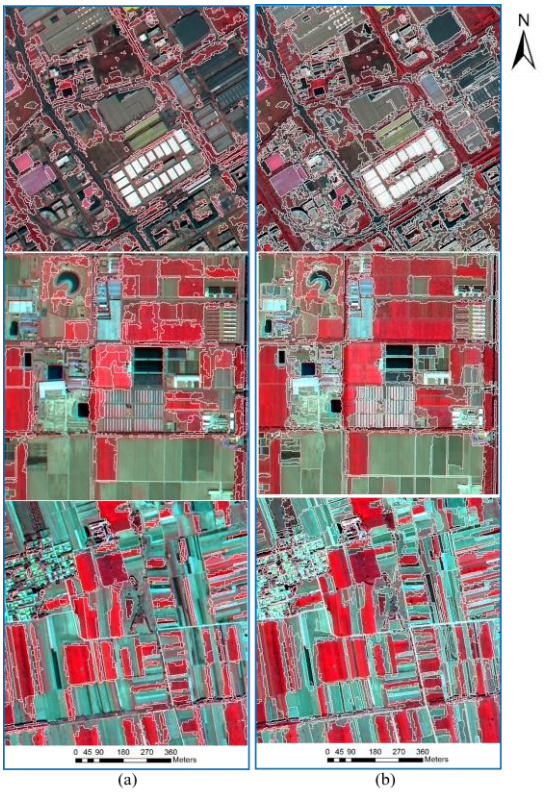


FIGURE 10. The local scale-guided region merging results of the test images at the optimal β: (a) vegetative merging results, and (b) non-vegetative merging results.

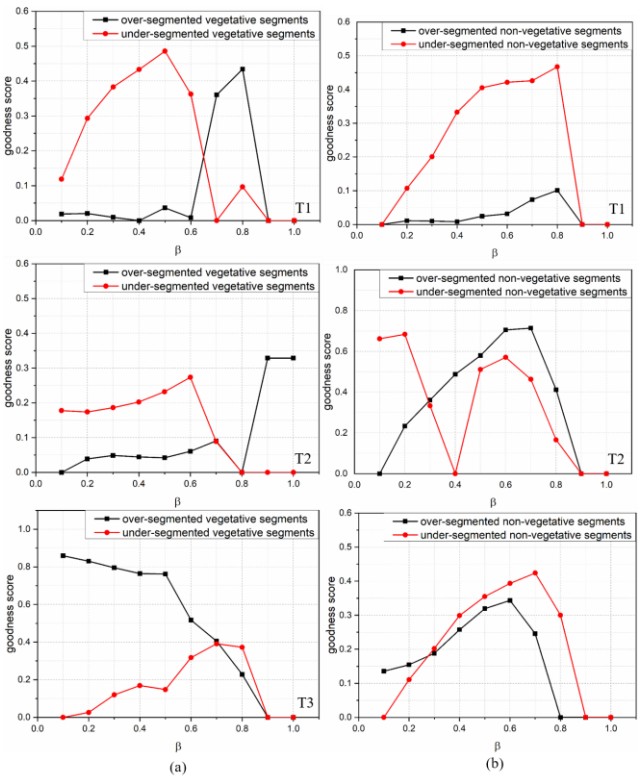


FIGURE 11. The goodness score variations of over- and under-segmented regions with increased β: (a) vegetative segments, and (b) non-vegetative segments.

This article has been accepted for publication in IEEE Access. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2022.3194047

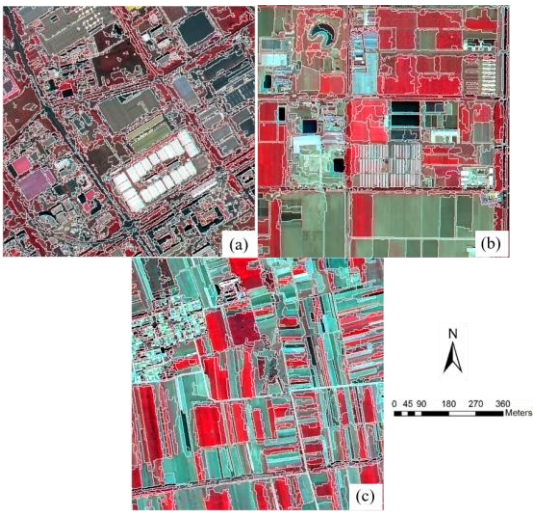


FIGURE 12. The ultimate segmentation results produced by the proposed method.

watershed transformation. To obtain the optimal global scale, the result of goodness score variations with increased β is shown in Figure 11. For T1, 0.8 is selected as the further merged threshold of vegetative segments and non-vegetative segments (over-segmented regions), and 0.5 and 0.8 are selected as the re-merged threshold of vegetative segments and non-vegetative segments (under-segmented regions). For T2, the further merged threshold of that is set at 0.9 and 0.7, respectively, and the re-merged threshold of that is set at 0.6 and 0.2, respectively. For T3, we determine 0.1 and 0.6 as the further merged threshold of that, and 0.7 as the re-merged threshold of that. Figure 12 shows the ultimate segmentation results produced by the proposed method.

C. COMPARISON EXPERIMENT

10 methods were compared and analyzed for fully validating the proposed method. The detailed information on the 10 methods can be found in Table II. Obviously, M1 and M2 are compared for analyzing the superiority of the proposed local scale approach; M3 is used to reveal the effect of over- and under-segmentation recognition and re-process; M4 aims at revealing the effect of primitive segmentation stratification; M5, M6, M7, M8, M9 and M10 are the competitive methods for further comparison. Note that all the methods are based on the same initial segments produced by the watershed transform, as shown in Figure 7(a). The segmentation parameters of these methods are determined by the parameter optimization method of Johnson et al. (2015), and the detailed parameter setting can be found in Table III.

Then, some subsets of the segmentation results of the three test images are shown in Figure 13 for visual analysis. For the first subset of T1, although all the methods wrongly merged grassland and bare soil land into one geo-objects, M1 and M3 performed better in segmenting tree belts than the other methods. M4 and M5 could not segment some tree

TABLE II
THE DETAILED INFORMATION OF COMPARED METHODS

Code	Description
M1	The proposed method.
M2	The proposed method with global approach.
M3	The proposed method with no over- and under-segmentation recognition and re-process.
M4	The proposed method with no primitive segmentation stratification.
M5	The FLSA-based region merging method. It is proposed by Robinson et al. (2002).
M6	The local spectral angle (SA)-based region merging method. It is proposed by Yang et al. (2017).
M7	The scale-variable region-merging method. It is proposed by Su (2019).
M8	The refined FLSA region merging method. It is proposed by Wang et al. (2020).
M9	The region merging method aided by inter-segment and boundary homogeneities. It is proposed by Zhang et al. (2019).
M10	The multiple merging method. It is proposed by Wang et al. (2021)

TABLE III
DETAILED PARAMETER SETTING OF THE 10 METHODS

Code	Image	parameter setting(s)						
		β						ρ
		Vegetative segments			Non-vegetative segments			
		Initial layer	OS layer	US layer	Initial layer	OS layer	US layer	
M1	T1	0.7	0.8	0.5	0.5	0.8	0.8	-
	T2	0.8	0.9	0.6	0.7	0.7	0.2	-
	T3	0.7	0.1	0.7	0.6	0.6	0.7	-
M2	T1	0.7	0.1	0.5	0.5	0.8	0.6	-
	T2	0.8	0.9	0.6	0.7	0.7	0.5	-
	T3	0.7	0.2	0.7	0.6	0.1	0.6	-
M3	T1	0.7	-	-	0.5	-	-	-
	T2	0.8	-	-	0.7	-	-	-
	T3	0.7	-	-	0.6	-	-	-
		β						ρ
		Initial layer		OS layer		US layer		
M4	T1	0.5		0.8		0.6		-
	T2	0.6		0.1		0.6		-
	T3	0.6		0.7		0.6		-
M5	T1	0.5		-		-		-
	T2	0.6		-		-		-
	T3	0.6		-		-		-
M6	T1	0.7		-		-		-
	T2	0.6		-		-		-
	T3	0.7		-		-		-
M7	T1	0.5		-		-		1
	T2	0.8		-		-		1
	T3	0.7		-		-		1
M8	T1	0.6		-		-		-
	T2	0.6		-		-		-
	T3	0.6		-		-		-
M9	T1	0.6		-		-		-
	T2	0.7		-		-		-
	T3	0.7		-		-		-
		Global scale		Merging criteria				
		Q		w_{color}		w_{shape}		
M10	T1	8		0.2		0.2		
	T2	8		0.1		0.5		
	T3	9		0.1		0.2		

[a] OS represents over-segmentation, and US represents under-segmentation.

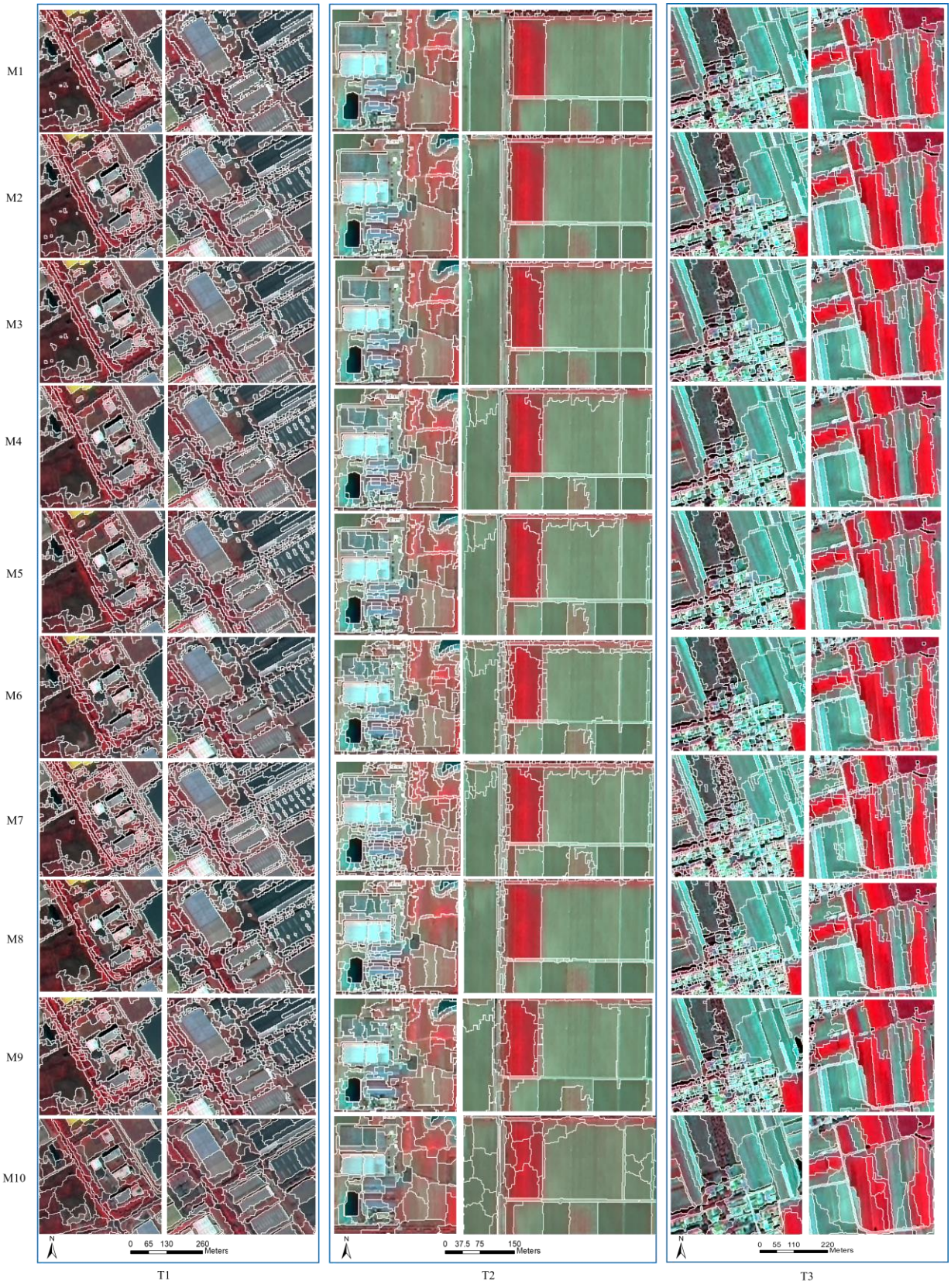


FIGURE 13. **The subsets of the segmentation results of the three test images with 10 segmentation methods.**

belts out with the bare soil land. M7 seriously over-segmented the tree belts. For the second subset of T1, M1

and M3 delineated differently sized geo-objects better than the other methods. Some buildings could not be segment out in the results of M2, M4, M5, M8 and M9. M6 and M7 over-segmented various geo-objects with different degrees. For the first subset of T2, T1 and T2 produced good segments for the bare soil land and buildings, whereas the other eight methods over-segmented these geo-objects with varying degrees. For the second subset of T2, T1 and T2 had a better overall performance in segmenting vegetative land and bare soil land, but they wrongly merged one small vegetative area and a path into the bare soil land. M3 had slight over-segmentation for the vegetative land. M4, M5, M6, M7, M9 and M10 almost had different levels of over-segmentation for the vegetative land and bare soil land. M8 under-segmented some bare soil land and a path. For the first subset of T3, M1 and M10 extracted the tree belt better than the other methods. M7 performed the worst in segmenting out the tree belt. Moreover, M4 did not distinguish some vegetative areas and bare soil areas well. For the second subset of T3, M1, M2, M4, M8, M9 and M10 had relatively satisfying segmentation for the bare soil land, whereas M3, M6, and M7 produced evident over-segmentation for this tpye of geo-object.

Finally, to further analyze the effectiveness of the proposed method, the quantitative evaluation results of three test images are provided in Table IV. For T1, The M1 can produce the segmentation with the best overall quality, as indicated by the lowest QR and TE values. The method with the second-best score is the M3. Although the indicators of OS and US cannot directly reflect the overall segmentation quality, they are competent to indicate the level of over- and under-segmentation. The lowest OS value and relatively lower US value reveal that there are less over- and under-segmentation in the segmentation produced by M1. For T2, the M1 had the lowest QR value and the second-lowest TE value, indicating good overall performance in generating satisfying segmentation. The M2 is followed, as indicated by the lowest TE value and the third-lowest QR value. The lowest OS and US values revealed less over-segmentation and under-segmentation in the M1 and M7 results, respectively. For T3, the M1 apparently outperformed the other nine methods, with the lowest QR and TE values. M4 produced less over-segmentation and M7 produced less under-segmentation, with the lowest OS and US values, respectively.

In addition, the running efficiency of the competing methods is shown in Table V. According to Table V, M1-M3 had less running time than the other methods, indicating the stratification strategy can effectively improving the running efficiency. More running time of M1 than M2 indicated the proposed local scale approach added the time complexity. Overall, the proposed method is a relatively timely and accurately efficient remote-sensing image segmentation method compared with the competing methods.

TABLE IV
QUANTITATIVE EVALUATION FOR THE SEGMENTATION RESULTS OF THE THREE TEST IMAGES WITH 10 SEGMENTATION METHODS.

Image	Method	QR	TE	OS	US
T1	M1	**0.4465**	**0.28**	**0.3487**	0.1877
	M2	0.4533	0.2834	0.3524	0.191
	M3	0.447	0.2808	0.3507	0.1862
	M4	0.4941	0.3188	0.3767	0.2477
	M5	0.4935	0.3171	0.3779	0.2415
	M6	0.617	0.3847	0.4073	0.3607
	M7	0.5447	0.3571	0.4808	**0.1545**
	M8	0.4955	0.3174	0.3791	0.2403
	M9	0.5143	0.3587	0.4287	0.2711
	M10	0.4793	0.298	0.2936	0.3023
T2	M1	**0.413**	0.2538	**0.2587**	0.2489
	M2	0.435	**0.2514**	0.3008	0.1897
	M3	0.4322	0.2571	0.2918	0.217
	M4	0.4405	0.2776	0.3599	0.1569
	M5	0.4401	0.2738	0.3515	0.1623
	M6	0.5987	0.403	0.4701	0.3221
	M7	0.5501	0.379	0.5277	**0.094**
	M8	0.5354	0.3428	0.4469	0.1879
	M9	0.4997	0.3218	0.4123	0.1925
	M10	0.5413	0.3494	0.4593	0.1824
T3	M1	**0.5019**	**0.3271**	0.4274	0.1771
	M2	0.5288	0.3502	0.4658	0.1684
	M3	0.5388	0.3577	0.4763	0.1702
	M4	0.5388	0.3393	**0.3499**	0.3284
	M5	0.5431	0.3527	0.4295	0.2537
	M6	0.588	0.3998	0.5043	0.2557
	M7	0.6512	0.453	0.6286	**0.1237**
	M8	0.5509	0.3585	0.4524	0.2289
	M9	0.5437	0.3519	0.4677	0.1701
	M10	0.5499	0.3533	0.4701	0.1692

TABLE V
THE RUNNING TIME OF COMPARED METHODS

Method	Running time (unit: second)		
	Test image		
	T1	T2	T3
M1	98.2	85.8	76.9
M2	93.1	83.1	72.4
M3	95.7	83.3	74.6
M4	107.1	91.9	81.3
M5	102.4	89.6	78.6
M6	602.6	489.6	425.7
M7	315.3	280.1	266.3
M8	464.9	419.9	389.8
M9	613.9	492.3	427.5
M10	293.4	270.7	258.6

IV. DISCUSSION

A good segmentation is very critical for subsequent remote-sensing image interpretation, since it is regarded as the core of GEOBIA. The hybrid image segmentation method is a good alternative to produce good segmentation that best matched the different sizes of geo-objects. However, the existing methods almost use a certain scale or other

segmentation parameters (SPs) to control the sizes and shapes of segments. This will lead to two issues: (1) one single scale is impossible to segment every geo-object well due to the land cover complexity within remote-sensing imageries; (2) over- and under-segmented regions still occur in the segmentation results, whatever using any advanced segmentation methods.

Multiple scale optimization and local scale approach are two main solutions to solve the first issue. But it is difficult to fuse the segmentation results obtained by the multiple optimal scales into one layer for the multiple scale optimization approaches. Considering the unique characteristics in geo-object size, shape, and spectral variability with different types of land covers, and using these characteristics to separate primitive segments into different layers may help achieve the multiple optimal scale determination and forming subsequent ultimate layer. For the local scale approaches, it adjusts the global scale based on local spectral structure for a more objective merging process. However, some methods neglect that the region merging process just happened in two adjacent segments. The spectral variance change after one pair of adjacent segments are merged is also a non-negligible factor to determine the reasonable local scale. Based on the above considerations, this paper developed the local scale-guided hierarchical region merging method. The multiple optimal scale determination and the local scale approach are skillfully combined in this method.

Over- and under-segmentation recognition and further process can solve the second issue. The existing methods can indicate segmentation quality but may not be helpful for further improving the segmentation results since the evaluation metrics is in a global manner. To further improve the segmentation quality, this paper developed one over- and under-segmentation recognition method based on the object-based paradigm, in which the CV is employed. Different from that the existing method uses different indicators to calculate the level of over- and under-segmentation, the CV is a comprehensive indicator that reveals the level of both over- and under-segmentation.

The proposed technology, which is more like the framework of one merge process, is independent of the related split- and -merge based segmentation methods. The readers can employ any split- and -merge method into our proposed technology, not just the watershed transformation and FLSA used in this paper. This undoubtedly increases the robustness of the proposed technology, extending its application scope. Moreover, the proposed framework will be applied to Multiresolution Segmentation (MRS) algorithm in the future, since the MRS is currently a popular segmentation method for GEOBIA.

To demonstrate the proposed technology, visual and quantitative analyses were given, as described in Section III. Figure 13 showed that the proposed method (M1) had a better overall performance in segmenting differently sized geo-objects compared with the competing methods. Table IV further demonstrated the conclusion since M1 had lower QR and TE values than M5-M10. The lower QR and TE values of M1 than M2 in most cases indicated the superiority of the proposed local scale approach. The same phenomena also occurred in the comparative results of T1 and T3, T1 and T4. The comparative results of T1 and T3 indicated that the primitive segmentation can effectively avoid wrong merging with different types of land covers, thus improving the segmentation quality. The comparative results of T1 and T4 revealed that the over- and under-segmentation recognition and re-process employing the CV is capable of reducing the over- under-segmentation. Overall, the proposed method has the potential to produce better segmentation results.

However, some problems were found in the experimental results. The detailed analyses are as follows:

(1) We observe that all the methods were incapable of segmenting out the vegetative area and bare soil area in the first subset of T1 of Figure 13, while the NDVI map of T1 in Figure 7(b) obviously separated them. The inappropriate initial segments may be the main cause of the above situation. Thus, using a more efficient segmentation method to obtain good initial segments is also non-negligible in the framework of split- and -merge strategy.

(2) The fact is found that the NDVI map of T2 (Figure 7(b)) did not recognize some vegetative parts very well, leading to that M1 and M2 could not distinguish the vegetative area and bare soil area in the second subset of T2 of Fig. 13. This paper just used the empirical threshold to calculate the NDVI map, which may result in segmentation error.

(3) It is observed that the QR and TE values of M3 are a little higher than those of M1 for T1 in Table IV, indicating a limited improvement of segmentation quality in T1 after the over- and under-segmentation recognition and re-process is implemented. The reason causing the phenomenon is that single over-segmented regions were recognized, and it was not helpful to reduce segmentation error since no neighboring segments can be further merged with the single over-segmented region. Existing recognition methods of over- and under-segmentation distinguish these regions based on certain mathematical criteria, such as the local variance and Moran's I in Johnson and Xie (2011), and CV adopted in this paper. They both ignore the context information between one over-segmented region and its neighbors. Thus, it is very critical for developing a more objective recognition method by considering the semantic information in the future.

Moreover, only NDVI was used in this paper to achieve the primitive segmentation stratification. Using NDVI for segmentation may be a little outdated and simple, but it is limited by the condition that the test images only contain four bands of blue, green, red, and near-infrared. With the spectral resolution improvement of the high spatial

resolution imagery in the future, more auxiliary parameters, such as the normalized difference impervious surface index (NDISI), Normalized Burn Ratio (NBR), and normalized difference building index (NDBI), are involved in primitive segmentation stratification for obtaining more accurate segmentation.

Although the multiple scale optimization and local scale approach is helpful to improve segmentation quality, the global scale needs to be artificially set and the optimal scale is searched within some ranges, which adds certain subjectivity and much work. In the future, developing one automatic segmentation stopping rule is also very critical to guarantee objective and efficient segmentation.

V. CONCLUSION

In this paper, a hybrid remote sensing segmentation method with local scale-guided hierarchical region merging and further over- and under-segmentation processing is proposed to segment differently sized geo-objects. The proposed technology skillfully integrates multiple optimal scale determination and the local scale approach, and then recognizes and processes the over- and under-segmented regions for generating more satisfying segmentation. To validate the proposed method, three test images of gaofen-1 were used and ten competing methods were compared. The visual and quantitative results demonstrated that the proposed method was more potential to delineate various sizes of geo-objects compared with the competing methods.

In the future, more auxiliary parameters used for primitive segmentation stratification, the automatic segmentation stopping rule, and the semantic information-based over- and under-segmentation recognition method will be our main focus.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions for further improving the work.

REFERENCES

[1] T. Blaschke *et al.*, "Geographic Object-Based Image Analysis - Towards a new paradigm," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 87, pp. 180-191, Jan 2014.

[2] M. Lei, G. Yin, Z. Zhou, H. Lu, and M. Li, *Uncertainty of Object-Based Image Analysis for Drone Survey Images*. in Drones - Applications, Dekoulis, G., Ed. Intech Open, UK: London, 2017.

[3] D. M. Johnson and R. Mueller, "Pre- and within-season crop type classification trained with archival land cover information," *Remote Sensing of Environment*, vol. 264, pp. 112576, Oct 2021.

[4] F. Low, P. Knofel, and C. Conrad, "Analysis of uncertainty in multi-temporal object-based classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 105, pp. 91-106, Jul 2015.

[5] T. F. Su and S. W. Zhang, "Object-based crop classification in Hetao plain using random forest," *Earth Science Informatics*, vol. 14, no. 1, pp. 119-131, Mar 2021.

[6] M. C. Li, L. Ma, T. Blaschke, L. Cheng, and D. Tiede, "A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments," *International Journal of Applied Earth Observation and Geoinformation*, vol. 49, pp. 87-98, Jul 2016.

[7] Q. L. Guo, J. P. Zhang, and Y. Zhang, "Multitemporal Images Change Detection Based on AMMF and Spectral Constraint Strategy," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 4, pp. 3444-3457, Apr 2021.

[8] P. F. Xiao, M. Yuan, X. L. Zhang, X. Z. Feng, and Y. W. Guo, "Cosegmentation for Object-Based Building Change Detection From High-Resolution Remotely Sensed Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 3, pp. 1587-1603, Mar 2017.

[9] R. Mv, J. Jones, H. Eicken, and C. Kambhamettu, "Extracting Quantitative Information on Coastal Ice Dynamics and Ice Hazard Events From Marine Radar Digital Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 5, pp. 2556-2570, May 2013.

[10] X. Long, X. D. Chen, A. L. Wang, Y. Liu, and S. Y. Ji, "Observation of Sea Ice Drift Characteristics with Marine Radar Images on an Offshore Platform in the Bohai Sea," *Journal of Waterway Port Coastal and Ocean Engineering*, vol. 147, no. 4, pp. 04021013, Jul 2021.

[11] L. Shujun and H. Lianzhi, *Remote Sensing Image Change Detection Based on Fully Convolutional Network With Pyramid Attention*. Presented at 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS.

[12] X. H. Li, Z. S. Du, Y. Y. Huang, and Z. Y. Tan, "A deep translation (GAN) based change detection network for optical and SAR remote sensing images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 179, pp. 14-34, Sep 2021.

[13] K. Johansen, L. A. Arroyo, S. Phinn, and C. Witte, "Comparison of Geo-Object Based and Pixel-Based Change Detection of Riparian Environments using High Spatial Resolution Multi-Spectral Imagery," *Photogrammetric Engineering and Remote Sensing*, vol. 76, no. 2, pp. 123-136, Feb 2010.

[14] C. Cleve, M. Kelly, F. R. Kearns, and M. Morltz, "Classification of the wildland-urban interface: A comparison of pixel- and object-based classifications using high-resolution aerial photography," *Computers Environment and Urban Systems*, vol. 32, no. 4, pp. 317-326, Jul 2008.

[15] C. Burnett and T. Blaschke, "A multi-scale segmentation/object relationship modelling methodology for landscape analysis," *Ecological Modelling*, vol. 168, no. 3, pp. 233-249, Oct 15 2003.

[16] T. F. Su and S. W. Zhang, "Local and global evaluation for remote sensing image segmentation," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 130, pp. 256-276, Aug 2017.

[17] J. Yang, Y. H. He, J. Caspersen, and T. Jones, "A discrepancy measure for segmentation evaluation from the perspective of object recognition," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 101, pp. 186-192, Mar 2015.

[18] L. Ma, L. Cheng, M. C. Li, Y. X. Liu, and X. X. Ma, "Training set size, scale, and features in Geographic Object-Based Image Analysis of very high resolution unmanned aerial vehicle imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 102, pp. 14-27, Apr 2015.

[19] L. Ma, M. C. Li, X. X. Ma, L. Cheng, P. J. Du, and Y. X. Liu, "A review of supervised object-based land-cover image classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 130, pp. 277-293, Aug 2017.

[20] T. Blaschke, "Object based image analysis for remote sensing," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 1, pp. 2-16, Jan 2010.

[21] B. A. Johnson and L. Ma, "Image Segmentation and Object-Based Image Analysis for Environmental Monitoring: Recent Areas of Interest, Researchers' Views on the Future Priorities," *Remote Sensing*, vol. 12, no. 11, pp. 1772, Jun 2020.

[22] J. Canny, "A Computational Approach to Edge-Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679-698, Nov 1986.

[23] R. Gaetano, G. Masi, G. Poggi, L. Verdoliva, and G. Scarpa, "Marker-Controlled Watershed-Based Segmentation of Multiresolution Remote Sensing Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 6, pp. 2987-3004, Jun 2015.

[24] D. F. He and Q. D. Wang, "Edge Detecting Method for Microscopic Image of Cotton Fiber Cross-Section Using RCF Deep Neural Network," *Information*, vol. 12, no. 5, pp. 196, May 2021.

[25] H. Zhang, Z. H. Tang, Y. F. Xie, X. L. Gao, and Q. Chen, "A watershed segmentation algorithm based on an optimal marker for bubble size measurement," *Measurement*, vol. 138, pp. 182-193, May 2019.

[26] J. C. Tilton, Y. Tarabalka, P. M. Montesano, and E. Gofman, "Best Merge Region-Growing Segmentation With Integrated Nonadjacent Region Object Aggregation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 11, pp. 4454-4467, Nov 2012.

[27] Implementation of a Fast Algorithm for Segmenting SAR Imagery, by D. J. Robinson, N. J. Redding, and D. J. Crisp. (2002). Patent DSTO-TR-1242.

[28] C. B. Zhou, D. M. Wu, W. H. Qin, and C. C. Liu, "An efficient two-stage region merging method for interactive image segmentation," *Computers & Electrical Engineering*, vol. 54, pp. 220-229, Aug 2016.

[29] T. F. Su, T. X. Liu, S. W. Zhang, Z. Y. Qu, and R. P. Li, "Machine learning-assisted region merging for remote sensing image segmentation," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 168, pp. 89-123, Oct 2020.

[30] T. F. Su, "Unsupervised evaluation-based region merging for high resolution remote sensing image segmentation," *Giscience & Remote Sensing*, vol. 56, no. 6, pp. 811-842, Aug 2019.

[31] Z. W. Hu, Q. Zhang, Q. Zou, Q. Q. Li, and G. F. Wu, "Stepwise Evolution Analysis of the Region-Merging Segmentation for Scale Parameterization," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 7, pp. 2461-2472, Jul 2018.

[32] T. F. Su, "A novel region-merging approach guided by priority for high resolution image segmentation," *Remote Sensing Letters*, vol. 8, no. 8, pp. 771-780, 2017.

[33] H. Y. Wang *et al.*, "Improvement of Region-Merging Image Segmentation Accuracy Using Multiple Merging Criteria," *Remote Sensing*, vol. 13, no. 14, pp. 2782, Jul 2021.

[34] Y. H. Zhang, X. Wang, H. S. Tan, C. Xu, X. Ma, and T. F. Xu, "Region Merging Method for Remote Sensing Spectral Image Aided by Inter-Segment and Boundary Homogeneities," *Remote Sensing*, vol. 11, no. 12, pp. 1414, Jun 2019.

[35] Y. J. Wang, Q. W. Qi, L. L. Jiang, and L. Ying, "Hybrid Remote Sensing Image Segmentation Considering Intrasegment Homogeneity and Intersegment Heterogeneity," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 1, pp. 22-26, Jan 2020.

[36] J. Yang, Y. H. He, and J. Caspersen, "Region merging using local spectral angle thresholds: A more accurate method for hybrid segmentation of remote sensing images," *Remote Sensing of Environment*, vol. 190, pp. 137-148, Mar 2017.

[37] X. L. Zhang, P. F. Xiao, X. Z. Feng, J. G. Wang, and Z. Wang, "Hybrid region merging method for segmentation of high-resolution remote sensing images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 98, pp. 19-28, Dec 2014.

[38] M. Wang and R. X. Li, "Segmentation of High Spatial Resolution Remote Sensing Imagery Based on Hard-Boundary Constraint and Two-Stage Merging," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5712-5725, Sep 2014.

[39] J. Wang, L. L. Jiang, Y. J. Wang, and Q. W. Qi, "An Improved Hybrid Segmentation Method for Remote Sensing Images," *ISPRS International Journal of Geo-Information*, vol. 8, no. 12, pp. 543, Dec 2019.

[40] Y. J. Wang, Q. Y. Meng, Q. W. Qi, J. Yang, and Y. Liu, "Region Merging Considering Within- and Between-Segment Heterogeneity: An Improved Hybrid Remote-Sensing Image Segmentation Method," *Remote Sensing*, vol. 10, no. 5, pp. 781, May 2018.

[41] B. Chen, F. Qiu, B. F. Wu, and H. Y. Du, "Image Segmentation Based on Constrained Spectral Variance Difference and Edge Penalty," *Remote Sensing*, vol. 7, no. 5, pp. 5980-6004, May 2015.

[42] J. Chen, M. Deng, X. M. Mei, T. Q. Chen, Q. B. Shao, and L. Hong, "Optimal segmentation of a high-resolution remote-sensing image guided by area and boundary," *International Journal of Remote Sensing*, vol. 35, no. 19, pp. 6914-6939, Oct 2014.

[43] X. L. Li, J. S. Chen, L. L. Zhao, S. X. Guo, L. Y. Sun, and X. M. Zhao, "Adaptive Distance-Weighted Voronoi Tessellation for Remote Sensing Image Segmentation," *Remote Sensing*, vol. 12, no. 24, pp. 4115, Dec 2020.

[44] B. A. Johnson, M. Bragais, I. Endo, D. B. Magcale-Macandog, and P. B. M. Macandog, "Image Segmentation Parameter Optimization Considering Within- and Between-Segment Heterogeneity at Multiple Scale Levels: Test Case for Mapping Residential Areas Using Landsat Imagery," *ISPRS International Journal of Geo-Information*, vol. 4, no. 4, pp. 2292-2305, Dec 2015.

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

[46] A. M. Vamsee, P. Kamala, T. R. Martha, K. V. Kumar, G. J. Sankar, and E. Amminedu, "A Tool Assessing Optimal Multi-Scale Image Segmentation," *Journal of the Indian Society of Remote Sensing*, vol. 46, no. 1, pp. 31-41, Jan 2018.

[47] T. R. Martha, N. Kerle, C. J. van Westen, V. Jetten, and K. V. Kumar, "Segment Optimization and Data-Driven Thresholding for Knowledge-Based Landslide Detection by Object-Based Image Analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 12, pp. 4928-4943, Dec 2011.

[48] T. F. Su, "Scale-variable region-merging for high resolution remote sensing image segmentation," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 147, pp. 319-334, Jan 2019.

[49] F. Canovas-Garcia and F. Alonso-Sarria, "A local approach to optimize the scale parameter in multiresolution segmentation for multispectral imagery," *Geocarto International*, vol. 30, no. 8, pp. 937-961, Sep 2015.

[50] Y. J. Wang, Z. H. Tian, Q. W. Qi, and J. Wang, "Double-Variance Measures: A Potential Approach to Parameter Optimization of Remote Sensing Image Segmentation," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 2314-2326, Mar 2021.

[51] Y. J. Wang, Q. W. Qi, Y. Liu, L. L. Jiang, and J. Wang, "Unsupervised segmentation parameter selection using the local spatial statistics for remote sensing image segmentation," *International Journal of Applied Earth Observation and Geoinformation*, vol. 81, pp. 98-109, Sep 2019.

[52] X. L. Zhang, X. Z. Feng, P. F. Xiao, G. J. He, and L. J. Zhu, "Segmentation quality evaluation using region-based precision and recall measures for remote sensing images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 102, pp. 73-84, Apr 2015.

[53] H. Grybas, L. Melendy, and R. G. Congalton, "A comparison of unsupervised segmentation parameter optimization approaches using moderate- and high-resolution imagery," *Giscience & Remote Sensing*, vol. 54, no. 4, pp. 515-533, May 2017.

[54] B. Johnson and Z. X. Xie, "Unsupervised image segmentation evaluation and refinement using a multi-scale approach," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 4, pp. 473-483, Jul 2011.

[55] A. Troya-Galvis, P. Gancarski, N. Passat, and L. Berti-Equille, "Unsupervised Quantification of Under- and Over-Segmentation for Object-Based Remote Sensing Image Analysis," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 5, pp. 1936-1945, May 2015.

[56] L. Vincent and P. Soille, "Watersheds in Digital Spaces - an Efficient Algorithm Based on Immersion Simulations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583-598, Jun 1991.

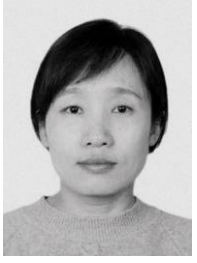
[57] K. Parvati, B. S. P. Rao, and M. M. Das, "Image Segmentation Using Gray-Scale Morphology and Marker-Controlled Watershed Transformation," *Discrete Dynamics in Nature and Society*,vol. 2008, pp. 384346, Jan 2008.

[58] P. J. Li, J. C. Guo, B. Q. Song, and X. B. Xiao, "A Multilevel Hierarchical Image Segmentation Method for Urban Impervious Surface Mapping Using Very High Resolution Imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, no. 1, pp. 103-116, Mar 2011.

[59] N. Clinton, A. Holt, J. Scarborough, L. Yan, and P. Gong, "Accuracy Assessment Measures for Object-based Image Segmentation Goodness," *Photogrammetric Engineering and Remote Sensing*, vol. 76, no. 3, pp. 289-299, Mar 2010.

Yongji Wang received the Ph.D. degree in Cartography and Geographical Information System at the State Key Laboratory of Resources and Environmental Information System, Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, China, in 2020.

He is currently a lecturer with School of Geoscience and Technology, Zhengzhou University. His study interests are remote sensing image processing and GIS spatial analysis.



Lili Wu received the Ph.D. degree in Cartography and Geographical Information System at the School of Geospatial Information, Information Engineering University, China, in 2013.

She is currently lecturer with School of Geoscience and Technology, Zhengzhou University. Her study interests are GIS spatial analysis and thematic mapping.

Qingwen Qi received the Ph.D. degree in Cartography and Geographical Information System from Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, China, in 1996.

He is currently a Professor with the Institute of Geographical Sciences and Natural Resources Research, Chinese Academy of Sciences, China. His research interests include methodology of geo-information sciences, analysis and application of geographic information and cartographic generalization automation.

Jun Wang is currently pursuing the Ph.D. degree in Cartography and Geographic Information Engineering from Shandong University of Science and Technology, Qingdao, Shandong, China.

From 2018 to the present, he was a Visiting Doctor, Institute of Geographic Science and Resources, Chinese Academy of Sciences. His research interest includes computer theory and technology, the theory of legal geography, geographic information science and methods, the theory and technology of agricultural informatization, remote sensing image segmentation technology.