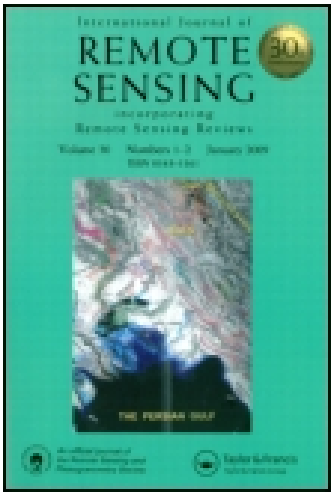


International Journal of Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tres20

A novel method for assessing the segmentation quality of high-spatial resolution remote-sensing images

Jiehai Cheng[abc], Yanchen Bo[ac], Yuxin Zhu[acd] & Xiaole Ji[ac]

[a] State Key Laboratory of Remote Sensing Science, Research Centre for Remote Sensing and GIS, and School of Geography, Beijing Normal University, Beijing 100875, China

[b] School of Surveying & Land Information Engineering, Henan Polytechnic University, Jiaozuo 454003, China

[c] Beijing Key Laboratory of Remote Sensing Science, Beijing Normal University, Beijing 100875, China

[d] School of Urban and Environmental Sciences, Huaiyin Normal University, Huaiyin 223300, China

Published online: 30 May 2014.

To cite this article: Jiehai Cheng, Yanchen Bo, Yuxin Zhu & Xiaole Ji (2014) A novel method for assessing the segmentation quality of high-spatial resolution remote-sensing images, International Journal of Remote Sensing, 35:10, 3816-3839, DOI: 10.1080/01431161.2014.919678

To link to this article: http://dx.doi.org/10.1080/01431161.2014.919678

International Journal of Remote Sensing, 2014
Vol. 35, No. 10, 3816–3839, http://dx.doi.org/10.1080/01431161.2014.919678

Taylor & Francis
Taylor & Francis Group

A novel method for assessing the segmentation quality of high-spatial resolution remote-sensing images

Jiehai Cheng[a,b,c], Yanchen Bo[a,c]*, Yuxin Zhu[a,c,d], and Xiaole Ji[a,c]

[a]*State Key Laboratory of Remote Sensing Science, Research Centre for Remote Sensing and GIS, and School of Geography, Beijing Normal University, Beijing 100875, China;* [b]*School of Surveying & Land Information Engineering, Henan Polytechnic University, Jiaozuo 454003, China;* [c]*Beijing Key Laboratory of Remote Sensing Science, Beijing Normal University, Beijing 100875, China;* [d]*School of Urban and Environmental Sciences, Huaiyin Normal University, Huaiyin 223300, China*

(*Received 26 August 2013; accepted 20 April 2014*)

Image segmentation quality significantly affects subsequent image classification accuracy. It is necessary to develop effective methods for assessing image segmentation quality. In this paper, we present a novel method for assessing the segmentation quality of high-spatial resolution remote-sensing images by measuring both area and position discrepancies between the delineated image region (DIR) and the actual image region (AIR) of a scene object. In comparison with the most frequently used area coincidence-based methods, our method can assess the segmentation quality more objectively in that it takes into consideration all image objects intersecting with the AIR of a scene object. Moreover, the proposed method is more convenient to use than the existing boundary coincidence-based methods in that the calculation of the distance between the boundary of the image object and that of the corresponding AIR of the scene object is not required. Another benefit of this method over the two types of method above is that the assessment procedure of the segmentation quality can be conducted with less human intervention. The obtained optimal segmentation result can ensure maximal delineation of the extent of scene objects and can be beneficial to subsequent classification operations. The experimental results have shown the effectiveness of this new method for both segmentation quality assessment and optimal segmentation parameter selection.

1. Introduction

Since the 1990s, more and more high-spatial resolution remote-sensors have been in operation. The massive remote-sensing images acquired by these sensors have been widely applied in many fields such as forest change detection (Desclée, Bogaert, and Defourny 2006) and land use monitoring (Myint et al. 2011). Remote-sensing images with high spatial resolution offer more detailed spatial information on the earth's surface than middle- and coarse-spatial resolution images. Traditional pixel-based image analysis approaches face serious problems in analysing and classifying high-spatial resolution remote-sensing images (Campbell 2002), because they only take spectral information (pixel values) as a basis to analyse and classify remote-sensing images while neglecting both spatial information and a group of pixels which should be considered together as an object (Benz et al. 2004; Walter, 2004; Castillejo-González et al. 2009). As an alternative approach, Object-based image analysis (OBIA) has been proposed (Blaschke et al. 2000; Schiewe, Tufte, and Ehlers 2001; Hay et al. 2005; Hay and Castilla 2006; Yu et al. 2006;

*Corresponding author. Email: boyc@bnu.edu.cn

Im, Jensen, and Tullis 2008; Hölbling et al. 2012). One of the fundamental steps in OBIA is image segmentation, in which the image is partitioned into a number of meaningful image objects (or segments) (Marçal and Rodrigues 2009). Ideally, an image object should match the corresponding actual image regions (AIRs) of a scene object completely. In practice, however, an image object often mismatches the corresponding AIR of a scene object because of the influence of image quality, landscape heterogeneity, and segmentation algorithms (Carleer, Debeir, and Wolff 2005; Lang, Schöpfer, and Langanke 2009). The degree of mismatch represents the segmentation quality, which significantly affects subsequent image classification accuracy (Dorren, Maier, and Seijmonsbergen 2003; Neubert and Meinel 2003; Blaschke, Burnett, and Pekkarinen 2004; Meinel and Neubert 2004; Weidner 2008; Kim, Madden, and Warner 2009; Blaschke 2010; Clinton et al. 2010). Therefore, it is necessary to develop effective methods for assessing image segmentation quality.

At present, there are two types of image segmentation quality assessment methods: unsupervised methods (or empirical goodness methods) and supervised methods (or empirical discrepancy methods) (Zhang 1996; Zhang, Fritts, and Goldman 2008; Corcoran, Winstanley, and Mooney 2010; Johnson and Xie 2011). Unsupervised methods (Chabrier et al. 2006; Espindola et al. 2006; Kim, Madden, and Warner 2009; Corcoran, Winstanley, and Mooney 2010; Johnson and Xie 2011) assess a segmented image based on how well it matches a broad set of characteristics of segmented images as desired by humans (Zhang, Fritts, and Goldman 2008), and use certain quality criteria which are typically established in agreement with human perception of what makes a good segmentation (Zhang 1997; Chabrier et al. 2006). Supervised methods (Delves et al. 1992; Yang et al. 1995; Abeyta and Franklin 1998; Lucieer and Stein 2002; Paglieroni 2004; Carleer, Debeir, and Wolff 2005; Zhan et al. 2005; Möller, Lymburner, and Volk 2007; Radoux and Defourny 2007; Tian and Chen 2007; Radoux and Defourny 2008; Weidner 2008) involve comparing a segmentation result to a ground truth. The reliability of unsupervised methods depends on the rationality of the proposed quality criteria. However, desirable quality criteria for unsupervised methods are usually chosen subjectively, and moreover, it should be assured that the criteria are not used by segmentation algorithms to avoid a bias assessment (Carleer, Debeir, and Wolff 2005). Supervised methods can be both objective and quantitative (Zhang 1996). As long as ground truths or reference segmentations can be obtained, supervised methods should be desirable for the assessment of segmentation results.

Among supervised methods, there are two main types of image segmentation quality assessment: area coincidence-based methods (Yang et al. 1995; Lucieer and Stein 2002; Zhan et al. 2005; Möller, Lymburner, and Volk 2007; Tian and Chen 2007; Weidner 2008) and boundary coincidence-based methods (Abeyta and Franklin 1998; Radoux and Defourny 2007, 2008). Area coincidence-based methods select the image object with the largest area or the image objects with the dominant intersection area to assess the segmentation quality by comparing its/their area(s) with that of the corresponding AIR of a scene object. In fact the image objects, which have not been utilized in the area coincidence-based methods, are not dispensable for the assessment of segmentation quality. For example, there is a segmentation result in which all image objects fall completely within the AIR of a scene object. Such segmentation result is beneficial to subsequent classification operation because there is no area discrepancy between the merging region of all image objects and the AIR of the scene object. If only one image object or parts are selected for the assessment of the segmentation

result, the conclusion is that the segmentation result is unsuitable. Therefore, most existing area coincidence-based methods are not objective in the assessment of image segmentation quality. Boundary coincidence-based methods are used to assess segmentation quality by calculating the distance between the boundary of the image object and that of the corresponding AIR of a scene object. The shorter the distance, the better the segmentation quality. The result of distance calculation depends on the way points are selected on the boundaries and the number of these points. However, there is no better way to select the points than by applying the experience of experts. As a consequence, the result of distance calculation could be unreliable. Another common problem with the use of boundary coincidence-based methods is that it is complicated to identify the boundary of an image object that corresponds to the boundary of the corresponding AIR of a scene object, because the geometric shape of the boundary of the image object is usually different from that of the boundary of the corresponding AIR of the scene object, which makes direct calculation of the distance difficult between the boundary of the image object and that of the corresponding AIR of the scene object. Because of the above-mentioned problems, the application of boundary coincidence-based methods is limited.

The goals of image segmentation are to perform the automated delineation of the image region of a scene object and to serve subsequent classification operations (Marçal and Rodrigues 2009; Blanchard, Jakubowski, and Kelly 2011). The quality of the segmentation result depends on the degree of discrepancy between the DIR and the corresponding AIR of a scene object. Many researchers have suggested that area- and position-based indices for segmentation quality assessment could be combined by a range of normalization or standardization methods (Levine and Nazif 1982; Möller, Lymburner, and Volk 2007; Weidner 2008; Clinton et al. 2010). According to the object-fate analysis method (Schöpfer and Lang 2006; Schöpfer, Lang, and Albrecht 2008; Lang, Schöpfer, and Langanke 2009; Albrecht, Lang, and Hölbling 2010), image objects intersecting with the corresponding AIR of a scene object can be categorized into three types: ‘good’, ‘expanding’, and ‘invading’ objects (Figure 1), and segmentation quality is assessed by comparing the numbers of the three types of image object. However, the object-fate analysis method only considers the number of the three types of image objects, which implies that either all the image objects are of the same area or that those of different size have the same effects on segmentation quality, but this is not the case in practice. Inspired by the principle of the object-fate analysis method, we propose a novel method for simultaneously assessing image segmentation quality and selecting the optimal segmentation result. Unlike the object-fate analysis method, our method considers both area and position discrepancies between the three types of image object and the corresponding AIR of the scene object for segmentation quality assessment. In regard to area discrepancies, both the area commission and area omission errors in delineating the image region of a scene object are involved. In regard to position discrepancies, the distances between the centroids of the three types of image object and that of the corresponding AIR of a scene object are computed (see Figure 1).

The remainder of this article is organized as follows. Section 2 describes the principle of the proposed method, while the data and segmentation operation used are described in Section 3; the experimental results are reported at the end of that section. Section 4 discusses the experimental results while Section 5 concludes.

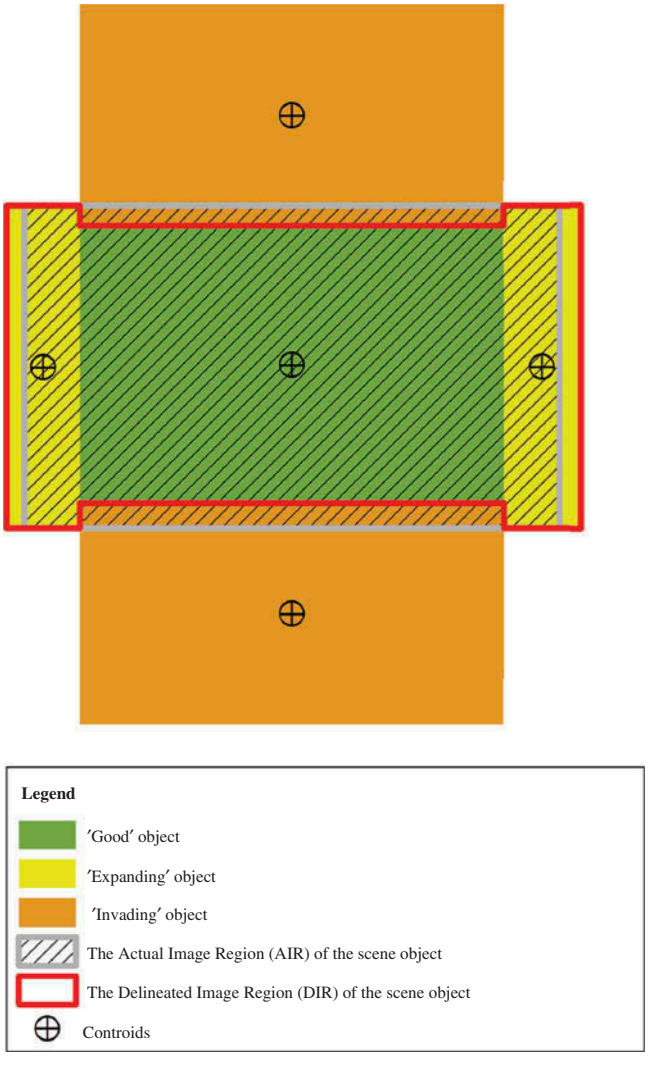


Figure 1. Spatial relationships between image objects and the corresponding AIR of the scene object (modified from Lang, Schöpfer, and Langanke (2009)).

2. Method

2.1. Object metrics

In this paper, the terms good object, expanding object, and invading object – derived from the object-fate analysis method – are exploited for segmentation quality assessment (Figure 1). A good object refers to an image object that falls completely within the corresponding AIR of the scene object, and is optimal for delineation of the corresponding AIR of the scene object because it does not lose its own regions and contains no neighbouring regions. An expanding object refers to an image object that exceeds the corresponding AIR of the scene object. Its centroid lies within the corresponding AIR of the scene object and the area of the overlapping region is greater than 50% of the area of the image object. An expanding object contributes to delineation of the corresponding

AIR of the scene object because its main part falls within the corresponding AIR of the scene object. An invading object refers to an image object that exceeds the corresponding AIR of the scene object, with its centroid lying outside the corresponding AIR of the scene object and the area of the overlapping region less than 50% of the area of the image object. An invading object is problematic to delineation of the corresponding AIR of the scene object because its main part does not fall within the corresponding AIR of the scene object. Thus, the merging region of good and expanding objects represents the DIR of the scene object (shown as the red-bordered polygon in Figure 1). This area, together with the position of the DIR of the scene object, determines segmentation quality. The key to segmentation quality assessment is to represent the area discrepancy and position discrepancy between the DIR and the corresponding AIR of the scene object.

We developed two metrics to represent the area discrepancy between the DIR and the corresponding AIR of a scene object – omission error (OE) and commission error (CE). OE refers to the area ratio of the overlapped regions between the invading objects and the corresponding AIR of the scene object to the corresponding AIR of the scene object, as shown in Equation (1). CE refers to the area ratio of the expanding objects minus the overlapped regions between the expanding objects and the corresponding AIR of the scene object to the corresponding AIR of the scene object, as shown in Equation (2).

$$\mathrm{OE} = \frac{\sum_{j=1}^{n} \{A_i(j) \cap A_r\}}{A_r} \times 100\%, \tag{1}$$

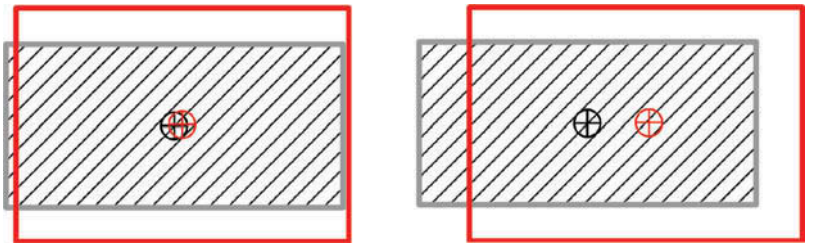
$$\mathrm{CE} = \frac{\sum_{k=1}^{m} \{A_e(k) - (A_e(k) \cap A_r)\}}{A_r} \times 100\%, \tag{2}$$

where n represents the number of invading objects, $A_i(j)$ is the area of the jth invading object, A_r is the area of the AIR of the scene object, m represents the number of expanding objects, and $A_e(k)$ is the area of the kth expanding object.

OE denotes the ratio of the omission areas (which are included in the neighbour image objects) to the area of the corresponding AIR of the scene object. CE denotes the ratio of the commission areas (which really belong to the neighbour image objects) to the area of the corresponding AIR of the scene object. The lower the values of OE and CE, the lower the area discrepancy between the DIR and the corresponding AIR of the scene object, and the better the segmentation quality. When the scene object is well segmented, the values of OE and CE are both close to 0.

Any changes in the OE and CE values indicate changes in the quality of the image segmentation, and can be presented in a two-dimensional space as (OE, CE). The distance between (OE, CE) and the origin represents the quality of the segmentation result. Based on the OE and CE metrics, we propose the area discrepancy index (ADI), as shown in Equation (3), to represent the area discrepancy between the DIR and the corresponding AIR of the scene object:

$$\mathrm{ADI} = \sqrt{\mathrm{OE}^2 + \mathrm{CE}^2}. \tag{3}$$



(a) (b)

Figure 2. The different segmentation results for the same scene object: (*a*) small position discrepancy between the DIR and the corresponding AIR of the scene object; (*b*) large position discrepancy between the DIR and the corresponding AIR of the scene object. The red-bordered polygon denotes the DIR of the scene object; the slashed polygon denotes the AIR of the scene object; the black crossed circle denotes the centroid of the AIR of the scene object; and the red crossed circle denotes the centroid of the DIR of the scene object.

When the segmentation operation is performed on an image with suitable segmentation parameter combinations, a smaller value of ADI is preferred, and when the scene object is well segmented, the value of ADI is close to 0.

The segmentation quality is also reflected by the position discrepancy (Molenaar 1998; Ragia and Winter 2000; Zhan et al. 2005; Möller, Lymburner, and Volk 2007; Ke, Quackenbush, and Im 2010; Pu and Landry 2012) between the DIR and the corresponding AIR of the scene object. As shown in Figure 2, the segmentation result of Figure 2(*a*) is superior to that of Figure 2(*b*) because the DIR in Figure 2(*a*) is closer to the centre of the corresponding AIR than that in Figure 2(*b*). We propose the position discrepancy index (PDI) to represent the position discrepancy between the DIR and the corresponding AIR of the scene object (Equation (4)):

$$\mathrm{PDI} = \frac{1}{N+M}\left(\sum_{k=1}^{N}\sqrt{(X(k)-X_r)^2+(Y(k)-Y_r)^2} + \sum_{l=1}^{M}\sqrt{(X(l)-X_r)^2+(Y(l)-Y_r)^2}\right), \tag{4}$$

where N represents the number of good objects, M represents the number of expanding objects, $X(k)$ and $Y(k)$ are the coordinates of the centroids of the kth good object, $X(l)$ and $Y(l)$ are the coordinates of the centroids of the lth expanding object, and X_r and Y_r are the coordinates of the centroid of the AIR of the scene object.

From Equation (4), note that the PDI is the average value of all the distances. As the DIR of the scene object is really composed of the good and expanding objects, in Equation (4) we compute the distances between the centroids of the good objects and the centroid of the corresponding AIR of the scene object, and the distances between the centroids of the expanding objects and the centroid of the corresponding AIR of the scene object.

For each segmentation result, the segmentation quality can be assessed based on the ADI and PDI metrics. The assessment procedure is shown in Figure 3.

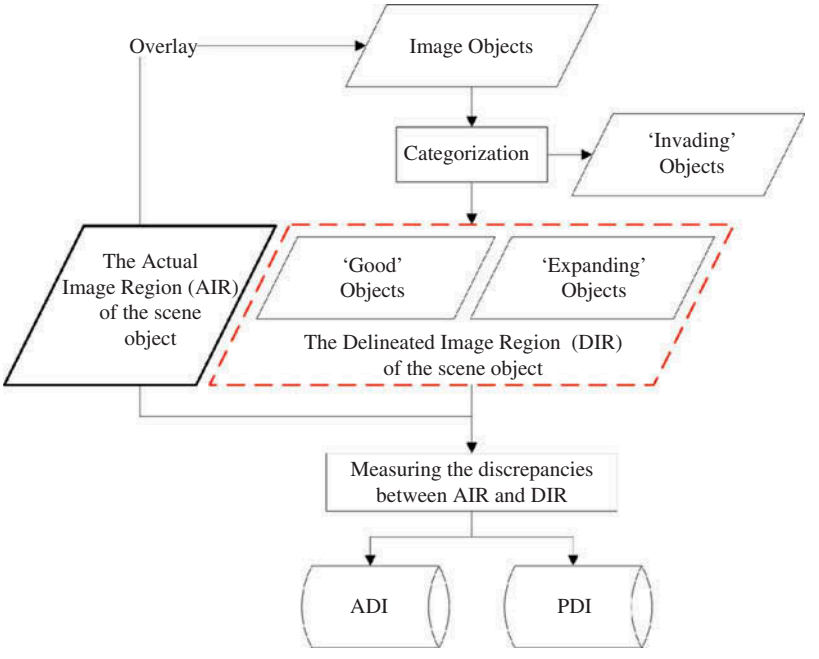


Figure 3. Assessment procedure for image segmentation quality.

2.2. *Assessment of a single-scene object*

As shown in Figure 3, assessment of the image segmentation quality of a single-scene object is straightforward. For each segmentation result of a single-scene object, the ADI and PDI values can be obtained as the segmentation quality.

2.3. *Assessment of a whole image*

Because scene objects of different size and different cover type in the image present different segmentation quality for each segmentation operation, the ADI and PDI metrics cannot be directly applied to the assessment of the segmentation quality of the whole image. Therefore, we define the $OE_{overall}$ (Equation (5)) and $CE_{overall}$ (Equation (6)) to represent the overall area discrepancy between the DIRs and the corresponding AIRs of all the scene objects of the whole image:

$$OE_{overall} = \frac{\sum_{i=1}^{n} (OE(i) * A_r(i))}{\sum_{i=1}^{n} A_r(i)} \times 100\%, \quad (5)$$

$$CE_{overall} = \frac{\sum_{i=1}^{n} (CE(i) * A_r(i))}{\sum_{i=1}^{n} A_r(i)} \times 100\%, \quad (6)$$

where n represents the number of selected scene objects of the whole image, $\mathrm{OE}(i)$ is the OE value of the ith scene object, $A_r(i)$ is the area of the AIR of the ith scene object, and $\mathrm{CE}(i)$ is the CE value of the ith scene object.

Based on the $\mathrm{OE}_{\mathrm{overall}}$ and $\mathrm{CE}_{\mathrm{overall}}$ metrics, we propose the $\mathrm{ADI}_{\mathrm{overall}}$ metric, as shown in Equation (7), to represent the overall area discrepancy between the DIRs and the corresponding AIRs of all the scene objects of the whole image:

$$\mathrm{ADI}_{\mathrm{overall}} = \sqrt{\mathrm{OE}_{\mathrm{overall}}{}^2 + \mathrm{CE}_{\mathrm{overall}}{}^2}. \tag{7}$$

We also propose the $\mathrm{PDI}_{\mathrm{overall}}$ metric to represent the overall position discrepancy between the DIRs and the corresponding AIRs of all the scene objects of the whole image, as shown in Equation (8):

$$\mathrm{PDI}_{\mathrm{overall}} = \frac{1}{n}\sum_{i=1}^{n} \mathrm{PDI}(i), \tag{8}$$

where n represents the number of the selected scene objects and $\mathrm{PDI}(i)$ is the PDI value of the ith scene object.

3. Experiments

3.1. Experimental procedures

High-spatial resolution remote-sensing images are segmented by using any type of segmentation algorithms and parameter combinations, and the obtained segmentation results can be assessed based on the ADI and PDI metrics (Figure 4). In order to screen out the optimal segmentation result, we must take a compromise between the ADI and PDI. The ideal optimal segmentation result can be obtained when both the ADI and the PDI are simultaneously at their minimum value; however, such a case rarely occurs. Area discrepancy information is more important than position discrepancy information for assessment of the segmentation result (Ji 2012). Thus, we can first select the segmentation results with relatively low ADI values as the better segmentation results because they contribute to minimization of the classification error caused by the image segmentation operation. Due to differences in spectral characteristics, land-cover types, and image quality, there is no fixed threshold of ADI that can be used as a reference standard and applied to all remote-sensing images for selecting better segmentation results. Through repeated testing, an empirical interval (ADI ≥ min(all(ADI)) and ADI ≤ 1.1 * min(all (ADI))) is used as the standard for selecting optimal segmentation results from the total. If the ADI value of the segmentation result falls within the interval (ADI ≥ min(all(ADI)) and ADI ≤ 1.1 * min(all(ADI))), the corresponding segmentation result would be classified as a better segmentation result. Then, from all of the better segmentation results selected, we can take the segmentation result with the lowest PDI value as the optimal segmentation result.

3.2. Data and segmentation

The image used for this study is a three-band (RGB) aerial image from Florida, USA (Figure 5), with a spatial resolution of 0.32 m. The major underlaying surface types in this image are building, frontage ground, road, grassland, forest, and water.

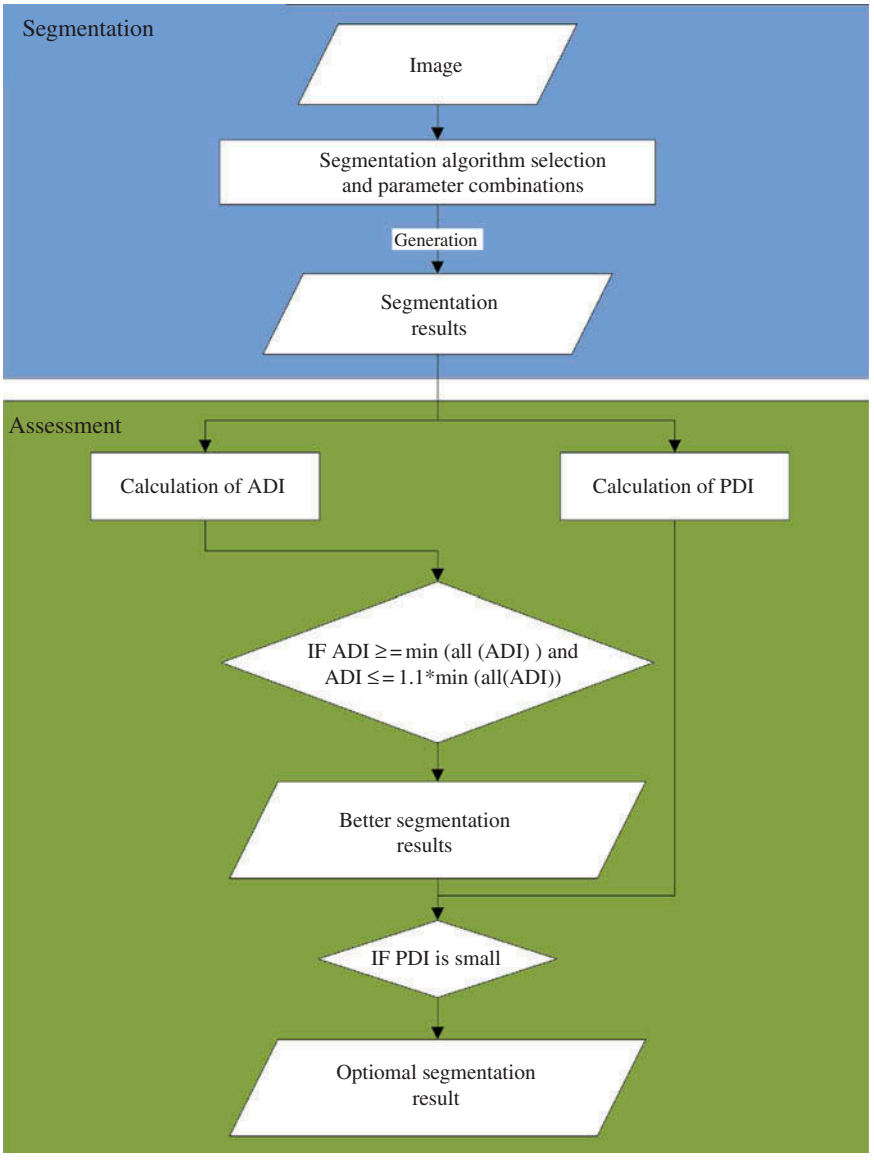


Figure 4. Experimental procedures.

We conducted the image segmentations by using Definiens 8.64 (formerly eCognition) (http://www.definiens.com). Definiens 8.64 uses multi-resolution segmentation based on the region merging algorithm (Benz et al. 2004). This algorithm uses three parameters that allow users to tune the segmentation results. The first parameter is scale parameter H. Higher H values result in larger image objects, and *vice versa*; the other two parameters are shape and compactness. The shape criterion determines to what degree shape influences segmentation compared with colour: for example, a shape weighting 0.1 results in a colour weighting 0.9. In the same way, the value for compactness gives it a relative weighting against smoothness.

Figure 5. The aerial image used for segmentation and quality assessment. The yellow-bordered polygons are the AIRs of the buildings; the purple-bordered polygons are the AIRs of the forests; the blue numbers are the codes for buildings; and the purple numbers are the codes for forest.

Different combinations of scale, shape, and compactness values produce different segmentation results. For an image, there is an optimal parameter combination by which image objects are very close to their corresponding AIRs of scene objects. However, it is a trial-and-error procedure to obtain an optimal parameter combination (Hay et al. 2003; Stein and de Beurs 2005; Möller, Lymburner, and Volk 2007). In this study, 175 parameter combinations were generated for scale, shape, and compactness according to {60, 70, 80, 90, 100, 110, 120} × {0.1, 0.3, 0.5, 0.7, 0.9} × {0.1, 0.3, 0.5, 0.7, 0.9}. Thirty-one buildings and seven forests in the image were randomly selected and vectorized as the reference segments (Figure 5).

3.3. Results

3.3.1. Segmentation quality assessment of a single-scene object

We selected building 16 as an example of single-scene object. Seven selected segmentation results with different segmentation scales are presented in Figure 6. For building 16, the ADI and PDI metrics were calculated for each segmentation result. The assessment results of 175 segmentation results are shown in Table 1, where the ADI value of 4.70% is the lowest. Thus, the four corresponding segmentation results with an ADI value of 4.70% are superior to the other segmentation results for building 16. Among the PDI values of the four segmentation results for building 16, the lowest PDI value is 0.34. According to the rule described in Section 2, the segmentation result with ADI of 4.70% and PDI of 0.34 is regarded as being optimal for building 16. Changes of ADI and PDI according to scale parameter shown in Figure 7 illustrate that ADIs under the condition of scale ≤ 90 are lower than those under the condition of scale >90. This means that the segmentation results under the condition of scale ≤ 90 are superior to those under the condition of scale

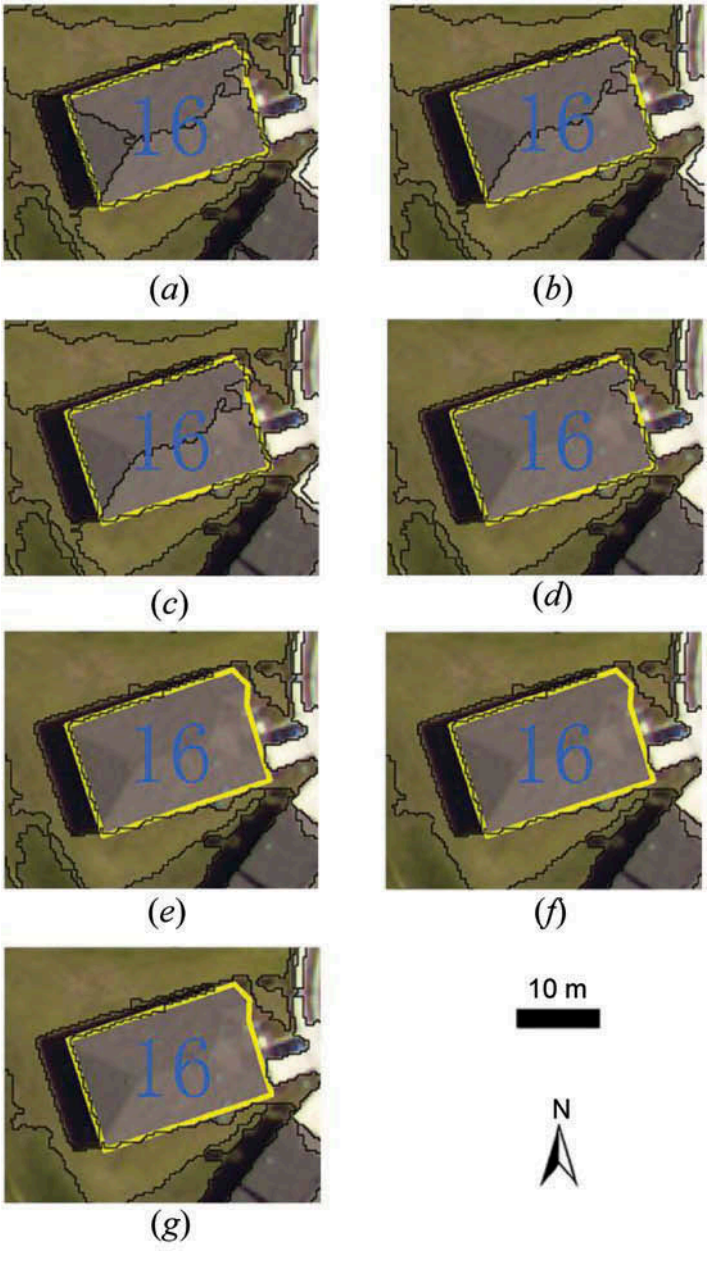


Figure 6. Illustrations of the segmentation results for building 16 at scales (*a*) 60, (*b*) 70, (*c*) 80, (*d*) 90, (*e*) 100, (*f*) 110, and (*g*) 120. Shape and compactness parameters are both 0.1. The yellow-bordered polygons are the AIRs of the buildings and the black-bordered polygons are the image objects.

>90 (Figures 6(*a*)–(*g*)). The PDI values are different at scale 60, 70, 80, and 90, while the corresponding ADI values are identical. By taking the combined ADI and PDI metrics into consideration, the scale 90 is regarded as the optimal segmentation scale for the segmentation of building 16.

Table 1. Assessment results of the segmentation results for building 16 from 175 parameter combinations using ADI and PDI, AFI, OL, and I.

Segmentation parameter combination									
Shape	Compactness	Scale	OE (%)	CE (%)	ADI (%)	PDI (m)	AFI	OL	I
0.1	0.1	60	4.02	2.43	4.70	5.54	0.561	1.0	0.25
0.1	0.1	70	4.02	2.43	4.70	3.66	0.486	1.0	0.33
0.1	0.1	80	4.02	2.43	4.70	3.66	0.486	1.0	0.33
0.1	0.1	90	4.02	2.43	4.70	0.34	0.016	1.0	0.50
0.1	0.1	100	1.04	29.68	29.70	2.69	−0.286	0.0	1.0
0.1	0.1	110	1.04	29.68	29.70	2.69	−0.286	0.0	1.0
0.1	0.1	120	1.04	29.68	29.70	2.69	−0.286	0.0	1.0
0.1	0.3	60	3.14	3.30	5.86	5.55	0.539	1.0	0.25
0.1	0.3	70	3.14	5.35	6.20	2.95	0.441	1.0	0.33
0.1	0.3	80	3.14	5.35	6.20	2.95	0.441	1.0	0.33
0.1	0.3	90	3.14	5.35	6.20	0.49	−0.022	1.0	0.50
0.1	0.3	100	3.14	5.35	6.20	0.49	−0.022	0.0	1.0
0.1	0.3	110	3.14	5.35	6.20	0.49	−0.022	0.0	1.0
0.1	0.3	120	3.14	5.35	6.20	0.49	−0.022	0.0	1.0
…	…	…	…	…	…	…	…	…	…
0.9	0.9	60	23.60	2.91	23.78	1.46	0.207	0.0	0.25
0.9	0.9	70	100.00[a]	0.00	100.00	NaN[b]	−0.204	0.0	1.0
0.9	0.9	80	100.00[a]	0.00	100.00	NaN[b]	−0.339	0.0	1.0
0.9	0.9	90	100.00[a]	0.00	100.00	NaN[b]	−0.339	0.0	1.0
0.9	0.9	100	100.00[a]	0.00	100.00	NaN[b]	−0.339	0.0	1.0
0.9	0.9	110	100.00[a]	0.00	100.00	NaN[b]	−0.339	0.0	1.0
0.9	0.9	120	100.00[a]	0.00	100.00	NaN[b]	−0.418	0.0	1.0

Notes: [a]OE 100% demonstrates that the scene object is completely partitioned into the adjacent image objects. [b]PDI NaN demonstrates that the value of the PDI metric in this case cannot be calculated because the DIR of the scene object does not exist.

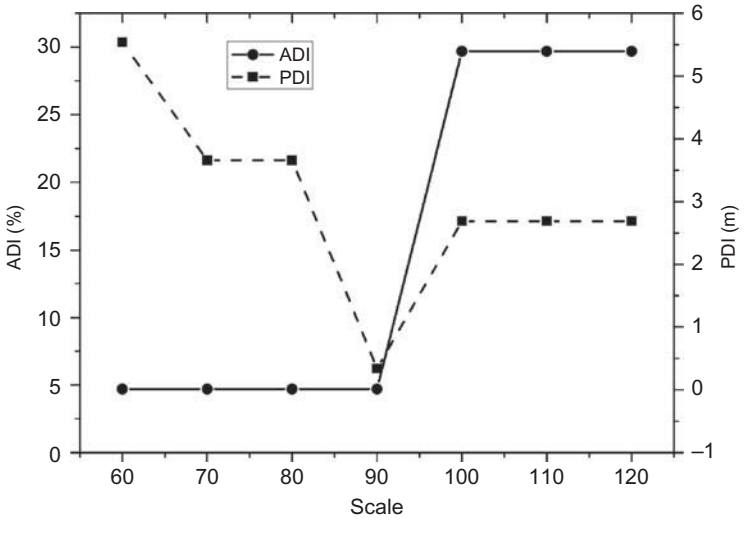


Figure 7. Diagram of the assessment result of seven selected segmentation results for building 16 shown in Figure 6 using ADI and PDI.

We also used the AFI, OL, and I metrics to assess the segmentation quality of building 16 from 175 parameter combinations. AFI (area fitness index) is defined in Equation (9) (Lucieer and Stein 2002). OL (offspring loyalty) and I (interference) are defined in Equation (10) and Equation (11), respectively (Schöpfer and Lang 2006; Schöpfer, Lang, and Albrecht 2008; Lang, Schöpfer, and Langanke 2009; Albrecht, Lang, and Hölbling 2010). The assessment results are also shown in Table 1. For each segmentation parameter combination, the assessment results are different from that obtained using the ADI and PDI metrics:

$$\mathrm{AFI} = \frac{A_r - A_{\text{largest image object}}}{A_r}, \tag{9}$$

where $A_{\text{largest image object}}$ is the area of the largest image object intersecting with the scene object. For a perfect segmentation result, AFI = 0.0. A scene object is over-segmented when AFI > 0.0. A scene object is under-segmented when AFI < 0.0:

$$\mathrm{OL} = \frac{n_{\mathrm{g}}}{n_{\mathrm{g}} + n_{\mathrm{e}}}, \tag{10}$$

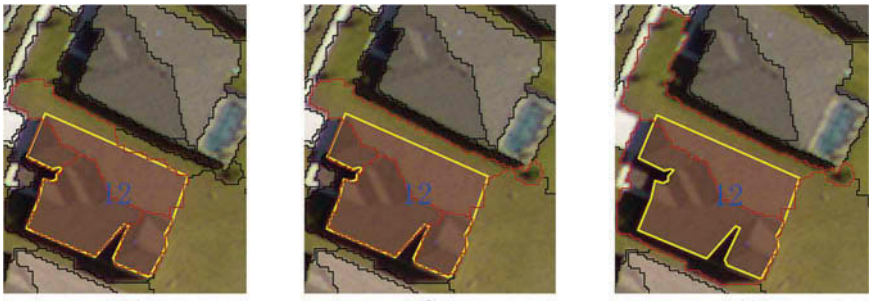
$$\mathrm{I} = \frac{n_{\mathrm{i}}}{n_{\mathrm{all}}}, \tag{11}$$

where n_{g} is the number of good objects, n_{e} is the number of expanding objects, n_{i} is the number of invading objects, and n_{all} is the number of all intersecting objects. For a perfect segmentation result, OL = 1.0 and I = 0.0.

The different optimal segmentation results obtained using ADI and PDI, AFI, OL, and I are shown in Figures 8–11. In Figures 8, 9, and 11, the OL values of results (a) and (b) are the same, as are the I values. Thus, results (a) and (b) should both be regarded as the optimal segmentation results according to the object-fate analysis method. However, compared with result (b), result (a) has lower ADI values. This means that result (a) has lower CE or OE than result (b) and therefore, result (a) in Figures 8, 9, and 11 is the real optimal segmentation result for subsequent classification. In Figures 8–11, compared with result (c), result (a) has lower ADI values. This means that result (a) has lower CE or OE than result (c) and therefore, in Figures 8–11 result (a) is more suitable for subsequent classification than result (c).

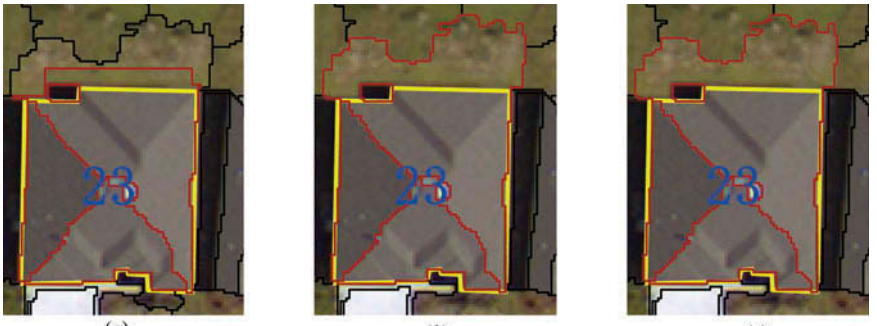
3.3.2. *Segmentation quality assessment of the whole image*

To simplify the discussion, we selected only building objects in the image as the study objects. To carry out an overall assessment of the segmentation result for all buildings, selecting all building objects is not practical in application because there could be many buildings in the image and acquiring their reference segments is time consuming. Thus, it is necessary to select building samples using the random sampling approach. The sampling proportion of buildings is the key factor that affects the overall assessment results of the segmentation quality of all buildings. In the research literature, there is no appropriate sampling scheme which can be directly used to determine the sampling proportion of research objects. For this reason, we propose a new sampling scheme (Figure 12) for determining the sampling proportion of buildings. This sampling scheme is also valid for the sample selection of other cover types. The experimental steps of this sampling scheme are explained as follows.



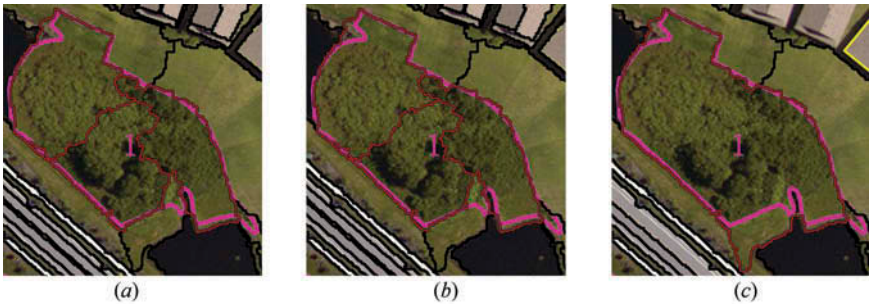
(a) (b) (c)

Figure 8. Optimal segmentation result for building 12 using (*a*) ADI and PDI in our method: ADI = 4.18%, PDI = 1.19, OL = 0.5, I = 0, AFI = 0.52; (*b*) OL and I in the object-fate analysis method: ADI = 15.27%, PDI = 4.23, OL = 0.5, I = 0, AFI = 0.47; and (*c*) AFI: ADI = 72.14%, PDI = 6.28, OL = 0, I = 0.5, AFI = 0.08. The yellow-bordered polygons are the AIRs of building 12; the black- and red-bordered polygons are the image objects.



(a) (b) (c)

Figure 9. Optimal segmentation result for building 23 using (*a*) ADI and PDI in our method: ADI = 1.34%, PDI = 0.59, OL = 0.75, I = 0, AFI = 0.46; (*b*) OL and I in the object-fate analysis method: ADI = 18.52%, PDI = 6.44, OL = 0.75, I = 0, AFI = 0.19; and (*c*) AFI: ADI = 18.52%, PDI = 6.44, OL = 0.75, I = 0, AFI = 0.19. The yellow-bordered polygons are the AIRs of building 23; the black- and red-bordered polygons are the image objects.



(a) (b) (c)

Figure 10. Optimal segmentation result for forest 1 using (*a*) ADI and PDI in our method: ADI = 2.05%, PDI = 2.41, OL = 1, I = 0.25, AFI = 0.58; (*b*) OL and I in the object-fate analysis method: ADI = 2.05%, PDI = 2.41, OL = 1, I = 0.25, AFI = 0.58; and (*c*) AFI: ADI = 5.69%, PDI = 5.38, OL = 0, I = 0, AFI = −0.14. The purple-bordered polygons are the AIRs of forest 1; the black- and red-bordered polygons are the image objects.

Figure 11. Optimal segmentation result for forest 6 using (*a*) ADI and PDI in our method: ADI = 5.15%, PDI = 1.27, OL = 0.67, I = 0, AFI = 0.44; (*b*) OL and I in the object-fate analysis method: ADI = 19.36%, PDI = 11.48, OL = 0.67, I = 0, AFI = 0.23; and (*c*) AFI: ADI = 18.74%, PDI = 2.57, OL = 0, I = 0, AFI = −0.16. The purple-bordered polygons are the AIRs of forest 6; the black- and red-bordered polygons are the image objects.

Step 1: without sampling, calculate the $ADI_{overall}$ and $PDI_{overall}$ metrics of all the building research objects. The obtained $ADI_{overall}$ and $PDI_{overall}$ values are regarded as the real values of $ADI_{overall}$ and $PDI_{overall}$ metrics of all buildings. Step 2: with the random sampling approach, building samples are selected from building research objects in various sampling proportions such as 5%, 10%, 15%, 20%, 25%, and 30%. For each sample operation, the corresponding $ADI_{overall}$ and $PDI_{overall}$ metrics are calculated. In this case, the obtained $ADI_{overall}$ and $PDI_{overall}$ values are the sampling values of $ADI_{overall}$ and $PDI_{overall}$ metrics of all buildings. Step 3: for each sample operation, calculate the deviations between the sampling values and real values of $ADI_{overall}$ and $PDI_{overall}$ metrics. The deviation percentages are obtained by dividing the deviations by the corresponding real values of $ADI_{overall}$ and $PDI_{overall}$ metrics. Step 4: repeat the sampling operation of each sampling proportion more than 50 times. In other words, steps 2 and 3 are carried out repeatedly more than 50 times. As a result, the histograms of the deviation percentages for each sampling proportion are obtained (Figure 13). The shapes of the histograms are similar to Gaussian distribution shapes. Step 5: merge the deviation percentage intervals of the same absolute values into positive deviation percentage intervals. Recount the sampling times falling within each positive deviation percentage interval for each sampling proportion. Then, calculate the cumulative percentages of sampling times for each sampling proportion. The calculation results are shown in Table 2. As shown in Table 2, if we want to constrain both the deviation percentage of $ADI_{overall}$ and that of $PDI_{overall}$ to less than 4%, the 20% building sample is desirable. Therefore, we selected the 20% building sample (i.e. 31 buildings in Figure 5) randomly and vectorized these as reference segments.

According to our method, we use the $ADI_{overall}$ and $PDI_{overall}$ metrics to assess the segmentation results for the buildings. The overall assessment result of the segmentation results of the 31 buildings is shown in Table 3, from which we can see that, with the parameter combination of shape 0.3, compactness 0.7, and scale 60, $ADI_{overall}$ presents the lowest value of 5.77% and $PDI_{overall}$ presents the lowest value of 3.81. Thus, the parameter combination of shape 0.3, compactness 0.7, and scale 60 is the optimal parameter combination that leads to the optimal segmentation result, as shown in Figure 14.

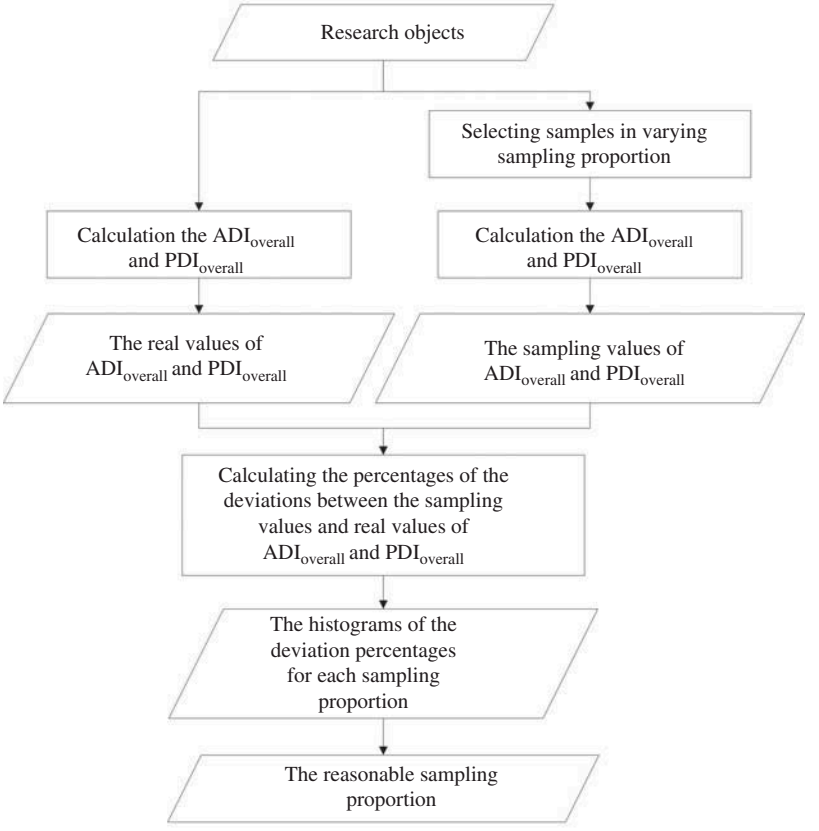


Figure 12. Sampling scheme for determining the sampling proportion of research objects.

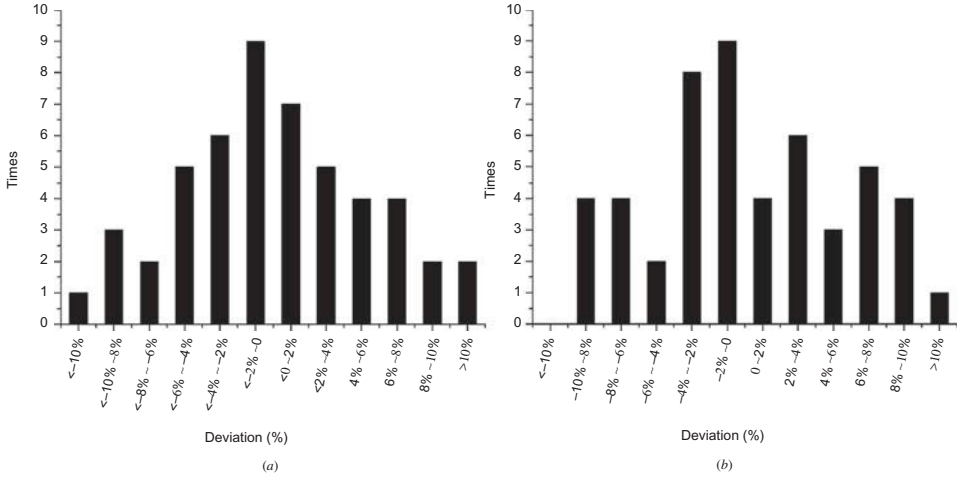


Figure 13. Histograms of the deviation percentages of 50-times building sampling in 5% sampling proportion: (*a*) ADI$_{overall}$, (*b*) PDI$_{overall}$.

Table 2. Calculation results for cumulative percentages of sampling times for each sampling proportion of all buildings.

Metrics	Deviation percentage interval	5% sampling proportion		10% sampling proportion		15% sampling proportion		20% sampling proportion		25% sampling proportion		30% sampling proportion	
		^{1}T	^{2}CP (%)	^{1}T	^{2}CP (%)	^{1}T	^{2}CP (%)	^{1}T	^{2}CP (%)	^{1}T	^{2}CP (%)	^{1}T	^{2}CP (%)
$ADI_{overall}$	<2%	16	32	23	46	29	58	37	74	39	78	46	92
	2–4%	11	54	13	72	9	76	13	100	11	100	4	100
	4–6%	9	72	10	92	11	98	0	100	0	100	0	100
	6–8%	6	84	2	96	1	100	0	100	0	100	0	100
	8–10%	5	94	2	100	0	100	0	100	0	100	0	100
	>10%	3	100	0	100	0	100	0	100	0	100	0	100
$PDI_{overall}$	<2%	13	26	17	34	27	54	33	66	41	82	44	88
	2–4%	14	54	14	62	8	70	17	100	9	100	6	100
	4–6%	5	64	6	74	9	88	0	100	0	100	0	100
	6–8%	9	82	8	90	3	94	0	100	0	100	0	100
	8–10%	8	98	5	100	3	100	0	100	0	100	0	100
	>10%	1	100	0	100	0	100	0	100	0	100	0	100

Note: ^{1}T refers to times, ^{2}CP refers to cumulative percentages.

Table 3. Overall assessment results for segmentation results of the 31 buildings from 175 parameter combinations using $ADI_{overall}$ and $PDI_{overall}$.

Segmentation parameter combination						
Shape	Compactness	Scale	$OE_{overall}$ (%)	$CE_{overall}$ (%)	$ADI_{overall}$ (%)	$PDI_{overall}$ (m)
0.1	0.1	60	6.34	5.25	8.22	5.06
0.1	0.1	70	7.75	4.99	9.22	4.83
0.1	0.1	80	8.09	5.47	9.77	4.73
0.1	0.1	90	8.99	6.83	11.30	4.71
0.1	0.1	100	9.02	8.27	12.23	5.08
0.1	0.1	110	10.25	8.84	13.54	4.77
0.1	0.1	120	11.46	8.68	14.37	4.71
0.1	0.3	60	5.22	4.55	6.93	5.48
0.1	0.3	70	6.01	4.45	7.48	4.89
…	…	…	…	…	…	…
0.3	0.7	60	4.11	4.05	5.77	3.81
0.3	0.7	70	4.75	4.28	6.39	4.06
0.3	0.7	80	5.43	4.26	6.90	4.17
…	…	…	…	…	…	…
0.9	0.9	60	22.41	12.71	25.76	5.25
0.9	0.9	70	24.02	16.75	29.28	5.41
0.9	0.9	80	31.88	14.78	35.14	5.68
0.9	0.9	90	36.01	15.14	39.07	6.88
0.9	0.9	100	40.85	12.74	42.80	7.65
0.9	0.9	110	45.56	10.32	46.71	8.23
0.9	0.9	120	48.05	9.00	48.89	9.33

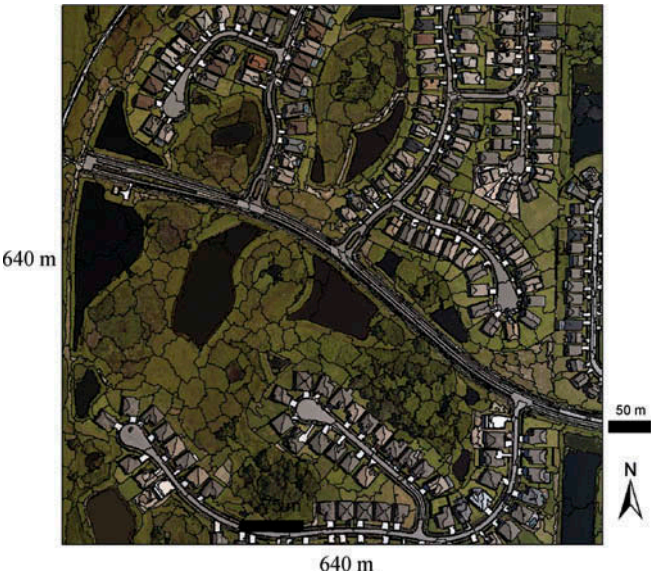


Figure 14. Overall optimal segmentation result for buildings according to shape 0.3, compactness 0.7, and scale 60.

4. Discussion

Use of boundary coincidence-based methods is limited, as shown by the assessment of the segmentation results for building 16. For the three segmentation results for building 16 in Figures 6(*a*)–(*c*), it is difficult to identify a boundary of an image object that corresponds to the boundary of the AIR of building 16 for distance calculation. For the four segmentation results of building 16 in Figures 6(*d*)–(*g*), although it is relatively easy to identify the boundary of the image object that corresponds to the boundary of the AIR of building 16 for distance calculation, both the means and number of selecting points on the identified boundary for the distance calculation mainly rely on human subjective judgement. When selection for the points is unreasonable, the assessment result is far from ideal.

The method of Lucieer and Stein (2002) and the object-fate analysis method cannot always express the quality of a segmentation result objectively. In Table 1, when shape and compactness parameters are both 0.1 and scale parameter changes from 60 to 120, AFI ranges from 0.561 to −0.286. Because AFI 0.016 is closest to 0.0, the segmentation parameter combination of shape 0.1, compactness 0.1, and scale 90 is optimal for building 16 according to Lucieer and Stein (2002) and the corresponding segmentation result is optimal (Figure 6(*d*)). This conclusion is consistent with that obtained by our method. However, for the segmentation results of building 16 in Figures 6(*a*)–(*c*), AFI is 0.561, 0.486, and 0.486, respectively. This means that the segmentation quality of building 16 in Figures 6(*a*)–(*c*) is very poor. In fact, area discrepancy between the DIR and AIR of building 16 in Figures 6(*a*)–(*c*) is the same as that in Figure 6(*d*). From the perspective of the delineation of building 16, the segmentation results for this in Figures 6(*a*)–(*c*) are as good as those in Figure 6(*d*). The reason that the segmentation result for building 16 in Figure 6(*d*) is preferred is that its position discrepancy is lower than that of the three segmentation results in Figures 6(*a*)–(*c*). Although users can obtain the optimal segmentation parameter according to the balance between the OL and I values, segmentation quality is not expressed by the OL and I values objectively. For example, we can see that the I value of the segmentation result for building 16 in Figure 6(*a*) is 0.25, whereas the I value of the segmentation result in Figure 6(*b*) is 0.33. According to the rules of the object-fate analysis method, with a lower I value the segmentation quality is better. If this is true, it can be concluded that the segmentation result for building 16 in Figure 6(*a*) is better than that in Figure 6(*b*). However, in reality, this is not the case. This situation occurs because the object-fate analysis method uses the numbers of different types of image object as the standard for assessing the segmentation quality without considering their area.

In our method, we take all the image objects intersecting with the AIR of building 16 into consideration to measure both area and position discrepancy between the DIR and the AIR. We can see that the ADIs of the segmentation results in Figure 6(*a*)–(*d*) are all 4.70%, whereas the PDIs are different. The quality of a segmentation result depends not only on ADI but also PDI. For a smaller PDI value, the segmentation quality is better. Thus, the conclusion is that the quality of the segmentation result for building 16 in Figure 6(*d*) is better than that in Figures 6(*a*)–(*c*), and the quality of the segmentation result for building 16 in Figures 6(*b*) and (*c*) is better than that in Figure 6(*a*). This matches the real situation very well.

Different assessment methods may derive different optimal segmentation results for the same scene object, as shown in Figures 8–11. Among all optimal segmentation results, that derived from our method is most suitable. In the method of Lucieer and Stein (2002),

the authors tried to find an image object that can represent the corresponding scene object, and the largest image object intersecting with the scene object was chosen to fulfil this role. If the area of this largest image object is close to the area of the corresponding scene object, the segmentation result is regarded as the optimal segmentation result. However, in some cases, the largest image object has little spatial overlap with the corresponding scene object even when the areas are fairly close (Figures 8(*c*) and 9(*c*)). In this case the largest image object cannot represent the corresponding scene object. Therefore, the so-called optimal segmentation result in the method of Lucieer and Stein (2002) is likely to be unsuitable. In the object-fate analysis method, the authors believe that the optimal segmentation result depends on the numbers of expanding and invading objects: when the numbers of both are low, the segmentation result is regarded as optimal. However, two segmentation results are found when the numbers of both are equal, whereas the areas of expanding objects are not the same (Figures 8(*a*) and (*b*), 9(*a*) and (*b*), 11(*a*) and (*b*)). Clearly, the two segmentation results have different effects on subsequent classification operations. Therefore, the so-called optimal segmentation result from the object-fate analysis method is likely to be unsuitable. In our method, both area and position discrepancy between the DIR and the corresponding AIR of a scene object are regarded as the standards for assessing segmentation quality. When both discrepancies between the DIR and corresponding AIR of a scene object are low, the extent of the scene object can be well delineated by the image objects. Even when the optimal segmentation results from our method have a tendency towards over-segmentation (Figures 8–11, result (a)), they are still beneficial to the subsequent classification operation in that over-segmentation can be easily handled through appropriate operations such as merging.

From Table 1, we can see that although the segmentation parameter combination (shape 0.1, compactness 0.1, scale 90) is optimal for building 16, this may not be suitable for other scene objects. Generally, different scene objects of the same type in an image may have different spectral characteristics, shapes, and sizes. As a result, the optimal segmentation parameters for these scene objects are different. According to the experimental results, we can see that the optimal segmentation parameter combination (shape 0.3, compactness 0.7, scale 60) for buildings is different from that (shape 0.1, compactness 0.1, scale 90) for a single building (building 16). This result is not contradictory, because assessment of the segmentation quality of a single-scene object only considers the segmentation result of that object, whereas overall assessment of the segmentation quality of the buildings takes the segmentation results of all the buildings into consideration.

As shown in Figure 14, some buildings are poorly segmented and some are even partitioned into many image objects. This conforms to the actual situation. The overall optimal segmentation result does not mean that all the buildings in the image are well segmented in this case, because of the differences in spectral characteristic, shapes, and sizes. Also in Figure 14, other land-cover types (e.g. water surfaces, roads) are not always well segmented. Marçal and Rodrigues (2009) pointed out that the segmentation result obviously depends on land-cover type, object size, and shape. By taking object samples for assessment, the optimal segmentation result for one land-cover type is statistically optimal for that land-cover type. It is possible that the optimal segmentation parameter combination for one land-cover type is not suitable for others. If the goal of segmentation assessment is to obtain the optimal segmentation result for a specific land-cover type, one need only apply our method to that specific type of object. However, if the goal is to obtain the optimal segmentation results of all land-cover types, one can apply our method to each type of object individually. If the goal of the segmentation assessment is to obtain

the global optimal segmentation result for the whole image, then one should apply our method to the object samples that cover all land-cover types.

By using our method, the cost in computation time is minimal. The time cost is mainly in regard to manually obtaining the AIRs of object samples, and it depends on sample size. Providing the AIRs of object samples are obtained, the following assessment procedure requires minimal human intervention and can be automatically conducted by utilizing computer programs or third-party software.

5. Conclusions

In this article, we propose a novel method for assessing the segmentation quality of high-spatial-resolution remote-sensing images, which is a type of supervised method. The method classifies image objects as 'good', 'expanding', or 'invading' objects in terms of the spatial relationships between image objects and the corresponding AIR of the scene object. As the main parts of both good and expanding objects fall within the AIR of the scene object, these represent the DIR (delineated image region) of the scene object. Discrepancies between the DIR and the corresponding AIR of the scene object reflect the quality of the segmentation result. We propose that ADI and PDI represent these discrepancies. One can use ADI and PDI to assess the segmentation quality of various segmentation results. Our method can be used to assess not only the segmentation quality of a single-scene object, but also the segmentation quality of a whole image. The assessment procedure for the segmentation result can be conducted by utilizing computer programs or third-party software, which minimizes human intervention. In comparison with most other methods used for assessing segmentation quality, our method assesses the segmentation result more objectively and is more convenient in practice. The obtained optimal segmentation result can ensure maximal delineation of the extent of scene objects and can be beneficial to subsequent classification operations.

The effectiveness of our method for assessment of image segmentation quality has been proven by the reported experimental results. Providing the reference regions of scene objects are provided, the method can be applied to all scenarios to assess image segmentation quality. Moreover, our method is highly applicable to the assessment of segmentation results from other segmentation algorithms, although we used only the multi-resolution segmentation algorithm to obtain the image objects in this study.

Funding

This research was jointly supported by the National Natural Science Foundation of China [grant number 41271347]; the National Key Basic Research Programme of China (973 Programme) Project [grant number 2013CB733403]; the Prior Research Project of High Spatial Resolution Application Demonstration System of Land and Resources [grant number E0202/1112/0102]; the Science and Technology Key Research Project of Henan Province Department of Education [grant number 14A420004]; and the Doctoral Programme Fund of Henan Polytechnic University [grant number B2014-014].

References

Abeyta, A. M., and J. Franklin. 1998. "The Accuracy of Vegetation Stand Boundaries Derived from Image Segmentation in a Desert Environment." *Photogrammetric Engineering & Remote Sensing* 64 (1): 59–66.

Albrecht, F., S. Lang, and D. Hölbling. 2010. "Spatial Accuracy Assessment of Object Boundaries for Object-Based Image Analysis." Paper presented at the proceedings of geographic object-

based image analysis – the international archives of the photogrammetry, remote sensing and spatial information sciences, Ghent, Belgium, June 29–July 2.

Benz, U. C., P. Hofmann, G. Willhauck, I. Lingenfelder, and M. Heynen. 2004. "Multi-Resolution, Object-Oriented Fuzzy Analysis of Remote Sensing Data for Gis-Ready Information." *ISPRS Journal of Photogrammetry & Remote Sensing* 58: 239–258. doi:10.1016/j.isprsjprs.2003.10.002.

Blanchard, S. D., M. K. Jakubowski, and M. Kelly. 2011. "Object-Based Image Analysis of Downed Logs in Disturbed Forested Landscapes Using Lidar." *Remote Sensing* 3: 2420–2439. doi:10.3390/rs3112420.

Blaschke, T. 2010. "Object Based Image Analysis for Remote Sensing." *ISPRS Journal of Photogrammetry and Remote Sensing* 65: 2–16. doi:10.1016/j.isprsjprs.2009.06.004.

Blaschke, T., C. Burnett, and A. Pekkarinen. 2004. "Image Segmentation Methods for Object-Based Analysis and Classification." In *Remote Sensing Image Analysis: Including the Spatial Domain*, edited by S. M. de Jong and F. D. van der Meer, 211–236. Dordrecht: Kluwer Academic.

Blaschke, T., S. Lang, E. Lorup, J. Strobl, and P. Zeil. 2000. "Object-Oriented Image Processing in an Integrated Gis/Remote Sensing Environment and Perspectives for Environmental Applications." In *Environmental Information for Planning, Politics and the Public*, edited by A. Cremers and K. Greve, 555–570. Marburg: Metropolis Verlag.

Campbell, J. B. 2002. *Introduction to Remote Sensing*. 3rd ed. New York: The Guilford Press.

Carleer, A. P., O. Debeir, and E. Wolff. 2005. "Assessment of Very High Spatial Resolution Satellite Image Segmentations." *Photogrammetric Engineering & Remote Sensing* 71 (11): 1285–1294. doi:10.14358/PERS.71.11.1285.

Castillejo-González, I. L., F. López-Granados, A. García-Ferrer, J. M. Peña-Barragán, M. Jurado-Expósito, M. S. de la Orden, and M. González-Audicana. 2009. "Object- and Pixel-based Analysis for Mapping Crops and Their Agro-Environmental Associated Measures Using Quickbird Imagery." *Computers and Electronics in Agriculture* 68: 207–215. doi:10.1016/j.compag.2009.06.004.

Chabrier, S., B. Emile, C. Rosenberger, and H. Laurent. 2006. "Unsupervised Performance Evaluation of Image Segmentation." *EURASIP Journal on Applied Signal Processing* 2006: 1–12. doi:10.1155/ASP/2006/96306.

Clinton, N., A. Holt, J. Scarborough, L. Yan, and P. Gong. 2010. "Accuracy Assessment Measures for Object-based Image Segmentation Goodness." *Photogrammetric Engineering & Remote Sensing* 76 (3): 289–299. doi:10.14358/PERS.76.3.289.

Corcoran, P., A. Winstanley, and P. Mooney. 2010. "Segmentation Performance Evaluation for Object-based Remotely Sensed Image Analysis." *International Journal of Remote Sensing* 31 (3): 617–645. doi:10.1080/01431160902894475.

Delves, L. M., R. Wilkinson, C. J. Oliver, and R. G. White. 1992. "Comparing the Performance of SAR Image Segmentation Algorithms." *International Journal of Remote Sensing* 13 (11): 2121–2149. doi:10.1080/01431169208904257.

Desclée, B., P. Bogaert, and P. Defourny. 2006. "Forest Change Detection by Statistical Object-based Method." *Remote Sensing of Environment* 102: 1–11. doi:10.1016/j.rse.2006.01.013.

Dorren, L. K. A., B. Maier, and A. C. Seijmonsbergen. 2003. "Improved Landsat-Based Forest Mapping in Steep Mountainous Terrain Using Object-based Classification." *Forest Ecology and Management* 183: 31–46. doi:10.1016/S0378-1127(03)00113-0.

Espindola, G. M., G. Camara, I. A. Reis, L. S. Bins, and A. M. Monteiro. 2006. "Parameter Selection for Region-Growing Image Segmentation Algorithms Using Spatial Autocorrelation." *International Journal of Remote Sensing* 27 (14): 3035–3040. doi:10.1080/01431160600617194.

Hay, G. J., T. Blaschke, D. J. Marceau, and A. Bouchard. 2003. "A Comparison of Three Image-object Methods for the Multiscale Analysis of Landscape Structure." *ISPRS Journal of Photogrammetry & Remote Sensing* 57: 327–345. doi:10.1016/S0924-2716(02)00162-4.

Hay, G. J., and G. Castilla. 2006. "Object-based Image Analysis: Strengths, Weaknesses, Opportunities and Threats (SWOT)." Paper presented at the proceedings of the object-based image analysis (OBIA) – The international archives of the photogrammetry, remote sensing and spatial information sciences, Salzburg, July 4–5.

Hay, G. J., G. Castilla, M. A. Wulder, and J. R. Ruiz. 2005. "An Automated Object-Based Approach for the Multiscale Image Segmentation of Forest Scenes." *International Journal of Applied Earth Observation and Geoinformation* 7: 339–359. doi:10.1016/j.jag.2005.06.005.

Hölbling, D., P. Füreder, F. Antolini, F. Cigna, N. Casagli, and S. Lang. 2012. "A Semi-Automated Object-based Approach for Landslide Detection Validated by Persistent Scatterer Interferometry Measures and Landslide Inventories." *Remote Sensing* 4: 1310–1336. doi:10.3390/rs4051310.

Im, J., J. R. Jensen, and J. A. Tullis. 2008. "Object-Based Change Detection Using Correlation Image Analysis and Image Segmentation." *International Journal of Remote Sensing* 29 (2): 399–423. doi:10.1080/01431160601075582.

Ji, X. 2012. "Research on the Method of Accuracy Assessment of the Object-based Classification from Remotely Sensed Data." Master diss., Beijing Normal University.

Johnson, B., and Z. Xie. 2011. "Unsupervised Image Segmentation Evaluation and Refinement Using a Multi-scale Approach." *ISPRS Journal of Photogrammetry and Remote Sensing* 66: 473–483. doi:10.1016/j.isprsjprs.2011.02.006.

Ke, Y., L. J. Quackenbush, and J. Im. 2010. "Synergistic Use of Quickbird Multispectral Imagery and LIDAR Data for Object-based Forest Species Classification." *Remote Sensing of Environment* 114: 1141–1154. doi:10.1016/j.rse.2010.01.002.

Kim, M., M. Madden, and T. A. Warner. 2009. "Forest Type Mapping Using Object-specific Texture Measures from Multispectral IKONOS Imagery: Segmentation Quality and Image Classification Issues." *Photogrammetric Engineering & Remote Sensing* 75 (7): 819–829. doi:10.14358/PERS.75.7.819.

Lang, S., E. Schöpfer, and T. Langanke. 2009. "Combined Object-based Classification and Manual Interpretation–Synergies for a Quantitative Assessment of Parcels and Biotopes." *Geocarto International* 24 (2): 99–114. doi:10.1080/10106040802121093.

Levine, M. D., and A. M. Nazif. 1982. "An Experimental Rule Based System for Testing Low Level Segmentation Strategies." In *Multi-Computers and Image Processing: Algorithms and Programs*, edited by K. Preston and L. Uhr, 149–160. New York, NY: Academic Press.

Lucieer, A., and A. Stein. 2002. "Existential Uncertainty of Spatial Objects Segmented from Satellite Sensor Imagery." *IEEE Transactions on Geoscience and Remote Sensing* 40 (11): 2518–2521. doi:10.1109/TGRS.2002.805072.

Marçal, A. R. S., and A. S. Rodrigues. 2009. "A Method for Multi-Spectral Image Segmentation Evaluation Based on Synthetic Images." *Computers & Geosciences* 35: 1574–1581. doi:10.1016/j.cageo.2008.11.008.

Meinel, G., and M. Neubert. 2004. "A Comparison of Segmentation Programs for High Resolution Remote Sensing Data." *International Archives of Photogrammetry and Remote Sensing* 35 Part B, Commission 4: 1097–1105.

Molenaar, M. 1998. *An Introduction to the Theory of Spatial Object Modelling for GIS*. London: Taylor & Francis.

Möller, M., L. Lymburner, and M. Volk. 2007. "The Comparison Index: A Tool for Assessing the Accuracy of Image Segmentation." *International Journal of Applied Earth Observation and Geoinformation* 9: 311–321. doi:10.1016/j.jag.2006.10.002.

Myint, S. W., P. Gober, A. Brazel, S. Grossman-Clarke, and Q. Weng. 2011. "Per-pixel Vs. Object-based Classification of Urban Land Cover Extraction Using High Spatial Resolution Imagery." *Remote Sensing of Environment* 115: 1145–1161. doi:10.1016/j.rse.2010.12.017.

Neubert, M., and G. Meinel. 2003. "Evaluation of Segmentation Programs for High Resolution Remote Sensing Applications." Paper presented at the proceedings of the joint ISPRS/EARSeL workshop "high resolution mapping from space 2003", Hannover, October 6–8.

Paglieroni, D. W. 2004. "Design Considerations for Image Segmentation Quality Assessment Measures." *Pattern Recognition* 37: 1607–1617. doi:10.1016/j.patcog.2004.01.017.

Pu, R., and S. Landry. 2012. "A Comparative Analysis of High Spatial Resolution IKONOS and Worldview-2 Imagery for Mapping Urban Tree Species." *Remote Sensing of Environment* 124: 516–533. doi:10.1016/j.rse.2012.06.011.

Radoux, J., and P. Defourny. 2007. "A Quantitative Assessment of Boundaries in Automated Forest Stand Delineation Using Very High Resolution Imagery." *Remote Sensing of Environment* 110: 468–475. doi:10.1016/j.rse.2007.02.031.

Radoux, J., and P. Defourny. 2008. "Quality Assessment of Segmentation Results Devoted to Object-Based Classification." In *Object Based Image Analysis – Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, edited by T. Blaschke, S. Lang, and G. J. Hay, 257–271. Berlin: Springer-Verlag Berlin Heidelberg.

Ragia, L., and S. Winter. 2000. "Contributions to a Quality Description of Areal Objects in Spatial Data Sets." *ISPRS Journal of Photogrammetry & Remote Sensing* 55: 201–213. doi:10.1016/S0924-2716(00)00020-4.

Schiewe, J., L. Tufte, and M. Ehlers. 2001. "Potential and Problems of Multi-Scale Segmentation Methods in Remote Sensing." *GIS* 6 (1): 34–39.

Schöpfer, E., and S. Lang. 2006. "Object Fate Analysis – A Virtual Overlay Method for the Categorisation of Object Transition and Object-based Accuracy Assessment." Paper presented at the international archives of photogrammetry, remote sensing and spatial information sciences, Salzburg, July 4–5.

Schöpfer, E., S. Lang, and F. Albrecht. 2008. "Object-fate Analysis-spatial Relationships for the Assessment of Object Transition and Correspondence." In *Object-based Image Analysis – Spatial Concepts for Knowledge-driven Remote Sensing Applications*, edited by T. Blaschke, S. Lang, and G. J. Hay, 785–801. Berlin: Springer-Verlag Berlin Heidelberg.

Stein, A., and K. de Beurs. 2005. "Complexity Metrics to Quantify Semantic Accuracy in Segmented Landsat Images." *International Journal of Remote Sensing* 26 (14): 2937–2951. doi:10.1080/01431160500057749.

Tian, J., and D. M. Chen. 2007. "Optimization in Multi-scale Segmentation of High Resolution Satellite Images for Artificial Feature Recognition." *International Journal of Remote Sensing* 28 (20): 4625–4644. doi:10.1080/01431160701241746.

Walter, V. 2004. "Object-based Classification of Remote Sensing Data for Change Detection." *ISPRS Journal of Photogrammetry & Remote Sensing* 58: 225–238. doi:10.1016/j.isprsjprs.2003.09.007.

Weidner, U. 2008. "Contribution to the Assessment of Segmentation Quality for Remote Sensing Applications." Paper presented at the international archives of the photogrammetry, remote sensing and spatial information sciences, Beijing, July 3–11.

Yang, L., F. Albregtsen, T. Lonnestad, and P. Grottum. 1995. "A Supervised Approach to the Evaluation of Image Segmentation Methods." Paper presented at the proceedings of the 6th international conference on computer analysis of images and patterns, Prague, September 6–8.

Yu, Q., P. Gong, N. Clinton, G. Biging, M. Kelly, and D. Schirokauer. 2006. "Object-based Detailed Vegetation Classification with Airborne High Spatial Resolution Remote Sensing Imagery." *Photogrammetric Engineering & Remote Sensing* 72 (7): 799–811. doi:10.14358/PERS.72.7.799.

Zhan, Q., M. Molenaar, K. Tempfli, and W. Shi. 2005. "Quality Assessment for Geo-spatial Objects Derived from Remotely Sensed Data." *International Journal of Remote Sensing* 26 (14): 2953–2974. doi:10.1080/01431160500057764.

Zhang, H., J. E. Fritts, and S. A. Goldman. 2008. "Image Segmentation Evaluation: A Survey of Unsupervised Methods." *Computer Vision and Image Understanding* 110: 260–280. doi:10.1016/j.cviu.2007.08.003.

Zhang, Y. J. 1996. "A Survey on Evaluation Methods for Image Segmentation." *Pattern Recognition* 29 (8): 1335–1346. doi:10.1016/0031-3203(95)00169-7.

Zhang, Y. J. 1997. "Evaluation and Comparison of Different Segmentation Algorithms." *Pattern Recognition Letters* 18: 963–974. doi:10.1016/S0167-8655(97)00083-4.