

Article

Exploration of Semantic Geo-Object Recognition Based on the Scale Parameter Optimization Method for Remote Sensing Images

Jun Wang^{1,2}, Lili Jiang^{2,*}, Qingwen Qi^{2,3} and Yongji Wang⁴ 

¹ College of Geodesy and Geomatics, Shandong University of Science and Technology, Qingdao 266590, China; wangjun@reis.ac.cn

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; qiqw@igsrr.ac.cn

³ University of Chinese Academy of Sciences, Beijing 100049, China

⁴ School of Geoscience and Technology, Zhengzhou University, Zhengzhou 450001, China; wangyji.17b@igsrr.ac.cn

* Correspondence: jiangll@igsrr.ac.cn

Abstract: Image segmentation is of significance because it can provide objects that are the minimum analysis units for geographic object-based image analysis (GEOBIA). Most segmentation methods usually set parameters to identify geo-objects, and different parameter settings lead to different segmentation results; thus, parameter optimization is critical to obtain satisfactory segmentation results. Currently, many parameter optimization methods have been developed and successfully applied to the identification of single geo-objects. However, few studies have focused on the recognition of the union of different types of geo-objects (semantic geo-objects), such as a park. The recognition of semantic geo-objects is likely more crucial than that of single geo-objects because the former type of recognition is more correlated with the human perception. This paper proposes an approach to recognize semantic geo-objects. The key concept is that a single geo-object is the smallest component unit of a semantic geo-object, and semantic geo-objects are recognized by iteratively merging single geo-objects. Thus, the optimal scale of the semantic geo-objects is determined by iteratively recognizing the optimal scales of single geo-objects and using them as the initiation point of the reset scale parameter optimization interval. In this paper, we adopt the multiresolution segmentation (MRS) method to segment Gaofen-1 images and tested three scale parameter optimization methods to validate the proposed approach. The results show that the proposed approach can determine the scale parameters, which can produce semantic geo-objects.

Keywords: GEOBIA; image segmentation; parameter optimization; semantic geo-object; Gaofen-1 images



Citation: Wang, J.; Jiang, L.; Qi, Q.; Wang, Y. Exploration of Semantic Geo-Object Recognition Based on the Scale Parameter Optimization Method for Remote Sensing Images. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 420. <https://doi.org/10.3390/ijgi10060420>

Academic Editors: Wolfgang Kainz, Amr Abd-Elrahman, Zoltan Szantoi and Tao Liu

Received: 23 February 2021

Accepted: 18 June 2021

Published: 20 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Advances in satellite sensor technologies have enabled the acquisition of images with different spatial resolutions. For remote sensing images with moderate and high spatial resolutions, the traditional pixel-based approach cannot satisfy the requirements of several remote sensing applications because the same geo-object with different spectra and different geo-objects with identical spectra are present in remote sensing images. With the development of geographic object-based image analysis (GEOBIA) techniques, image classification has been enhanced due to the reduction in spectral variability within geo-objects [1–3].

Image segmentation is the first critical step in the GEOBIA framework, and the quality of image segmentation determines the accuracy of subsequent image classification [4–8]. It is challenging to perform image segmentation on remote sensing images involving complex land covers, and many segmentation methods have been developed,

such as the multiresolution segmentation (MRS) technique [9], watershed segmentation technique [10], deep learning method [11], fractal network evolutionary approach [12], and spectral angle segmentation approach [13]. In general, in these image segmentation methods, parameters must be set to control the segmentation size, shape, and attributes of an object [9,10,12,14–16]. Therefore, parameter optimization (PO) is of significance to obtain satisfactory segmentation results. PO methods have been extensively and intensively studied [17–27].

Most PO methods assume that to attain a satisfactory segmentation result, the inside of a segmented unit must be homogeneous, and the adjacent segmented units must be heterogeneous [17,18,22,25,28–33]. Therefore, most PO methods calculate the homogeneity and heterogeneity by considering certain criteria and combine these indicators into overall indicators to determine the suitable segmentation parameter(s). For example, Espindola et al. (2006) proposed a segmentation measure that adopted the area-weighted variance (WV) and global Moran's I (MI) [28]. Johnson et al. (2015) used the same indicators as those used by Espindola et al. (2006) but with different fusing strategies [22]. Specifically, the former researchers applied a sum approach, and the latter researchers applied the F-measure method. Zhang et al. (2012) used two metrics (T and D) to measure intrasegment homogeneity and intersegment heterogeneity, respectively, and achieved satisfactory results in the corresponding research area [18]. Wang et al. (2018) used the two indicators of WV and Jeffries–Matusita (JM) distance to assess the segmentation quality and determine the optimal segmentation parameter(s) [34].

At present, many PO methods have been developed and successfully applied to identify single geo-objects. However, only a few studies have attempted to recognize the union of different types of single geo-objects (semantic geo-objects), such as parks. The recognition of semantic geo-objects is likely more crucial than that of single geo-objects because the recognition of semantic geo-objects is more highly correlated with the human perception. For instance, if a person travels to a community to visit a person, the first objective is to identify the community (semantic geo-object), and the second objective is to search for the residential building and room within the community. Thus, it is meaningful to develop an approach to determine the scale parameter that recognizes semantic geo-objects.

This study aimed to establish an approach to determine the appropriate segmentation parameter(s) that recognize semantic geo-objects. The proposed approach was developed in three stages. (1) We selected several remote sensing images with different land covers and obtained a series of segmentation results in a certain interval by implementing the segmentation method. (2) We determined the appropriate segmentation scale parameters of single geo-objects by using three PO methods. (3) The segmentation results corresponding to the appropriate scale parameters were analyzed, and the iterative operation was implemented with the reset scale parameter optimization interval until the optimal scale parameters and segmentation results that satisfied the semantic geo-objects were obtained. The organizational structure of this paper is as follows. Section 1 introduces the research background, intent, and research significance of this paper. Section 2 describes the research area and workflow of the proposed approach. Section 3 presents the experimental results and describes the objective analysis of the results. Section 4 discusses the research findings and several relevant ideas, along with the limitations of this paper. Section 5 presents the concluding remarks.

2. Materials and Methods

2.1. Study Area

In this paper, we selected one scene acquired by Gaofen-1 (GF-1) on 7 August 2015, in Shenzhen, China. In general, the GF-1 satellite is equipped with six cameras: panchromatic and multispectral cameras with spatial resolutions of 2 and 8 m, respectively, and four multispectral wide cameras with a spatial resolution of 16 m [35]. Other technical specifications of the GF-1 satellite are presented in Table 1. The NNDiffuse pansharpening function of ENVI 5.3 was used to fuse the multispectral and panchromatic images with

spatial resolutions of 8 m and 2 m, respectively, into a 4-band pansharpened multispectral image with a spatial resolution of 2 m.

Table 1. Technical specifications of the GF-1 satellite.

Parameter	2-m resolution panchromatic/ 8-m resolution multispectral camera		16-m resolution multispectral camera
Spectral range (μm)	Panchromatic	0.45–0.90	0.45–0.90
		0.45–0.52	0.45–0.52
	Multispectral	0.52–0.59	0.52–0.59
		0.63–0.69	0.63–0.69
		0.77–0.89	0.77–0.89
Spatial Resolution	Panchromatic	2 m	16 m
	Multispectral	8 m	
Width	60 km (combination of two cameras)		800 km (combination of four cameras)
Revisit Period	4 d		2 d
Coverage period	41 d		4 d

Shenzhen is a coastal city in southern China adjacent to Hong Kong. The city is located south of the Tropic of Cancer, between $113^{\circ}43'$ and $114^{\circ}38'$ east longitude and between $22^{\circ}24'$ and $22^{\circ}52'$ north latitude. Shenzhen is located in the south of Guangdong Province; the eastern coast of the Pearl River Estuary is bordered by Daya Bay and Dapeng Bay in the east, Pearl River Estuary and Lingdingyang Bay in the west, Shenzhen River in the south, connected with Hong Kong, and Dongguan and Huizhou in the north. The total land area of Shenzhen is 1996.85 km^2 . The weather of Shenzhen corresponds to a dry, mild climate with abundant rainfall. The main landforms of Shenzhen include low mountains, flat platforms, and terraced hills. Plains account for 22.1% of the land area, and the forest coverage rate is 44.6%.

Four experimental areas were selected for this study and included traditional urban and suburban areas containing various land cover types. Roads, trees, water bodies, vegetation, various buildings, and other objects are present in the experimental areas. Small geo-objects are relatively clear because the experimental images have a high spatial resolution of 2 m. Four test images are shown in Figure 1. The image shown in Figure 1a contains factories, residential buildings, vegetation, and roads. The image shown in Figure 1b contains forests, rivers, factories, houses, and roads. The shown image in Figure 1c contains houses, small water bodies, vegetation, roads, and unconstructed land. The image shown in Figure 1d contains a small section of rivers, ponds, vegetation, and roads. Different combinations of geo-objects can provide several references for follow-up research. The areas of the 4 images are all $1.6 \times 1.6 \text{ km}^2$.

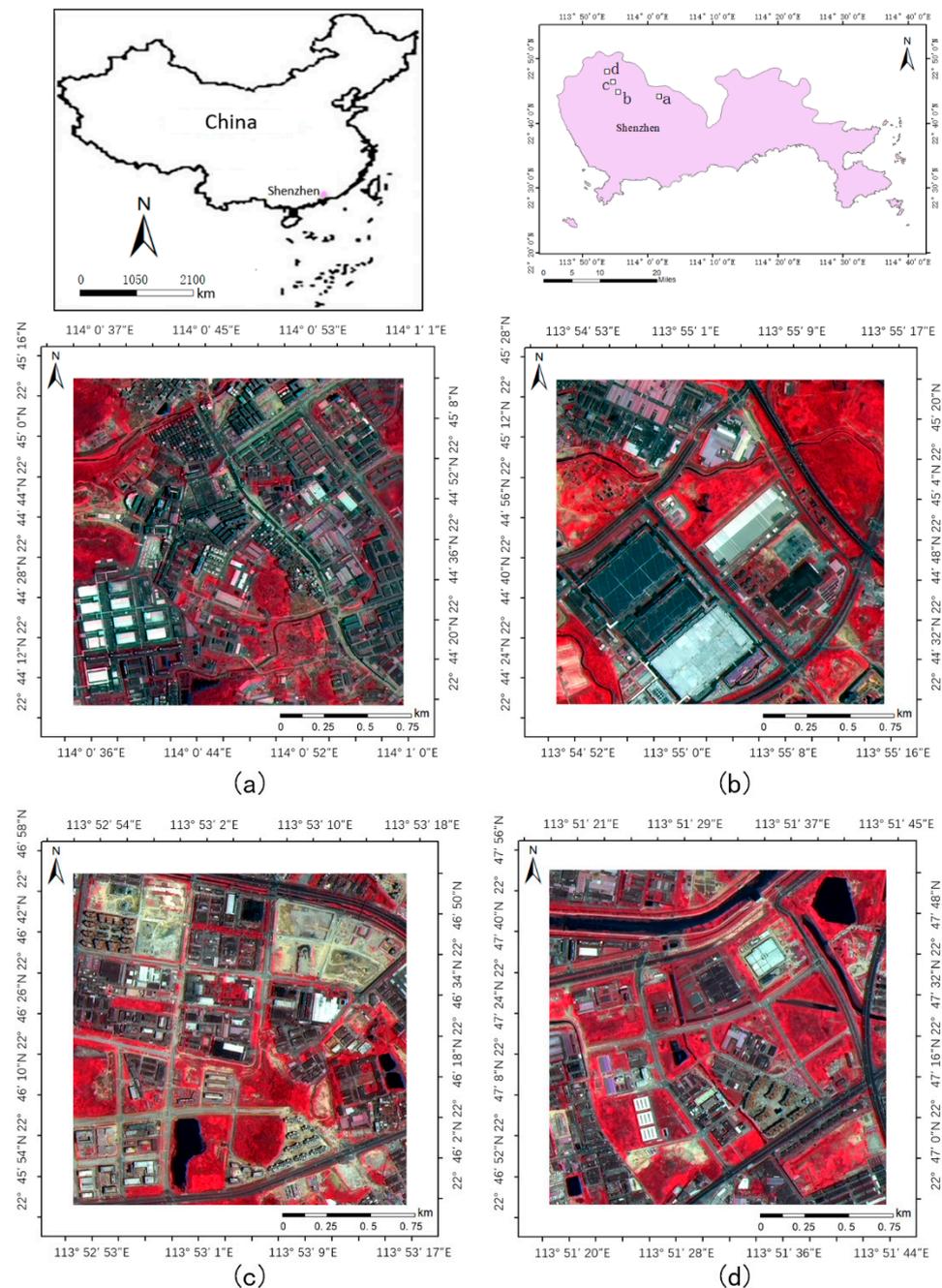


Figure 1. Images with a spatial resolution of 2 m: (a) P1, urban area, (b) P2, suburban area, (c) P3, urban area, and (d) P4, suburban area.

2.2. Methods

2.2.1. Overview

A semantic geo-object represents the union of different single geo-objects; consequently, the optimization of the scale parameters of semantic geo-objects is based on the optimization of single geo-objects. Thus, the proposed approach to recognize semantic geo-objects can be divided into three steps. The first step involves segmentation; we obtain a series of segmentation results by using MRS method in the experiment. The second step involves the scale parameter optimization of single geo-objects; we use the three PO methods reported by Johnson et al. (2015), Wang et al. (2018), and Zhang et al. (2012), referred to as Johnson's method (JSM), Wang's method (WM), and Zhang's method (ZM), respectively, to obtain the optimal scales of single objects. The third step involves the scale parameter

optimization of semantic geo-objects, and the implementation of an iterative process to determine the scale parameter(s) of semantic geo-objects. The approach to recognize the segmentation parameter(s) that can produce semantic geo-objects is illustrated in Figure 2.

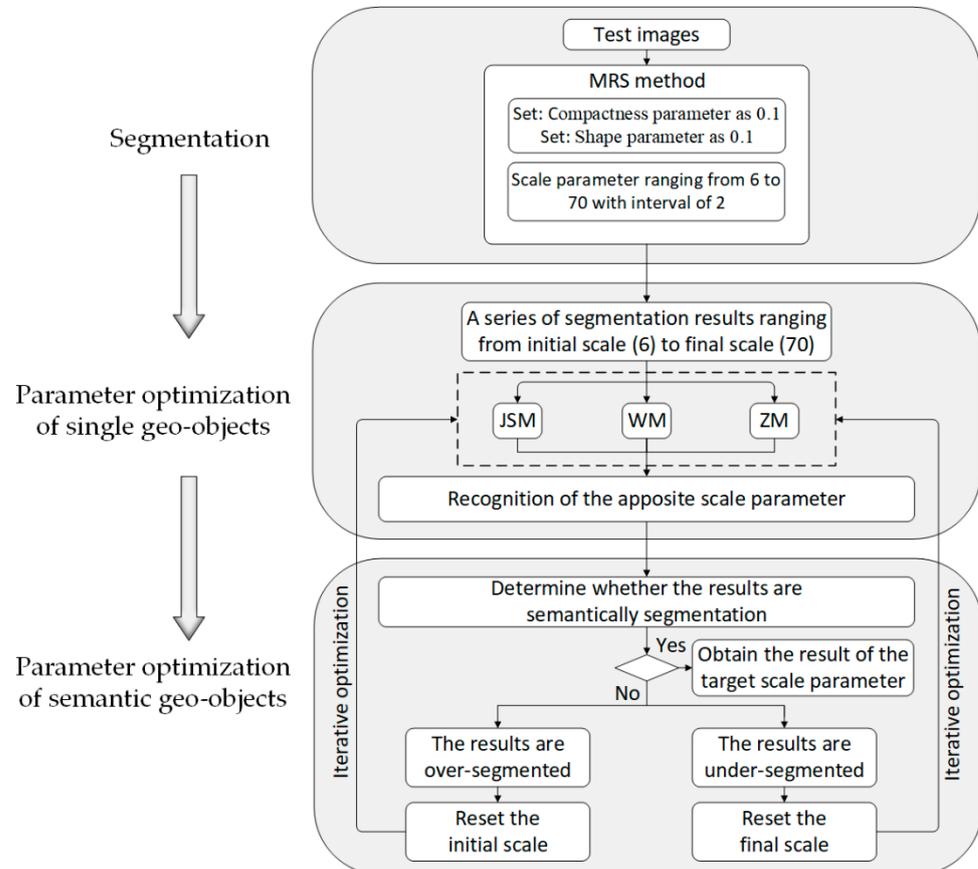


Figure 2. Process flow to determine the compatible semantic segmentation scale (the formulation of the workflow refers to the paper of Wang et al. (2021) [36]).

2.2.2. Semantic Geo-Object PO

(i) Segmentation

The multiresolution segmentation (MRS) method is used to segment the test images. The MRS method, which is embedded in eCognition Developer 9.0 software, is a bottom-up approach based on a region-merging technique; the approach selects each pixel and considers the shape, size, and attributes of the pixels within the object [37]. The method stops merging when the heterogeneity threshold is reached. The MRS method involves three parameters: scale, shape, and compactness. The scale parameter determines the maximum allowable heterogeneity, the shape parameter controls the shape and color, and the compactness parameter controls the smoothness of the image. The shape and compactness parameters are both set as 0.1 through visual analysis. The focus of this study is to determine a suitable scale parameter.

(ii) PO of single geo-objects

Based on a series of segmentation results, we use the following three PO methods to search for the appropriate segmentation scale parameter of single geo-objects.

The first PO method is JSM, proposed by Johnson et al. (2015) [22]. This method uses the WV and MI to measure the intrasegment homogeneity and intersegment heterogeneity,

respectively [22,28,38]. The WV can clarify the differences in a region. A low WV value indicates a high homogeneity. The WV can be calculated as follows:

$$WV = \frac{\sum_{i=1}^n a_i \cdot v_i}{\sum_{i=1}^n a_i} \quad (1)$$

where a_i and v_i denote the area and variance in region i , respectively, and n is the number of segments.

MI is an autocorrelation index that reflects the degree of spatial correlation [38]. A low MI value indicates a high heterogeneity. The MI can be determined as follows:

$$MI = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \hat{y}) (y_j - \hat{y})}{\left(\sum_{i=1}^n (y_i - \hat{y})^2 \right) \left(\sum_{i \neq j} w_{ij} \right)} \quad (2)$$

where n is the total number of segments; y_i and y_j are the mean gray values of segments i and j ; \hat{y} is the mean gray value of the entire image; and w_{ij} is a measure of the spatial adjacency of segments i and j [23,28]. If regions i and j are adjacent, $w_{ij} = 1$; otherwise, $w_{ij} = 0$ [28].

The MI and WV values should be normalized to 0–1 before implementing the F-measure. The normalized formula can be defined as follows:

$$WV_{norm}(MI_{norm}) = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (3)$$

where WV_{norm} and MI_{norm} represent the normalized WV and MI values, respectively; X is the WV or MI value; and X_{max} and X_{min} represent the maximum and minimum WV or MI values of all generated segmentations, respectively [22]. High WV_{norm} value represents low intrasegment homogeneity, and low MI_{norm} value represents high intersegment heterogeneity. Furthermore, WV_{norm} and MI_{norm} are calculated for each spectral band and subsequently averaged [22,39]. Finally, the F-measure is used to combine the WV and MI values to measure the “overall goodness” (OG), as follows:

$$OG_f = \left(1 + b^2 \right) \frac{MI_{norm} WV_{norm}}{b^2 MI_{norm} + WV_{norm}} \quad (4)$$

where b is the relative contribution of WV_{norm} and MI_{norm} . In this paper, we consider WV_{norm} and MI_{norm} have identical weights, i.e., $b = 1$.

The second PO method is WM, proposed by Wang et al. (2018) [34]. The homogeneity indicator and the combination strategy of this approach are similar to those in JSM, although a different heterogeneity indicator is adopted [34]. The JM distance, which is used as the heterogeneity indicator, has been demonstrated to be effective in evaluating the segmentation quality. The JM distance is typically used to measure the spectral separability between two class density functions [40,41]. Thus, the spectral heterogeneity of two adjacent fragments can be measured using the JM distance. For more information regarding the JM distance, please refer to the work of Wang et al. (2018) [34].

The third PO method is ZM, proposed by Zhang et al. (2012) [18]. This method uses two metrics (T and D) to measure the intrasegment homogeneity and intersegment heterogeneity, respectively. T is calculated as follows:

$$T(I) = \frac{1}{10S} \sqrt{R} \sum_{i=1}^R \frac{E_i}{1 + \log A_i} \quad (5)$$

The segmented image is represented by I , the image size is represented by S , the number in region i is represented by R , the mean error of the feature vector is denoted by E_i , and the area of region i is denoted by A_i .

D, which is used to measure the intersegment heterogeneity, represents a normalized variance that considers the mean feature vector [18]. D is calculated as follows:

$$D(I) = \frac{\sum_{i=1}^c \sum_{j=1}^R (m_{ij} - mm_i) / R}{\sqrt{R}} \quad (6)$$

where m_{ij} is the mean spectral value of band i in region j ; mm_i is the mean value of all the spectral mean values for the band I for all regions; and c is the number of spectral bands. The variance increases with the number of regions in the segmentation result; therefore, D can be scaled by \sqrt{R} [18].

The T and D values are normalized to 0–1 before implementing the OG strategy. OG_z is the weighted sum of T and D , calculated as follows:

$$OG_z = T + \lambda D \quad (7)$$

The value of T and D are large when oversegmentation and undersegmentation occur, respectively. Because another normalization operation is implemented in the original method, the optimal segmentation pertains to the result with the maximum OG_z . The change rate of T with respect to D can help determine the weight λ [26], which is calculated as follows:

$$\lambda = \frac{T_{max} - T_{min}}{D_{max} - D_{min}} \quad (8)$$

(iii) PO of semantic geo-objects

Based on the scale parameter optimization results of single geo-objects, we perform the PO of semantic geo-objects.

The PO of semantic geo-objects is an iterative process that proceeds from small single geo-objects to large single geo-objects to semantic geo-objects. In the optimization process, the results of the first optimization are usually associated with small and medium single geo-objects. The second optimization produces results pertaining to medium single geo-objects and semantic geo-objects. The third and subsequent optimizations gradually produce the semantic segmentation results. Usually, semantic geo-objects are formed by a combination of single geo-objects, and the choice of semantic geo-objects is derived from single geo-objects. For example, geo-objects such as residential buildings, green belts, and small pools are present in one semantic geo-object (a community). To more accurately recognize semantic geo-objects, we develop an approach to identify the semantic segmentation scale parameters. The details of the proposed approach are presented in Table 2.

Table 2. Approach to determine the optimal semantic segmentation scale.

Input: a series of segmentation results with the MRS from the initial scale (6) to the final scale (70)
Procedure:
(1) Calculate the values of the five indicators (WV, MI, JM distance, T , and D).
(2) Identify the scale with an OF value equal to the maximum values of JSM, WM, and ZM.
(3) Determine whether the segmentation results with the scale satisfy the semantic geo-objects; if the requirements are satisfied, output the result.
(4) If the segmentation results with the scale do not satisfy the semantic geo-objects, and the MRS result is oversemantic, we consider the scale to be the initial scale; otherwise, the scale is the final scale. The scale is re-evaluated from the initial scale to the final scale, and the iterative process is implemented until the target scale parameters pertaining to the semantic segmentation results are identified.
Output: target scale parameters and corresponding segmentation results

3. Results

3.1. Experimental Process

The main experimental process to recognize semantic geo-objects is as follows. First, a series of segmentation results are produced using the MRS method. The analysis of the segmentation results indicates that the test images are considerably oversegmented and undersegmented when the scale parameter is set as 6 and 70, respectively; therefore, we adjust the scale parameter to range from 6 to 70 in increments of 2. Both the compactness and shape parameters are set as 0.1. Second, to recognize single geo-objects, the PO methods of Johnson et al. (2015), Wang et al. (2018), and Zhang et al. (2012), i.e., JSM, WM, and ZM, are adopted to verify the proposed approach [18,22,34]. Traditionally, the JSM, WM, and ZM assume that the scale with maximum value of the objective function (OF) corresponds to the optimal segmentation results. Third, we determine whether the first PO results conform to semantic geo-objects. If the first PO results are in accordance with the semantic geo-objects, the scale is considered to be the most suitable for segmenting segment semantic geo-objects. If the result is oversemantic, we consider the scale to be the initial scale; otherwise, the scale is the final scale. The scale is examined iteratively until the semantic segmentation requirements are satisfied.

3.2. Results of Scale Parameter Optimization for Single Geo-Objects

To obtain the segmentation scale parameters of single geo-objects, we perform a series of calculations. We obtain 33 segmentation results by using the MRS method embedded in the eCognition Developer 9.0 software by varying the scale parameters from 6 to 70 in steps of 2 (6, 8, 10, 12, etc.). Corresponding to the scale parameters, we obtain 33 OF values of the three PO methods by calculating the WV, MI, JM distance, T, and D. The values of OF based on the JSM, WM, and ZM that correspond to the scale parameters are shown in Figure 3.

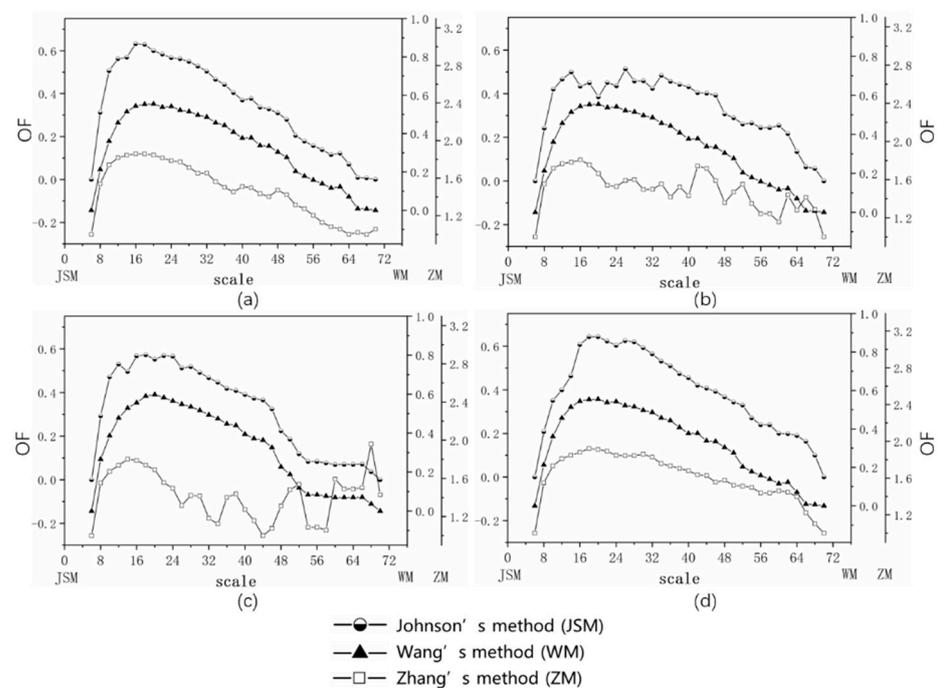


Figure 3. Segmentation optimization results with a scale of 6-70 for the four study areas: (a) P1, (b) P2, (c) P3, and (d) P4.

As shown in Figure 3, the curves exhibit a similar trend. Specifically, the OF values first increase and subsequently decrease with increasing scale. In certain cases, these values fluctuate. In addition, ZM yields OF values with several large fluctuations for P3,

which reflects the instability of the method for this image; however, ZM performs stable segmentation on the other test images. In addition, the scale with the highest OF value is considered the optimal scale.

The optimal scales and corresponding OF values obtained using the three methods are listed in Table 3. The maximum OF values for P1–P4, obtained using JSM, WM, and ZM are 0.6324, 0.5543, and 1.8595; 0.5138, 0.5846, and 1.8071; 0.5727, 0.5978, and 1.9607; and 0.6432, 0.5730, and 1.9612, respectively. The scale parameters corresponding to the maximum OF values are considered optimal. Therefore, the first optimal scales of P1–P4 using the JSM, WM, and ZM are 16, 20, and 16; 26, 24, and 16; 18, 20, and 68; and 20, 20, and 18, respectively. Because the internal indicators of the considered PO methods are different and ZM adopts different combination methods, we obtain different results for the first scale optimization.

Table 3. Maximum value of the objective functions obtained using the three PO methods and the corresponding scales.

Test	Method	OF	Scale
P1	JSM	0.6324	16
	WM	0.5543	20
	ZM	1.8595	16
P2	JSM	0.5138	26
	WM	0.5846	24
	ZM	1.8071	16
P3	JSM	0.5727	18
	WM	0.5978	20
	ZM	1.9607	68
P4	JSM	0.6432	20
	WM	0.5730	20
	ZM	1.9212	18

Subsets of the segmentation results are shown in Figure 4 to enable a visual comparison with the results of the first scale optimization performed using the JSM, WM, and ZM. For clear observation of the effect, partial regions of the four test images are shown. Figure 4 indicates that the JSM, WM, and ZM can effectively segment small and medium single geo-objects. For example, in Figure 4a–c, the small rooftops are well segmented. In Figure 4g–i, different rooftop shapes are well segmented. Figure 4m,n exhibit a satisfactory segmentation of small and medium rooftops and a grass path. In the middle parts of Figure 4s–u, three medium nature geo-objects are well segmented. However, oversegmentation occurs for several large single geo-objects and semantic geo-objects. For example, in Figure 4j–i, the rooftops of large factories are oversegmented. Figure 4d–f show the occurrence oversegmentation for forests. In Figure 4p,q, the semantic areas of unused land are segmented into small fragments. In addition, grass on the side of the houses is also oversegmented in Figure 4s–u. Figure 4v–x show that a piece of unused land containing vegetation is segmented into fragmented segments. In addition to Figure 4o,r, other subsets of the segmentation results in Figure 4 clearly demonstrate that the semantic geo-objects are not well recognized. In several remote sensing applications, because a larger area must be considered, certain semantic geo-objects are more meaningful than a single geo-object. For example, in a certain study, we may need to obtain the scope of a residential area on the remote sensing image, because different types of buildings are present in the residential area along with small green belts and other geo-objects. In this scenario, we must consider a way to directly segment the scope of the residential area, as a more effective strategy. Specifically, we must develop an approach to recognize semantic geo-objects. Because the first scale optimization does not yield satisfactory semantic segmentation results, the proposed approach searches for suitable scale parameters to obtain the semantic segmentation results, as described in the following sections.



Figure 4. Subsets of the first optimal segmentation results for the four test images: Subsets of (a–f) P1, (g–i) P2, (m–r) P3, and (s–x) P4. (the first and fourth columns correspond to the JSM results, the second and fifth columns correspond to the WM results, and the third and sixth columns correspond to the ZM results).

3.3. Results of Scale Parameter Optimization for Semantic Geo-Objects

The iterative process of searching for the scale parameter that can optimally segment semantic geo-objects is shown in Table 4. For P1, the three PO methods include four stages of scale parameter optimization, and the semantic segmentation results are generated in the fourth stage of the optimized scale determination. After the scale parameter optimal selection by the ZM, we obtain the maximum scale parameter of the undersegmented results in the fourth stage; therefore, it is inferred that the target scale parameter of the ZM corresponds to the third stage. Thus, the scale parameters of the JSM, WM, and ZM are set as 52, 58, and 48 for P1, P2, and P4, respectively, considering the three stages of scale parameter optimization, and the segmentation results are generated in the third stage. The optimized scale parameters of the JSM, WM, and ZM are set as 60, 64, and 66 for P2 and 48, 54, and 48 for P4, respectively. For P3, both the JSM and WM implement four stages of scale parameter optimization, and we use ZM to optimize the scale parameters after two stages. The optimized scale parameters of the JSM, WM, and ZM are set as 44, 52, and 68 for P3, respectively. In addition, the scale parameters produced by the four-stage scale parameter optimization are not larger than those produced by the three-stage optimization in certain cases for different remote sensing images.

To further validate the proposed approach, Figure 5 shows the segmentation results with the scales produced using the proposed approach. To observe the overall effect, we show the entire area of the four test images.

Table 4. Process of semantic scale parameter optimization.

	First Scale Optimization (JSM, WM, ZM)	Second Scale Optimization (JSM, WM, ZM)	Third Scale Optimization (JSM, WM, ZM)	Fourth Scale Optimization (JSM, WM, ZM)
P1	16, 20, 16	30, 36, 26	48, 50, 48	52, 58, 70
P2	26, 24, 16	42, 48, 42	60, 64, 66	—, —, —
P3	18, 20, 68	40, 36, 68	44, 46, —	44, 52, —
P4	20, 20, 18	36, 34, 30	48, 54, 48	—, —, —



Figure 5. Semantic segmentation results for the four test images. Segmentation results for (a–c) P1, (d–f) P2, (g–i) P3, and (j–l) P4. The first, second, and third columns show the results of JSM, WM, and ZM, respectively.

As shown in Figure 5, most semantic geo-objects are well recognized. In Figure 5a–c, several residential areas containing different types of single geo-objects are effectively identified. Figure 5d–f exhibit satisfactory semantic segmentation for large homogenous rooftops, large yards, and vegetation belts. In the upper-middle part of Figure 5h,i, a residential zone, which is a semantic geo-object with different spectral features that contain houses, a green belt, and pools, is well segmented. In the upper-right area, Figure 5h,i show the effective delineation of the unused land containing different types of geo-objects. Figure 5g displays a slightly inferior result. Figure 5j–l show the successful identification of a long channel in the image range; this type of artificial channel is a typical semantic

geo-object in reality. In Figure 5a–c, small rooftops with notable spectral features in the test images are separated from the surrounding geo-objects, and a large rooftop is not fully recognized in Figure 5d–f, because of the large spectral contrast. These cases correspond to unavoidable situations in the current experiments because effective segmentation is difficult when the spectral contrast is extremely high on the surface of a large single geo-object or between two adjacent geo-objects [42]. Although minor imperfections are noted for the JSM, WM, and ZM, satisfactory semantic segmentation can be realized using the optimized target scale parameters in the four test images.

4. Discussion

Image segmentation is a crucial task because it can provide objects for GEOBIA. Effective segmentation must be ensured to enable subsequent image interpretation. Intuitively, it is meaningful to obtain target objects, such as roads, houses, and transportation [43–46]. Hence, scale parameter optimization is a key step in achieving the desirable segments. In general, the segmentation result obtained using the PO method can be quantitatively evaluated using discrepancy measurement methods. Many discrepancy measurement approaches have been successfully applied to quantitatively evaluate single geo-objects produced by PO methods, as described in Section 2.2.2 [29,47–56]. However, only a few studies have quantitatively evaluated semantic segmentation results because the understanding of semantic geo-objects is subjective, and experts often differ in their opinions regarding the definition of semantic geo-objects. Currently, it is difficult to quantitatively evaluate the results of semantic segmentation based on a single criterion. Therefore, we used a visual evaluation method in this paper. Future research can be aimed at establishing a quantitative evaluation method for semantic segmentation.

In certain cases, to recognize semantic geo-objects, a weakened spectral difference is required within a semantic geo-object. From this perspective, low and medium-spatial-resolution images may be more suitable for segmenting semantic geo-objects. However, the geo-objects in cities are often relatively small, and high spatial resolution images can provide abundant and detailed geo-object information. In such cases, low- and medium-spatial-resolution images are suboptimal for segmenting single geo-objects. For example, images with a high spatial resolution can help identify small urban geo-objects, although this identification cannot be realized through images with a low spatial resolution. In practical applications, the combination of single and semantic geo-objects is most useful. For example, if the objective is to visit a building in a park, the location of the park is first identified. When we arrive at the park, our focus shifts to the building. Future work will be focused on the recognition and union of single and semantic geo-objects.

Our results (Table 4) show that most of the optimal scales were obtained in the third or fourth stage of the PO. Thus, generally, semantic geo-objects can be effectively segmented after three iterative stages of PO. In addition, most of the scales identified in the previous stages caused the oversegmentation of semantic geo-objects and slight undersegmentation. Thus, the scale obtained in the first optimization was used as the initial scale in most cases. However, when using the ZM for optimization, the scales recognized in the first stage produced larger segments than the semantic criterion for one image. In such cases, it is necessary to reset the final scale of the scale range and optimize the parameter again. In addition, this paper only assessed three PO methods. In future research, more PO methods can be considered to validate the proposed approach and enhance the method of recognizing semantic geo-objects.

A widespread phenomenon in remote sensing images is that the semantic geo-objects considered in this paper, such as a park, may have different types of geo-objects, although the features of the park are general and manifest as a certain homogeneity in the types of geo-objects in a park. This issue can be observed in real scenes, in which many features in parks exhibit similarities. This information is a realistic theoretical basis for exploring the feasibility of the proposed technique. A semantic geo-object is often formed by a series of single geo-objects, e.g., a semantic area contains several geo-objects with a common shape,

texture, and color. These common features are a manifestation of the homogeneity of a semantic geo-object. Future work will be aimed at techniques to investigate the internal homogeneity of a semantic geo-object and to enhance the heterogeneity of a semantic geo-object relative to its surroundings to more accurately recognize semantic geo-objects.

The key objective of this study was to establish a technique to search for the suitable scale parameter of semantic geo-objects in high spatial resolution images. In this paper, three scale parameter optimization methods were used to evaluate the feasibility of the proposed approach. After the first PO, the optimized scale parameters of single geo-objects were obtained. Based on these optimized scales, the optimization was iteratively performed. Finally, satisfactory semantic segmentation results were obtained for four test images. However, at present, no comparable techniques are available for the proposed exploratory approach. We believe that this study can provide a concept for further research and encourage other researchers to provide a similar strategy for comparison. The proposed approach was developed by considering actual research and applications to obtain the spatial scope of a region in a real scene and is thus believed to be of significance for application. We hope that this paper can provide references for future semantic segmentation research.

5. Conclusions

An approach to search for the optimal scale parameters of semantic segmentation was developed. The scales were searched through iterative PO by continuously reducing the range of optimization. This paper used the MRS algorithm as the segmentation method and considered three PO methods (JSM, WM, and ZM) to validate the proposed approach. GF-1 images were used as the test images, and the visual experiments demonstrated the efficiency of the proposed approach in determining the scale that can generate satisfactory semantic geo-objects. Future work will be aimed at enhancing semantic segmentation scale parameter determination and evaluation methods.

Author Contributions: Conceptualization, Jun Wang, and Lili Jiang; methodology, Jun Wang; validation, Jun Wang; investigation, Jun Wang, and Qingwen Qi; data curation, Yongji Wang; writing—original draft preparation, Jun Wang; writing—review and editing, Jun Wang, and Lili Jiang; visualization, Jun Wang; supervision, Lili Jiang. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the National Key Research and Development Program of China, project number 2017YFB0503500; and the Strategic Priority Research Program of the Chinese Academy of Sciences, project number XDA19040402.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors would like to thank the reviewers and editors for valuable comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Johansen, K.; Arroyo, L.A.; Phinn, S.; Witte, C. Comparison of Geo-Object Based and Pixel-Based Change Detection of Riparian Environments using High Spatial Resolution Multi-Spectral Imagery. *Photogramm. Eng. Rem. S* **2010**, *76*, 123–136. [[CrossRef](#)]
2. Blaschke, T.; Hay, G.J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Feitosa, R.Q.; van der Meer, F.; van der Werff, H.; van Coillie, F.; et al. Geographic Object-Based Image Analysis—Towards a new paradigm. *ISPRS J. Photogramm.* **2014**, *87*, 180–191. [[CrossRef](#)]
3. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm.* **2010**, *65*, 2–16. [[CrossRef](#)]
4. Liu, J.; Li, P.J.; Wang, X. A new segmentation method for very high resolution imagery using spectral and morphological information. *ISPRS J. Photogramm.* **2015**, *101*, 145–162. [[CrossRef](#)]

5. Yang, J.; He, Y.H.; Caspersen, J.; Jones, T. A discrepancy measure for segmentation evaluation from the perspective of object recognition. *ISPRS J. Photogramm.* **2015**, *101*, 186–192. [[CrossRef](#)]
6. Cleve, C.; Kelly, M.; Kearns, F.R.; Morlitz, M. Classification of the wildland-urban interface: A comparison of pixel- and object-based classifications using high-resolution aerial photography. *Comput. Environ. Urban* **2008**, *32*, 317–326. [[CrossRef](#)]
7. Myint, S.W.; Gober, P.; Brazel, A.; Grossman-Clarke, S.; Weng, Q.H. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sens. Environ.* **2011**, *115*, 1145–1161. [[CrossRef](#)]
8. Gao, Y.; Mas, J.F.; Maathuis, B.H.P.; Zhang, X.M.; Van Dijk, P.M. Comparison of pixel-based and object-oriented image classification approaches—A case study in a coal fire area, Wuda, Inner Mongolia, China. *Int. J. Remote Sens* **2006**, *27*, 4039–4055. [[CrossRef](#)]
9. Benz, U.C.; Hofmann, P.; Willhauck, G.; Lingenfelder, I.; Heynen, M. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS J. Photogramm.* **2004**, *58*, 239–258. [[CrossRef](#)]
10. Wagner, B.; Dinges, A.; Muller, P.; Haase, G. Parallel Volume Image Segmentation with Watershed Transformation. *Lect. Notes Comput. Sci.* **2009**, *5575*, 420–429.
11. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE T Pattern Anal.* **2018**, *40*, 834–848. [[CrossRef](#)]
12. Happ, P.N.; Ferreira, R.S.; Bentes, C.; Costa, G.A.O.P.; Feitosa, R.Q. Multiresolution Segmentation: A Parallel Approach for High Resolution Image Segmentation in Multicore Architectures. In *Int Arch Photogramm*; Addink, E.A., VanCoillie, F.M.B., Eds.; Copernicus Gesellschaft Mbh: Göttingen, German, 2010; p. 38-4-C7.
13. Yang, J.; He, Y.H.; Caspersen, J. Region merging using local spectral angle thresholds: A more accurate method for hybrid segmentation of remote sensing images. *Remote Sens. Environ.* **2017**, *190*, 137–148. [[CrossRef](#)]
14. Doulamis, A.D.; Doulamis, N.D.; Kollias, S.D. A fuzzy video content representation for video summarization and content-based retrieval. *Signal Process.* **2000**, *80*, 1049–1067. [[CrossRef](#)]
15. Vincent, L.; Soille, P. Watersheds in Digital Spaces—An Efficient Algorithm Based on Immersion Simulations. *IEEE T Pattern Anal.* **1991**, *13*, 583–598. [[CrossRef](#)]
16. Shi, J.B.; Malik, J. Normalized cuts and image segmentation. *IEEE T Pattern Anal.* **2000**, *22*, 888–905. [[CrossRef](#)]
17. Johnson, B.; Xie, Z.X. Unsupervised image segmentation evaluation and refinement using a multi-scale approach. *ISPRS J. Photogramm.* **2011**, *66*, 473–483. [[CrossRef](#)]
18. Zhang, X.L.; Xiao, P.F.; Feng, X.Z. An Unsupervised Evaluation Method for Remotely Sensed Imagery Segmentation. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 156–160. [[CrossRef](#)]
19. Belgiu, M.; Dragut, L. Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. *ISPRS J. Photogramm.* **2014**, *96*, 67–75. [[CrossRef](#)] [[PubMed](#)]
20. Bock, S.; Immitzer, M.; Atzberger, C. On the Objectivity of the Objective Function-Problems with Unsupervised Segmentation Evaluation Based on Global Score and a Possible Remedy. *Remote Sens.* **2017**, *9*, 769. [[CrossRef](#)]
21. Dragut, L.; Csillik, O.; Eisank, C.; Tiede, D. Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS J. Photogramm.* **2014**, *88*, 119–127. [[CrossRef](#)] [[PubMed](#)]
22. Johnson, B.A.; Bragais, M.; Endo, I.; Magcale-Macandog, D.B.; Macandog, P.B.M. Image Segmentation Parameter Optimization Considering Within- and Between-Segment Heterogeneity at Multiple Scale Levels: Test Case for Mapping Residential Areas Using Landsat Imagery. *ISPRS Int. Geo-Inf.* **2015**, *4*, 2292–2305. [[CrossRef](#)]
23. Grybas, H.; Melendy, L.; Congalton, R.G. A comparison of unsupervised segmentation parameter optimization approaches using moderate- and high-resolution imagery. *GISci. Remote Sens.* **2017**, *54*, 515–533. [[CrossRef](#)]
24. Dragut, L.; Tiede, D.; Levick, S.R. ESP: A tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 859–871. [[CrossRef](#)]
25. Chabrier, S.; Emile, B.; Rosenberger, C.; Laurent, H. Unsupervised performance evaluation of image segmentation. *Eur. J. Appl. Signal Process.* **2006**. [[CrossRef](#)]
26. Kim, M.; Madden, M.; Warner, T. *Estimation of Optimal Image Object Size for the Segmentation of Forest Stands with Multispectral IKONOS Imagery*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 291–307.
27. Wang, Y.; Tian, Z.; Qi, Q.; Wang, J. Double-Variance Measures: A Potential Approach to Parameter Optimization of Remote Sensing Image Segmentation. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 1.
28. Espindola, G.M.; Camara, G.; Reis, I.A.; Bins, L.S.; Monteiro, A.M. Parameter selection for region-growing image segmentation algorithms using spatial autocorrelation. *Int. J. Remote Sens.* **2006**, *27*, 3035–3040. [[CrossRef](#)]
29. Yang, J.; Li, P.J.; He, Y.H. A multi-band approach to unsupervised scale parameter selection for multi-scale image segmentation. *ISPRS J. Photogramm.* **2014**, *94*, 13–24. [[CrossRef](#)]
30. Yang, J.; He, Y.H.; Weng, Q.H. An Automated Method to Parameterize Segmentation Scale by Enhancing Intra-segment Homogeneity and Inter-segment Heterogeneity. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1282–1286. [[CrossRef](#)]
31. Kim, M.; Madden, M.; Warner, T.A. Forest Type Mapping using Object-specific Texture Measures from Multispectral Ikonos Imagery: Segmentation Quality and Image Classification Issues. *Photogramm. Eng. Remote Sens.* **2009**, *75*, 819–829. [[CrossRef](#)]
32. Faur, D.; Gavat, I.; Datcu, M. Salient Remote Sensing Image Segmentation Based on Rate-Distortion Measure. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 855–859. [[CrossRef](#)]
33. Corcoran, P.; Winstanley, A.; Mooney, P. Segmentation performance evaluation for object-based remotely sensed image analysis. *Int. J. Remote Sens.* **2010**, *31*, 617–645. [[CrossRef](#)]

34. Wang, Y.J.; Qi, Q.W.; Liu, Y. Unsupervised Segmentation Evaluation Using Area-Weighted Variance and Jeffries-Matusita Distance for Remote Sensing Images. *Remote Sens.* **2018**, *10*, 1193. [[CrossRef](#)]
35. Wang, J.; Jiang, L.; Wang, Y.; Qi, Q. An Improved Hybrid Segmentation Method for Remote Sensing Images. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 543. [[CrossRef](#)]
36. Wang, C.-K.; Fareed, N. Mapping Drainage Structures Using Airborne Laser Scanning by Incorporating Road Centerline Information. *Remote Sens.* **2021**, *13*, 463. [[CrossRef](#)]
37. Baatz, M.; Schäpe, A. Multiresolution Segmentation—An Optimization Approach for High Quality Multi-Scale Image Segmentation. In *Angewandte Geographische Informations-Verarbeitung*; Strobl, J., Blaschke, T., Griesebner, G., Eds.; Wichmann Verlag: Karlsruhe, Germany, 2000; pp. 12–23.
38. Mikelbank, B.A. Quantitative geography: Perspectives on spatial data analysis. *Geogr. Anal.* **2001**, *33*, 370–372. [[CrossRef](#)]
39. Wang, Y.J.; Meng, Q.Y.; Qi, Q.W.; Yang, J.; Liu, Y. Region Merging Considering Within- and Between-Segment Heterogeneity: An Improved Hybrid Remote-Sensing Image Segmentation Method. *Remote Sens.* **2018**, *10*, 781. [[CrossRef](#)]
40. Richards, J.A. *Remote Sensing Digital Image Analysis: An Introduction*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 47–54.
41. Schmidt, K.S.; Skidmore, A.K. Spectral discrimination of vegetation types in a coastal wetland. *Remote Sens. Environ.* **2003**, *85*, 92–108. [[CrossRef](#)]
42. Canovas-Garcia, F.; Alonso-Sarria, F. A local approach to optimize the scale parameter in multiresolution segmentation for multispectral imagery. *Geocarto Int.* **2015**, *30*, 937–961. [[CrossRef](#)]
43. Chaudhuri, D.; Kushwaha, N.K.; Samal, A. Semi-Automated Road Detection From High Resolution Satellite Images by Directional Morphological Enhancement and Segmentation Techniques. *IEEE J. Stars* **2012**, *5*, 1538–1544. [[CrossRef](#)]
44. Wang, M.; Yuan, S.G.; Pan, J. Building Detection in High Resolution Satellite Urban Image Using Segmentation, Corner Detection Combined with Adaptive Windowed Hough Transform. *Int. Geosci. Remote Sens.* **2013**, 508–511. [[CrossRef](#)]
45. Chen, C.T.; Su, C.Y.; Kao, W.C. An enhanced segmentation on vision-based shadow removal for vehicle detection. In Proceedings of the 2010 International Conference on Green Circuits and Systems (ICGCS), Shanghai, China, 21–23 June 2010; IEEE: Shanghai, China, 2010; pp. 679–682.
46. Liu, G.; Zhang, Y.S.; Zheng, X.W.; Sun, X.; Fu, K.; Wang, H.Q. A New Method on Inshore Ship Detection in High-Resolution Satellite Images Using Shape and Context Information. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 617–621. [[CrossRef](#)]
47. Liu, Y.; Bian, L.; Meng, Y.H.; Wang, H.P.; Zhang, S.F.; Yang, Y.N.; Shao, X.M.; Wang, B. Discrepancy measures for selecting optimal combination of parameter values in object-based image analysis. *ISPRS J. Photogramm.* **2012**, *68*, 144–156. [[CrossRef](#)]
48. Cheng, J.H.; Bo, Y.C.; Zhu, Y.X.; Ji, X.L. A novel method for assessing the segmentation quality of high-spatial resolution remote-sensing images. *Int. J. Remote Sens.* **2014**, *35*, 3816–3839. [[CrossRef](#)]
49. Costa, H.; Foody, G.M.; Boyd, D.S. Integrating User Needs on Misclassification Error Sensitivity into Image Segmentation Quality Assessment. *Photogramm. Eng. Remote Sens.* **2015**, *81*, 451–459. [[CrossRef](#)]
50. Zhang, X.L.; Xiao, P.F.; Feng, X.Z.; Feng, L.; Ye, N. Toward Evaluating Multiscale Segmentations of High Spatial Resolution Remote Sensing Images. *IEEE T Geosci. Remote* **2015**, *53*, 3694–3706. [[CrossRef](#)]
51. Su, T.F.; Zhang, S.W. Local and global evaluation for remote sensing image segmentation. *ISPRS J. Photogramm.* **2017**, *130*, 256–276. [[CrossRef](#)]
52. Yang, J.; He, Y.H.; Caspersen, J.P.; Jones, T.A. Delineating Individual Tree Crowns in an Uneven-Aged, Mixed Broadleaf Forest Using Multispectral Watershed Segmentation and Multiscale Fitting. *IEEE J. Stars* **2017**, *10*, 1390–1401. [[CrossRef](#)]
53. Moller, M.; Birger, J.; Gidudu, A.; Glasser, C. A framework for the geometric accuracy assessment of classified objects. *Int. J. Remote Sens.* **2013**, *34*, 8685–8698. [[CrossRef](#)]
54. Tian, J.; Chen, D.M. Optimization in multi-scale segmentation of high-resolution satellite images for artificial feature recognition. *Int. J. Remote Sens.* **2007**, *28*, 4625–4644. [[CrossRef](#)]
55. Clinton, N.; Holt, A.; Scarborough, J.; Yan, L.; Gong, P. Accuracy Assessment Measures for Object-based Image Segmentation Goodness. *Photogramm. Eng. Remote Sens.* **2010**, *76*, 289–299. [[CrossRef](#)]
56. Persello, C.; Bruzzone, L. A Novel Protocol for Accuracy Assessment in Classification of Very High Resolution Images. *IEEE T Geosci. Remote* **2010**, *48*, 1232–1244. [[CrossRef](#)]