ISPRS Journal of Photogrammetry and Remote Sensing 106 (2015) 28–41

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

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

Scale parameter selection by spatial statistics for GeOBIA: Using mean-shift based multi-scale segmentation as an example

Dongping Ming [a,*], Jonathan Li [b], Junyi Wang [b], Min Zhang [a]

[a] School of Information Engineering, China University of Geosciences (Beijing), 29 Xueyuan Road, Haidian, Beijing 100083, China
[b] Department of Geography and Environmental Management, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2L 3G1, Canada

ARTICLE INFO

Article history:
Received 13 August 2014
Received in revised form 9 March 2015
Accepted 20 April 2015
Available online 20 May 2015

Keywords:
GEOBIA
Scale selection
Statistics
Mean-shift segmentation
Average local variance

ABSTRACT

Geo-Object-Based Image Analysis (GEOBIA) is becoming an increasingly important technology for information extraction from remote sensing images. Multi-scale image segmentation is a key procedure that partitions an image into homogeneous parcels (image objects) in GEOBIA. Hierarchical image objects also provide a better representation result than a single-scale representation. However, scale selection in multi-scale image segmentation is always difficult for high-performance GEOBIA. This paper first generalizes the commonly used segmentation scale parameters into three aspects: spatial bandwidth (spatial distance between classes), attribute bandwidth (difference between classes) and merging threshold. Next, taking mean-shift multi-scale segmentation as an example, this paper proposes a spatial and spectral statistics-based scale parameter selection method for object-based information extraction from high spatial resolution remote sensing images. The main idea of this proposed method is to use the ALV graph to replace the semivariogram to pre-estimate the optimal spatial bandwidth. Next, the selection of the optimal attribute bandwidth and the merging threshold are based on the ALV histogram and simple geometric computation, respectively. This study uses Ikonos, Quickbird and aerial panchromatic images as the experimental data to verify the validity of the proposed scale parameter selection method. Experiments based on quantitative multi-scale segmentation evaluation testify to the validity of this method. This pre-estimation-based scale parameter selection method is practically helpful and efficient in GEOBIA. The idea of this method can be further extended to other segmentation algorithms and other sensor data.

1. Introduction

With the improvements in satellite sensor technology, the spatial resolution of remote sensing images has significantly increased. In very high spatial resolution images, it often occurs that different classes have the same spectral reflectance or the same class has different spectral reflectances (strong spectral variability within class). Thus, pixel-based multi-spectral image classification not only leads to misclassification but also results in broken patches. GEOBIA (Hay and Castilla, 2008), the geo-related sub-discipline of OBIA (Object-Based Image Analysis), has become increasingly commonplace over the last decade (Blaschke, 2010), and its popularity continues to sharply increase (Blaschke et al., 2014) because it can effectively incorporate spatial information and expert knowledge into the classification, and the classified image objects are a useful link on which remote sensing and GIS can be integrated (Blaschke et al., 2008; Blaschke, 2010). Multi-scale image segmentation is the foundational procedure of OBIA in which the digital image is transformed from discrete pixels into homogeneous image object primitives (Vieira et al., 2012). Blaschke et al. (2004) provide an overview of numerous segmentation techniques used in remote sensing. However, the real challenge is to define appropriate segmentation parameters (Hay et al., 2005). Object-based scale selection (scale parameter selection in the image segmentation, see Ming et al. (2011)) is the key to GEOBIA because an inappropriate scale will lead to over-segmentation or insufficient segmentation (Ming et al., 2012), which will directly reduce the accuracy and efficiency of multi-scale information extraction from high spatial resolution remote sensing images (Myint et al., 2011; Ming et al., 2011; Dronova et al., 2012). Although GEOBIA is becoming increasingly prominent in remote sensing science (Blaschke et al., 2008), the selection of segmentation scale parameters is often dependent on

* Corresponding author.
E-mail address: mingdp@cugb.edu.cn (D. Ming).

http://dx.doi.org/10.1016/j.isprsjprs.2015.04.010

subjective trial-and-error methods. A ready-to-use application allowing the user to evaluate the scale parameter as a function of the intrinsic spatial structure of images before segmentation is not common.

Kim et al. (2008) noted that defining the most suitable scale for image segmentation is still problematic as no objective method currently exists for setting the scale parameter in segmentation algorithms. Currently, several multi-scale segmentation algorithms have been proposed and applied in remote sensing image analysis, including watershed segmentation (Vincent and Soille, 1991; Scheunders and Sijbers, 2002), multi-resolution segmentation (Definiens Developer®) (Baatz and Schäpe, 2000; Peña-Barragán et al., 2011; Duro et al., 2012) and mean-shift segmentation (Comaniciu and Meer, 2002; Comaniciu, 2003; Rao et al., 2009; Zhou et al., 2013). Among these algorithms, watershed segmentation has specific scale parameter, namely, the sampling window size, valley threshold and catchment area threshold; however, the relationship between segmentation results on different scales cannot be easily determined. Multi-resolution segmentation (Baatz and Schäpe, 2000; Baatz et al., 2000; Benz et al., 2004) is based on the idea of Fractal Net Evolution approach (Baatz and Shäpe, 1999), and it not only provides a scale parameter but also incorporates the hierarchical relationship between different levels. The scale parameter in multi-resolution segmentation is very important and it determines the average image object size by determining the upper limit for a permitted change of heterogeneity throughout the segmentation process. However, the scale parameter is an abstract value used to determine the maximum possible change of heterogeneity caused by fusing several objects, and it is difficult for a user to quantitatively select the optimal scale parameter without repetitious trials because the practical meaning of scale parameter is different in different segmentation mode (Normal, Spectral difference and Sub obj.line analysis) provided by eCognition. The relation between the scale parameter and the image data is complicatedly tacit and cannot be directively and easily built.

Because the mean-shift segmentation algorithm has the advantages of a specific scale parameter and a hierarchical relationship between segmentation levels, this paper adopts the mean-shift multi-scale segmentation algorithm to extract the homogenous parcels from high spatial resolution images. Based on the essence of spatial dependence for scale, this paper combines the theories of geospatial statistics and pattern recognition and proposes an optimal scale parameter selection method based on average local variance (ALV) for mean-shift image segmentation. Compared with existing studies on scale parameter selection that are based on post-evaluation, the characteristic feature of this method is that it is based on pre-estimation. The method is further validated based on multi-scale segmentations of Ikonos and Quickbird panchromatic image data.

2. Object-based scale and optimal object-based scale selection

Scale is a broadly used term in geoscience and has a variety of meanings in different contexts. Although GEOBIA has drawn considerable attention in remote sensing image processing and analysis, a specific concept of an object-based scale of remote sensing images has not been given. Ming et al. (2011) generalized the three levels of connotation of the spatial scale of remote sensing images (pixel-based, object-based and pattern-based scales). The object-based scale is considered the size of the meaningful unit (image primitive). An image object is defined (Definiens, 2007) as a group of connected pixels that have homogenous features. Thus, the object-based scale refers to the spatial extent or the size of image object, and the optimal object-scale refers to the optimal size of the smallest class or the optimal size for different classes. From the viewpoint of using an algorithm to extract the image objects, the object-based scale corresponds to the scale parameters in the multi-scale image segmentation, and the optimal object-based scale corresponds to the optimal segmentation level that contains the most pure objects and the least mixed objects, and as a result, the subsequent object-based classification can achieve high accuracy.

In image processing, some attempts have been made to select the optimal scale parameters for multi-scale segmentation. For example, Claudio (2007), Tian and Chen (2007), Tan et al. (2007), Kim et al. (2008), He et al. (2009) and Johnson and Xie (2011) used the indices of the homogeneity within the segmentation parcels and the heterogeneity between the segmentation parcels to select the optimal scale parameter. Drăgut et al. (2009) presented a focal mean statistics-based procedure to optimize the parameterization and to scale for terrain segmentation. Drăgut et al. (2010) proposed the idea that the local variance (LV) of object heterogeneity within a scene can indicate the appropriate scale level, and they introduced an ESP (Estimation of Scale Parameter) tool to find optimal parameters for the multi-resolution segmentation. Karl and Maurer (2010) use variogram-based spatial dependency prediction to determine appropriate segmentation scales for producing land-management information. The rule Eid et al. (2010) use to detect the optimal scale is that if the area of an object is kept static or nearly static in a set of successive scales, then any of these scales can be chosen as the appropriate scale. Anders et al. (2011) evaluate the quality of the segmentation results for each specific geomorphological feature type to optimize segmentation parameters. Zhao et al. (2012) employed the changed ROC-LV method, similar to the ESP tool by Drăgut et al. (2010), to judge the optimal scales in the slope segmentation by using multi-resolution segmentation and eCognition software. However, these methods and applications are actually intended to select the scale by post-evaluation of the multi-scale segmentation, not pre-estimation of the optimal scale parameters. Additionally, most of them are basically based on the multi-resolution segmentation provided by Definiens Developer®, in which the meaning of the scale parameter is complicated, making it practically difficult to understand the relationship between the scale selection indictors and the scale parameter.

Scale parameters selection for multi-scale image segmentation in image processing is also called bandwidths selection for mode clustering in pattern recognition. In pattern recognition, some bandwidth selection methods in this context also have been explored for multi-scale clustering. Comaniciu (2003) discussed the optimal bandwidths (based on asymptotic bias-variance trade-off) originally derived for the purpose of multivariate normalized density estimation, however this idea was discarded quickly for practical considerations (Einbeck, 2011). An alternative family of methods, which is tailored toward the extraction of reliable scale information for multi-dimensional data, attempts to maximize the stability of the partitioning under variation of the bandwidth (Comaniciu et al., 2001; Comaniciu, 2003; Einbeck, 2011). However, it is not clear whether a bandwidth which is optimal for the clustering is necessarily optimal for the problem of finding local modes. What's more, these works are mostly based on discrete simulated data sets which seldom present in reality, therefore the resulting bandwidth is often of little practical use (Li et al., 2005; Li et al., 2007; Zhang et al., 2012). In addition, (Park et al., 2009; Sun and Xu, 2010; Vojir et al., 2014) employ estimation of local structure to select the bandwidth or scale parameter for image segmentation or filtering. However, these methods are mainly based on natural color images that are greatly differing from remote sensing images, their performance on complex remote sensing image segmentation need to be further verified.

In fact, the essence of determining the optimal scale parameters is the statistical pre-estimation of the spatial autocorrelation of objects within the image. The optimal scale is essentially the critical point at which the spatial dependence exists or does not exist. The critical point is just reflects the meaning of the spatial statistical term "range". Based on this idea, Ming et al. (2012) combined the classical semivariogram method with the segmentation scale selection of remote sensing images, which achieves good performance. Spatial statistics is deemed as a feasible and inevitable approach for scale selection, especially in the geo-application of remote sensing images. This paper applies statistical methods (especially geospatial statistics) to pattern recognition to select the optimal bandwidths. This approach can be explained by classical geospatial statistical theory and is practically meaningful for multi-scale information extraction from remote sensing images.

3. Multi-scale segmentation based on mean-shift algorithm

The mean-shift algorithm by Fukunaga and Hostetler (1975) is a robust and adaptive clustering algorithm with non-parametric density estimation, and it does not require a priori knowledge of the number of clusters. This algorithm can shift the points in the feature space to the local maxima of the density function by effective iterations and can achieve fast convergence. Thus, it has been successfully used in image segmentation (Comaniciu et al., 2001; Zhou et al., 2013; Yang et al., 2013).

3.1. Introduction to mean shift clustering

Given n data points x_i, $i = 1, \ldots, n$ in the d-dimensional space R^d, the kernel density estimation at the location x can be calculated by

$$\hat{f}_K(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h_i^d}k\left(\left\|\frac{x-x_i}{h_i}\right\|^2\right) \tag{1}$$

with the bandwidth parameter $h_i > 0$. The kernel K is a spherically symmetric kernel with bounded support (Polat and Ozden, 2006),

$$K(x) = c_{k,d}k(\|x\|^2) > 0 \quad \|x\| \leqslant 1 \tag{2}$$

where the normalization constant $c_{k,d}$ ensures that $K(x)$ integrates to one. The function $k(x)$ is called the profile of the kernel. Assuming derivative of the kernel profile $k(x)$ exists, using $g(x) = -k'(x)$ as the profile, the kernel $G(x)$ is defined as $G(x) = c_{g,d}$ $g(\|x\|^2)$. The following property can be proven by taking the gradient of Eq. (1) as follows,

$$m_G(x) = C\frac{\hat{\nabla}f_K(x)}{\hat{f}_G(x)} \tag{3}$$

where $m_G(x)$ is called mean shift vector. C is a positive constant and, it shows that, at location x, the mean shift vector computed with kernel G is proportional to the normalized density gradient estimate obtained with kernel K. The mean shift vector is defined as follows

$$m_{h,G}(x) = \frac{\sum_{i=1}^{n}x_i G(\|\frac{x-x_i}{h}\|^2)}{\sum_{i=1}^{n}G(\|\frac{x-x_i}{h}\|^2)} - x \tag{4}$$

The mean shift vector thus points toward the direction of maximum increase in the density (Comaniciu and Meer, 2002; zheng et al., 2010). The mean shift procedure is obtained by successive computation of the mean shift vector and translation of the kernel $G(x)$ by the mean shift vector. At the end of the procedure, it is guaranteed to converge at a nearby point where the estimate has zero gradient (Cheng, 1995). In other words, it is a hill climbing technique to the nearest stationary point of the density. The iterative equation is given by

$$y_{j+1} = \frac{\sum_{i=1}^{n}\frac{x_i}{h_i^{d+2}}G\left(\left\|\frac{y_j-x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n}\frac{1}{h_i^{d+2}}G\left(\left\|\frac{y_j-x_i}{h}\right\|^2\right)} \quad j = 1, 2, \ldots \tag{5}$$

The initial position of the kernel (starting point to calculate y_1) can be chosen as one of the data points x_i. Usually, the modes (local maximum) of the density are the convergence points of the iterative procedure. For more details, please refer to Cheng (1995), Polat and Ozden (2006) and Wu and Yang (2007).

3.2. Mean-shift based image segmentation and the meanings of the bandwidths

Most conventional color clustering algorithms for image segmentation are based on the information in the color space. However, clustering in the color space usually does not provide satisfactory performance because it lacks information about the spatial configuration. To resolve this limitation, in addition to color information, the spatial information is often incorporated into the feature space representation, and the mean-shift algorithm-based color segmentation method can be applied to obtain reliable segmentation results.

Mean-shift-based segmentation is a straightforward extension of the discontinuity preserving smoothing algorithm (Comaniciu and Meer, 2002). When the multi-variate kernels (h_s, h_r) are used to replace the single bandwidth parameter of the kernel function, formula (2) can be expressed by

$$K_{h_s,h_r}(x) = \frac{c}{h_s^2 h_r^2}k\left(\left\|\frac{x_s}{h_s}\right\|^2\right)k\left(\left\|\frac{x_r}{h_r}\right\|^2\right) \tag{6}$$

Spatial bandwidth h_s is the spatial distance between classes in the spatial domain, and it indicates the spatial window size in the segmentation. Attribute bandwidth (or spectral bandwidth) h_r represents the spectral difference between classes in the spectral domain. The mean-shift algorithm segments the image by grouping together all pixels that are closer than h_s in the spatial domain and h_r in the spectral domain and then connecting the corresponding convergence points. In the multi-scale segmentation, class labels for all pixels are available for building the relationships between different levels of segmentation results. An optional step of eliminating spatial regions containing less than M pixels is another approach to performing multi-scale segmentation. For further details, please refer to Comaniciu and Meer (2002).

In mean-shift-based multi-scale segmentation, there are three scale parameters (h_s, h_r, M) that determine the scale level of the segmentation. More concretely, the scale parameters h_s, h_r and M relate to the spatial size of the patches making up a class, the spectral "size" or compactness of classes and the minimum acceptable spatial size of patches in a class, respectively.

4. Object-based scale selection by spatial statistics: using mean-shift-based multi-scale segmentation as an example

Selection of the optimal scale parameters is the key to multi-scale segmentation because they can determine the pattern recognition resolution and the precision of the information extraction from the remote sensing image. However, it must be kept in mind that there is not absolutely ideal scale for the variously sized, shaped, and spatially distributed image-objects composing a scene, which also reflects the substance of the Modifiable Area Unit Problem (MAUP, Openshaw, 1984). A compromise solution is to compute a theoretically ideal scale on which the geographic information can be recognized by an GEOBIA-based classification, and the overall classification accuracy can be theoretically guaranteed by using the selected optimal scale parameters.

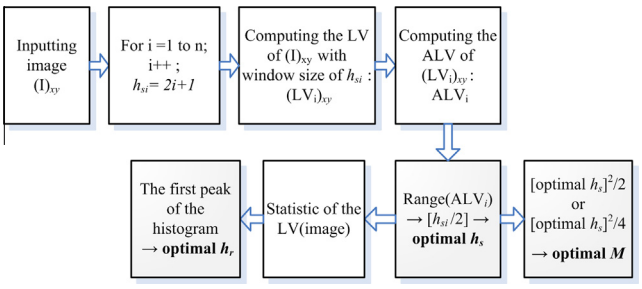


Fig. 1. Workflow of pre-estimating the optimal scale parameters.

Additionally, the three scale parameters in mean-shift-based multi-scale segmentation (h_s, h_r, M) are universal for representing the scale meanings in GEOBIA and are actually independent of the segmentation algorithms. This paper only uses mean-shift-based multi-scale segmentation as an example, and it applies classical geospatial and spectral statistics to pattern recognition to select the optimal scale parameters (h_s, h_r, M) in GEOBIA. Fig. 1 demonstrates the workflow of pre-estimating the optimal scale parameters based on the Local Variance (LV) (Woodcock and Strahler, 1987) and spectral statistical methods, where $(I)_{xy}$ represents an image with x rows and y columns, and i indicates the repeating times of the computation.

4.1. Spatial bandwidth selection based on local variance

Spatial bandwidth h_s has a distinct effect on the clustering output compared to the spectral bandwidth h_r (Comaniciu and Meer, 2002). Only features with large spatial support are represented in the segmentation when h_s increases. How to predict the optimal spatial bandwidth h_s or the appropriate spatial window size is an important issue for image segmentation and GEOBIA. Combining the theories of classical geospatial statistics and pattern recognition, Ming et al. (2012) propose an optimal spatial bandwidth selection method based on a semivariogram for mean-shift image segmentation and achieve good performance. The idea behind their approach is that the range can be deemed as the measurement of similarity between variables, and it can indicate the size of a spatial object, a spatial phenomenon or a spatial pattern (Wang et al., 2001). Similarly, because spatial bandwidth h_s corresponds to the size of the spatial window (abbreviated as w_s and $w_s = 2 * h_s + 1$), the LV method can also be used to estimate the optimal spatial bandwidth.

Local variance (LV) was introduced by Woodcock and Strahler (1987) to reveal the spatial structure of images and, thus, to characterize the relationship between the spatial resolution and objects in the scene. In previous studies, local variance calculated the mean value of the standard deviation by passing an n pixel by n pixel moving window for each pixel and then taking the mean of all local variance (Average Local Variance, ALV) over the entire image as an indication of the local variability in an image. This idea was later introduced in the context of OBIA by Kim et al. (2008) and Drăgut et al. (2010). However, the related works are mostly based on multi-resolution segmentation of the eCognition software, and the scale selections are a post-evaluation of the segmentation in nature. This paper proposes to employ the local variance on the pixel level before segmentation to pre-estimate the optimal spatial bandwidth in clustering-based pattern recognition.

Fig. 1 demonstrates the workflow of estimating the optimal spatial bandwidth h_s based on the LV method. The computation of LV is actually similar to the computation of synthetic SV (semivariance), considering the different directions (horizontal, vertical and two diagonal directions). Therefore, using LV to estimate the optimal spatial bandwidth is rational and feasible and is actually an extension of the semivariogram-based method proposed by Ming et al. (2012) because, from the computation of LV, LVs with different window sizes are synthetic variances in four directions and with different lags. Fig. 2 demonstrates the calculation of semivariance and local variance at one pixel, where both SV and ALV are the average values of accumulation on the whole image. For further details on the computing formulae, please refer to Ming et al. (2012) and Ming et al. (2010). Consequently, the principle of optimal spatial bandwidth selection based on ALV is similar to that of SV. If the ALV chart, namely, the curve of ALV varying with the window size, is used to substitute the semivariogram, the range of the ALV chart can be used to select the optimal spatial bandwidth h_s. Here the ALV chart is called ALvariogram. Next, the key of this task is to derive the range of ALvariogram.

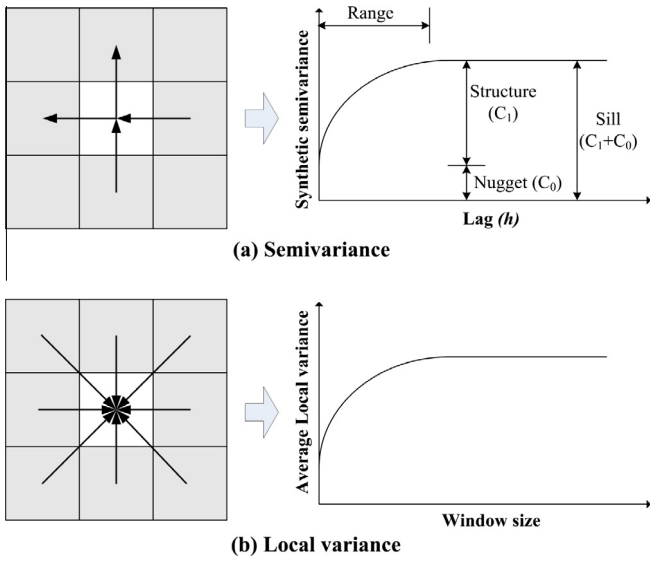


Fig. 2. Schematic diagrams for the computation of semivariance and local variance.

4.2. Attribute bandwidth selection based on spectral statistics

The direct estimation of attribute bandwidth is generally difficult because one needs a priori knowledge of the neighborhood size in which the fitting parameters are to be estimated (Comaniciu, 2003). However, once the spatial bandwidth h_s is selectively determined, attribute bandwidth h_r can be further estimated based on the result of the spatial bandwidth selection. Within the large sample approximation, the estimation bias can be canceled, and the estimation of the true local covariance of the underlying mode distribution is allowed (Comaniciu, 2003). The attribute bandwidth selection criterion is that the underlying mode distribution has the local variance equal to the attribute bandwidth that maximizes the magnitude of the normalized mean shift vector, which means the magnitude of the bandwidth-normalized mean-shift vector is maximized when the bandwidth is equal to the covariance matrix. Comaniciu (2003) mathematically proves the bandwidth selection theorem in detail.

Considering the simplicity of the method and the particularity of image segmentation, this paper simplifies the attribute bandwidth selection method and uses the matrix of local variance to estimate the optimal attribute bandwidth. Based on the selected optimal spatial bandwidth h_s, an LV matrix with x rows and y columns (marked with LV_{xy}) can be obtained. Then, the first left peak of the histogram of LV_{xy} can be used as the indicator of the optimal attribute bandwidth h_r. The workflow of the method is

demonstrated in Fig. 3, in which a schematic example is shown in the right part.

In this workflow, d refers to the bit depth of the original image. Because LV approximates the mean square of the difference values between the central pixel and its neighbors, the minimum bin is set as 4, so that the optimal h_r, the square root of LV at the first peak location, is absolutely more than 2. Of course, if the radiometric resolution of the original image is very high, such as d is16 bit, the minimum bin can be set as a larger value.

Additionally, to avoid the pseudo peaks, histogram curve smoothing is necessary. There are many approaches to smooth the histogram curve, such as interpolation or curve fitting. However, for simplicity, some pseudo peaks can be discarded according to the number of locations at ascending slope and descending slope. Then, the numbers of the bins need to be increased until the first peak of the LV_{xy} histogram is authentic and practically meaningful.

4.3. Merging threshold selection based on semivariance

The essence of merging threshold M is the area or pixel number of the minimum meaningful object. The segmented parcel that is

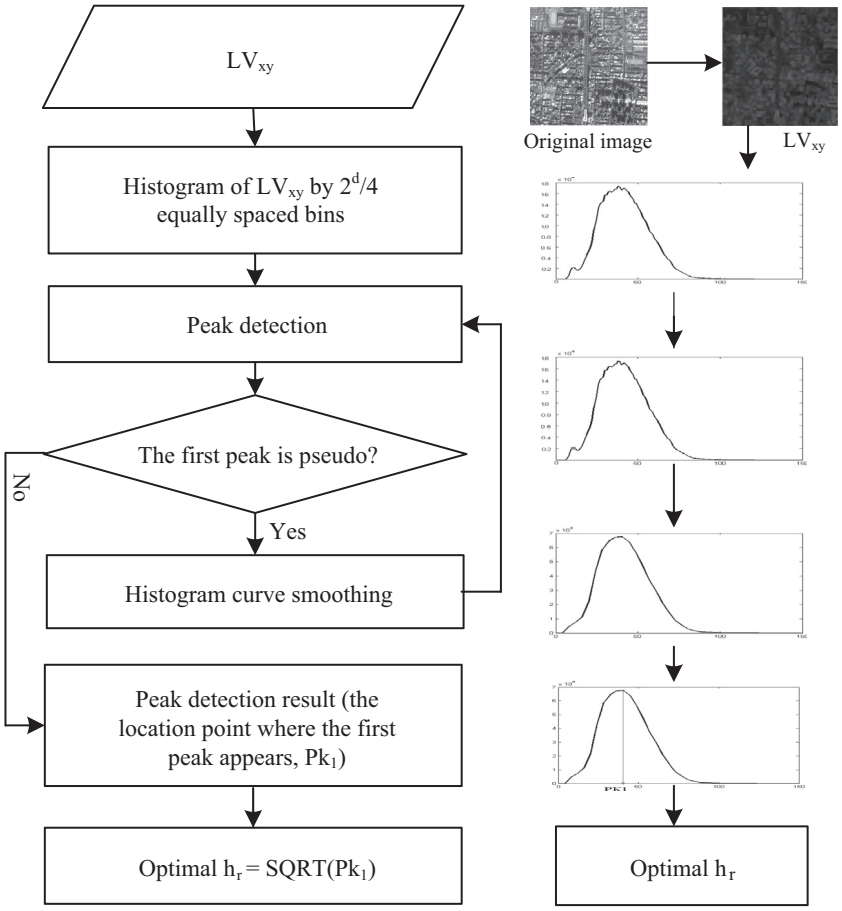


Fig. 3. Workflow of optimal bandwidth selection.

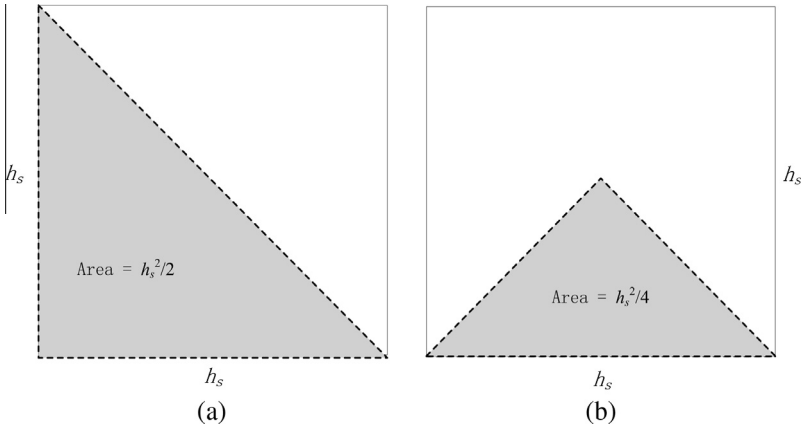


Fig. 4. Sketch map of computing the merging threshold.

smaller than M is deemed meaningless and will be merged into the most similar adjacent parcel with the least attribute difference. As mentioned above, the range of the semivariogram can indicate the size of a spatial object, and it can be predicted by ALvariogram, so M can be roughly predicted by a semivariogram or ALvariogram.

As is generally known, the object types in an image are usually complicated and an absolutely ideal object size that suits different classes does not exist. Therefore some priori knowledge and assumptions should be involved. The basic principle of setting M is that the value of M should be neither too low nor too high.

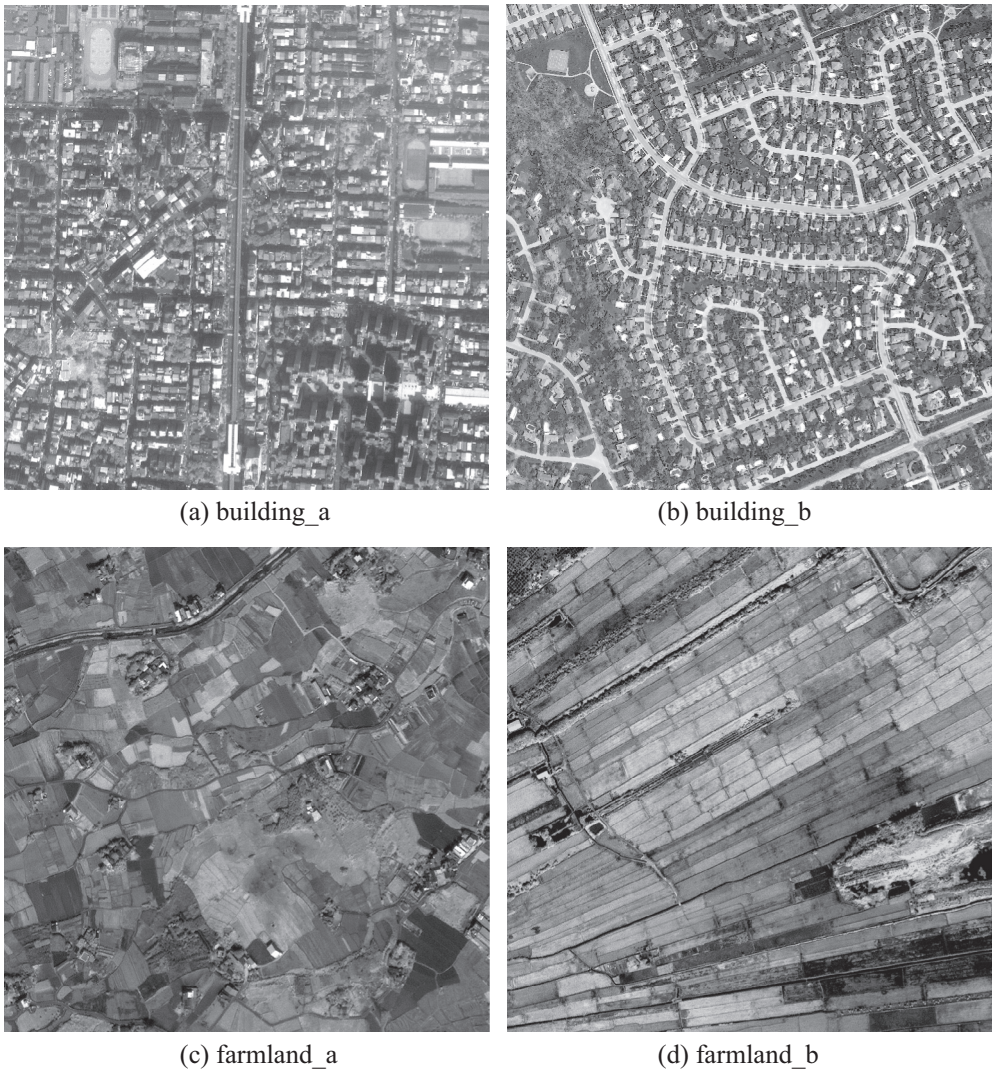


Fig. 5. Experimental panchromatic high spatial resolution images.

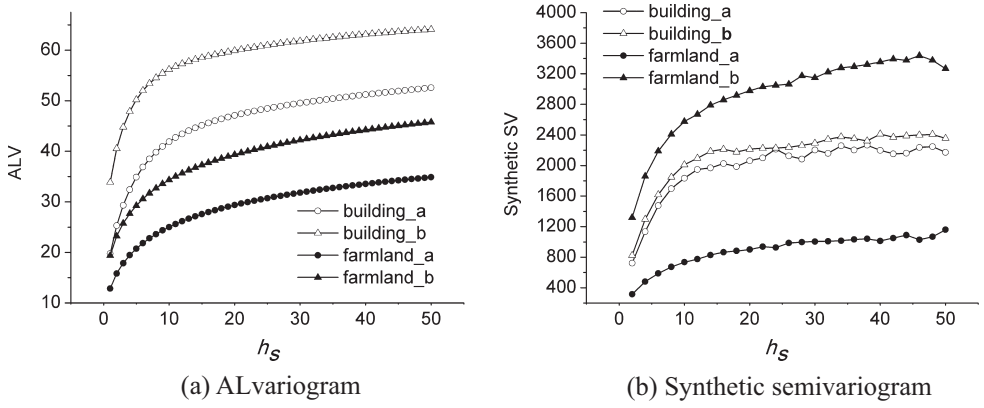


Fig. 6. ALV and synthetic SV of experimental images.

Firstly, to ensure that the meaningful parcels with small size will not be merged in segmentation, the value of M should not be too high. However, the value of M should not be too low otherwise the segmented parcels with low value of M will be over fractional, which will results in low heterogeneity between the parcels and further results in low separability in GeOBIA. Consequently, the following rules are involved.

(1) When the shape of object is regular and close to a square or rectangle, or when the image mainly covers an artificial building area, $M = \mathrm{INT}(h_s^2/2)$. That means when the shapes of objects are regular, the shape of objects are likely close to square or rectangle. Of course it also may be triangle, which has the minimum area of all the possible cases. As shown in Fig. 4(a), the area of the triangle can be approximated by the $h_s^2/2$, half of the square with size length h_s.

Table 1
The optimal merging thresholds for the experimental images.

Image	building_a	building_b	farmland_a	farmland_b
Object shape	Regular	Regular	Irregular	Regular
Optimal merging threshold M	144	113	121	200

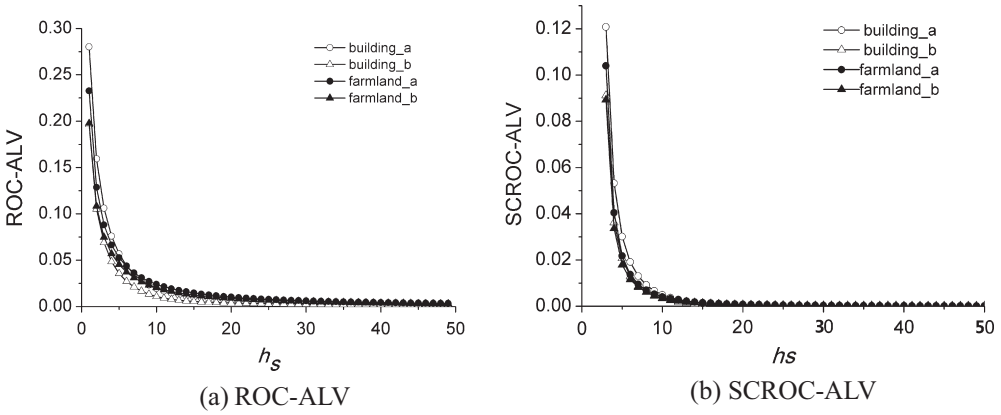


Fig. 7. ROC-ALV and SCROC-ALV of experimental images.

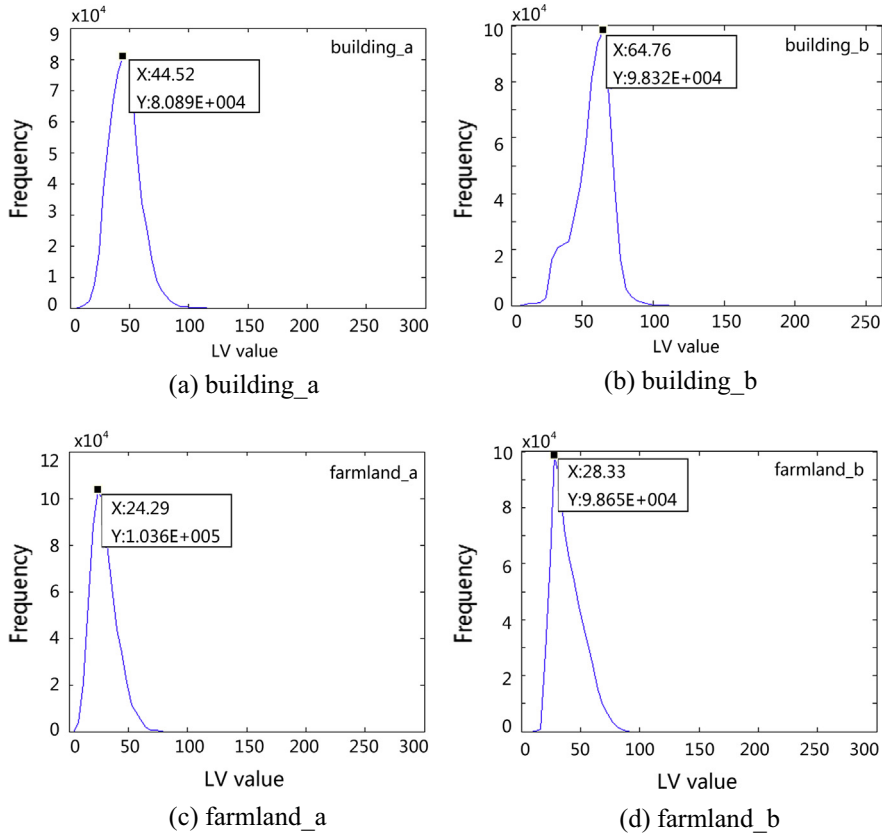


Fig. 8. Histograms of LV.

(2) When the shape of the object is irregular, $M = \mathrm{INT}(h_s^2/4)$. That means when the shapes of objects are irregular, the shapes of object could not be simply supposed. However, even when the shapes of objects are irregular, the value of M should not be too low. For sake of this, as shown in Fig. 4(b), the area of the triangle can be approximated by the $h_s^2/4$, quarter of the square with size length h_s.
(3) If not involving any a priori knowledge and not considering the fragmentation of the object shape, M can be set as INT $(h_s^2/4)$. That means $h_s^2/4$ can be used as the threshold of the minimum of the meaningful object.

5. Experiments

As shown in Fig. 5, four high spatial resolution images with a size of 800×800 pixels are used for the experiments. Fig. 5(a) and (b) are covered by building area. Fig. 5(a) is a subset Ikonos panchromatic image with a 1-meter spatial resolution and an area of 0.64 km^2. Fig. 5(b) is a subset aerial panchromatic image obtained from the Erdas Imagine example dataset (the spatial resolution and area are unknown). Fig. 5(c) and (d) are covered by cropland that had been partly harvested. These are Quickbird panchromatic images with a 0.7-meter spatial resolution and an area of 0.45 km^2. Fig. 5(b) is a subset aerial panchromatic imkm^2.

5.1. Computing the optimal scale parameter based on statistics

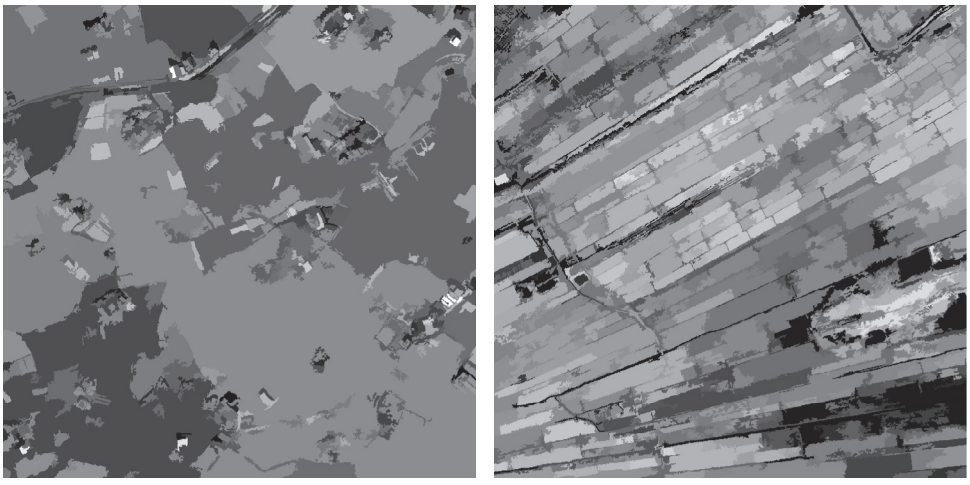
From the viewpoint of information extraction of the high spatial resolution images, the fine segmentation and the precise information are expected. To precisely reflect the change of ALV along the window size w_s, the w_s are set from 3 to 101, and the lag step is set as 2, which means the h_s are set from 1 to 50, and the lag step is set as 1. To compare the role of ALV and SV in selecting the optimal spatial bandwidth h_s, Fig. 6(a) and (b) show the computing results of the ALV and synthetic SV for the four experimental images. For more details on synthetic SV, please refer to Ming et al. (2012).

Fig. 6(a) shows that the ALvariogram is very smooth. However, as shown as in Fig. 6(b), there are many dithers on the synthetic semivariogram, which means the ALV method more easily indicates the optimal spatial band width than the SV method. If ALV and SV are taken as random variants, ALV is more stationary than SV. Next, determining the range of the ALvariogram is the key to selecting the optimal scale parameters in image segmentation.

5.1.1. Estimating the optimal spatial bandwidth

Function fitting is a commonly used way to determine the range of a semivariogram. However, in some practical applications, it is difficult to find a valid solution because the type of joint distribution of the spatial statistic character is usually unknown. Additionally, SV values with short lags play a more important role

(a) building_a with (17, 7, 144) (b) building_b with (15, 8, 113)



(c) farmland_a with (22, 5, 121) (d) farmland_b with (20, 5, 200)

Fig. 9. Segmentation results with the selected optimal scale parameters.

than those with long lags in determining the range of a semivariogram (Zhang, 2005). Based on the methods proposed by Drăgut et al. (2010) and Ming et al. (2012), this paper proposes a simplified range determining method based on the change of ALV. Two measures, namely, the first-order rate of change in ALV (ROC-ALV) and the second-order change in ALV (SCROC-ALV), are used to assess the dynamics of ALV along the spatial bandwidth h_s.

$$[\text{ROC-ALV}]_i = \frac{\text{ALV}_i - \text{ALV}_{i-1}}{\text{ALV}_{i-1}} \tag{7}$$

$$[\text{SCROC-ALV}]_i = [\text{ROC-ALV}]_{i-1} - [\text{ROC-ALV}]_i \tag{8}$$

where $[\text{ROC-ALV}]_i$ is the rate of change in ALV at level h_{si}, and the value of $[\text{ROC-ALV}]_i$ is usually between [0,1]. SCROC-ALV_i is the change of $[\text{ROC-ALV}]_i$, and the value of $[\text{SCROC-ALV}]_i$ is also usually between [0,1]. Most of the $[\text{SCROC-ALV}]_i$ are small fractions.

The thresholds of ROC-ALV and SCROC-ALV are respectively set as 0.01 and 0.001, which means the optimal spatial bandwidth h_{si} is determined by the window size w_{si} at which the $[\text{ROC-ALV}]_i$ is first less than 0.01 and $[\text{SCROC-ALV}]_i$ is first less than 0.001. There are three main reasons for setting the two thresholds in this way. The first reason is that it is based on large amount experiments. The second is the ALV values with short lags play a more important role than those with long lags. The third is that the condition of the ROC-ALV value of 0.01 and the SCROC-ALV value of 0.001 can guarantee that the change of ALV is tiny enough that the corresponding h_{si} can be roughly taken as the range of the ALvariogram.

Fig. 7 demonstrates the change of ROC-ALV and SCROC-ALV along h_s, according to which the optimal spatial bandwidths of building_a, building_b, farmland_a and farmland_b are, respectively, 17, 15, 22 and 20. The optimal spatial bandwidth selection results of farmland_a and farmland_b are very close to those of Ming et al. (2012), who used the same two experimental images but with a smaller size ($512 * 512$). The optimal spatial bandwidth selection results in the current study are 22 and 20, and those in Ming et al. (2012) are 22 and 18.

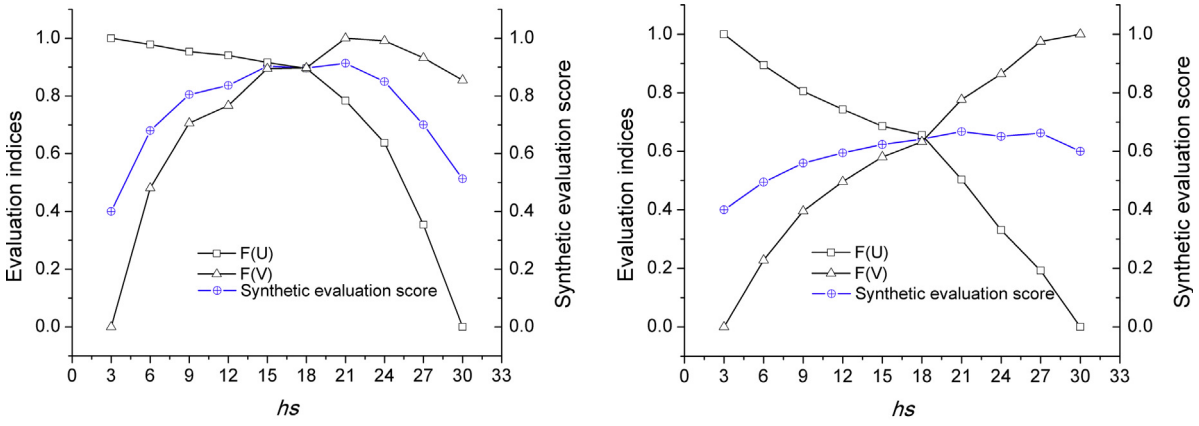
5.1.2. Estimating the optimal attribute bandwidth

Based on the optimal spatial bandwidth selection results, the corresponding histograms of the LV matrices are illustrated in Fig. 8.

According to the attribute bandwidth selection method mentioned in Section 4.2, the optimal attribute bandwidths h_r for the four experimental images are 6.67, 8.05, 4.93 and 5.32. In practical application, the attribute bandwidth is usually an integer. Thus, after the rounding-off process, the optimal h_r for the four experimental images are 7, 8, 5 and 5, respectively.

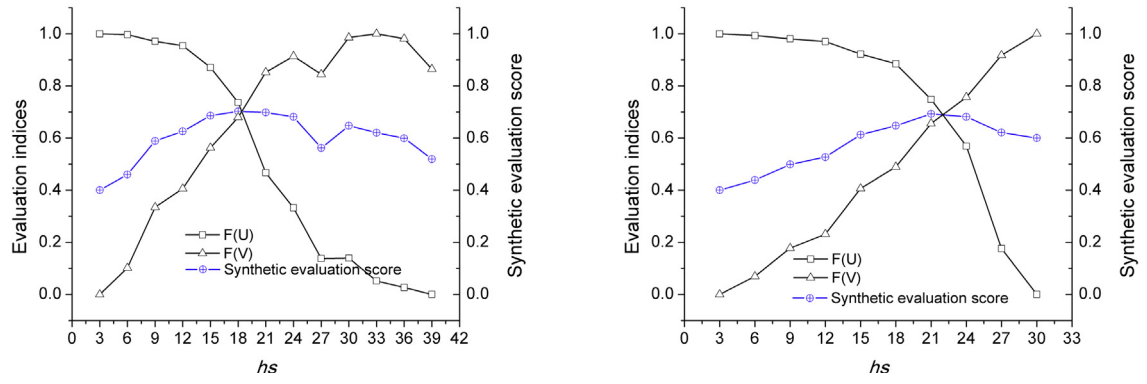
Table 2
Peak ranges and optimal h_s for the four experimental images.

Image	building_a	building_b	farmland_a	farmland_b
Peak range (h_s)	15–21	15–21	18–21	21–24
Peak point (h_s)	21	15	18	21
Optimal h_s by pre-estimation	17	15	22	20



(a) building_a with (h_s, 7, 10)

(b) building_b with (h_s, 8, 10)



(c) farmland_a with (h_s, 5, 10)

(d) farmland_b with (h_s, 5, 10)

Fig. 10. Segmentation evaluations changing with h_s of the four experimental images.

5.1.3. Estimating the merging threshold

Through visual interpretation, the object shapes in the four experimental images can be determined. According to the merging threshold selection method mentioned in Section 4.3, the optimal merging threshold M can be pre-estimated as shown in Table 1.

The segmentation results with the selected optimal scale parameters for the four images are demonstrated in Fig. 9.

5.2. Verification and analysis based on segmentation evaluation

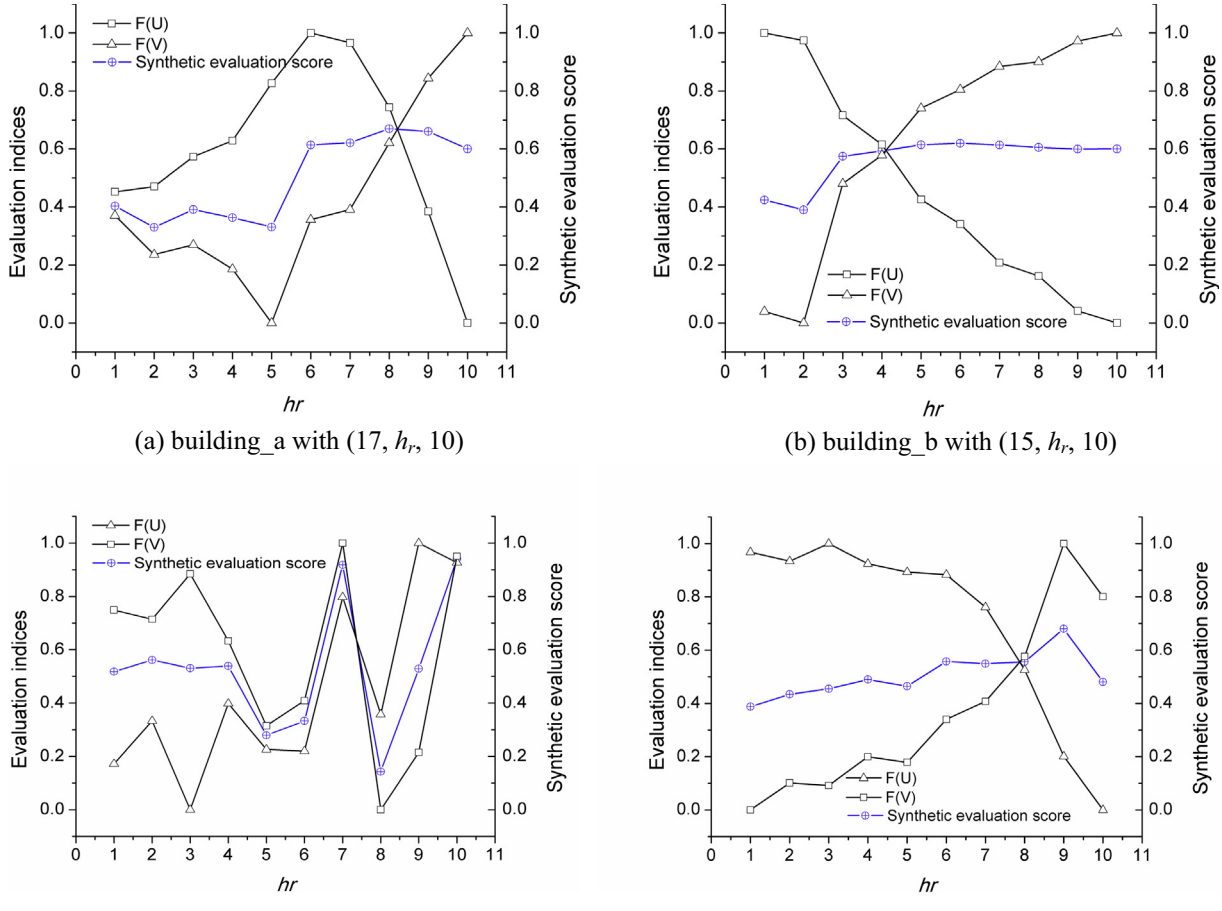
Generally, there are three types of evaluation methods: visual evaluation, quantitative evaluation and indirect evaluation (or application-based evaluation) (Li and Xiao, 2007). This paper uses quantitative evaluation to verify the validity of the scale parameter selection method proposed in this paper. Theoretically, the optimal segmentation scale is a segmentation level that contains the most pure objects and the least mixed objects, that is, the homogeneity within the segmentation parcels is high, and the heterogeneity between the segmentation parcels is low. Therefore, this paper uses the synthetic evaluation model based on homogeneity within the segmentation parcels ($F(U)$) and the heterogeneity between the segmentation parcels ($F(V)$) to test the optimal scale parameters. For more details about the synthetic evaluation model, please refer to Espindola et al. (2006), Kim et al. (2008) and Ming et al. (2012).

This paper sequentially segments the four experimental images with a series of scale parameters using the mean-shift algorithm. First, h_s and h_r are determined, and the fixed value of M is set as a small value of 10 to reduce the impact on segmentation caused by patch merging. Then, the merging threshold M *is* determined based on the selected optimal (h_s, h_r).

According to the research findings of (Ma, 2014), the following two aspects are considered. First, the heterogeneity between the image objects plays a more important role in object-oriented image classification, so the weight of the heterogeneity index is set 0.6. Second, considering the accuracy of object-oriented image classification, neither the homogeneity within the segmentation parcels nor the heterogeneity between the segmentation parcels should be poor, as also proven by (Ma, 2014). That means the segmented parcels with both high homogeneity and heterogeneity are desired. However, it is actually difficult to maintain both high homogeneity and high heterogeneity. A large amount of experiments have proved that high homogeneity often is accompanied by low heterogeneity or high heterogeneity often is accompanied by low homogeneity. Therefore, a compromised solution is proposed based on a large amount of experiments. The optimal bandwidth should meet the following two requirements.

Table 3
Peak ranges and optimal h_r for the four experimental images.

Image	building_a	building_b	farmland_a	farmland_b
Peak range (h_r)	7–9	5–6	4,7,10	7–8
Peak point (h_r)	8	6	10	8
Optimal h_r by pre-estimation	7	8	5	5



(a) building_a with (17, h_r, 10)

(b) building_b with (15, h_r, 10)

(c) farmland_a with (22, h_r, 10)

(d) farmland_b with (20, h_r, 10)

Fig. 11. Segmentation evaluations changing with h_r of the four experimental images.

(1) The synthetic evaluation score at the optimal bandwidth location should be within the peak ranges of the evaluation curve.
(2) The $F(U)$ and $F(V)$ should be greater than 0.4.

5.2.1. Verification of optimal h_s with fixed h_r and M

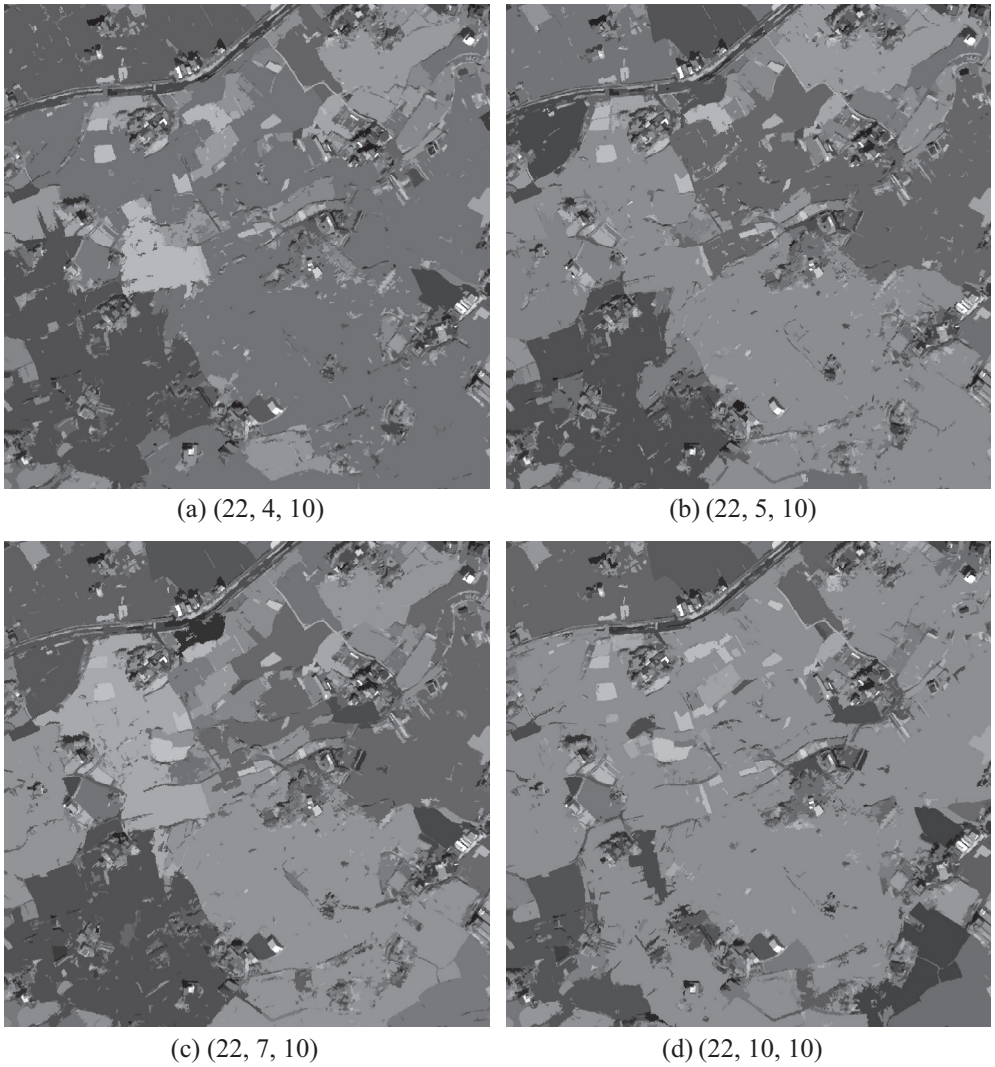
To reduce the computation of verification, the evaluation of the segmentations is based on spatial bandwidth from 3 to 30 with a step of 3. For the sake of comparison, uniform values of h_r and M are used. The value of h_r is set as the selected optimal value. To reduce the impact on segmentation caused by patch merging, the value of M is set as a small value of 10. The segmentation evaluation results with fixed h_r and M ($M = 10$) are demonstrated in Fig. 10.

Fig. 10 shows the peak ranges of the synthetic segmentation evaluation score for the four experimental images. The peak range indicates that the change of the segmentation evaluation scores within the range is very tiny. Table 2 further summarizes the peak ranges meeting the two requirements mentioned above and also lists the optimal h_s determined by LV. Because the h_s interval in the segmentation is 3 and there are inevitable random disturbances in the segmentation, the h_s at the peak points are not absolutely equal to the pre-estimated optimal ones. However, the selected optimal h_s based on LV are within the peak ranges and are close to the maximum points or close to the lower limits of the peak ranges. The results prove that the LV-based spatial bandwidth selection is effective at estimating the appropriate spatial bandwidth before segmentation.

5.2.2. Verification of optimal h_r with fixed h_s and M

The evaluation of the segmentations is based on spectral bandwidth from 1 to 10 with step of 1. For the sake of comparison, uniform values of h_s and M are used. The value of h_s is set as the selected optimal value and M is also set as 10. The segmentation evaluation results with fixed h_s and M ($M = 10$) are demonstrated in Fig. 11.

Table 3 summarizes the peak ranges meeting the two requirements mentioned above and lists the optimal h_r determined by the proposed statistical method. Similarly, the h_r at the peak points are not absolutely equal to the selected optimal ones. However, the selected optimal h_r of two of the four images (building_a and building_b) are within the peak ranges, another two are close to the peak ranges. Taking farmland_a as an example (the difference between pre-estimation and verification is the most distinguished), as shown in Fig. 12, the optimal segmentation



(a) (22, 4, 10) (b) (22, 5, 10)

(c) (22, 7, 10) (d) (22, 10, 10)

Fig. 12. Segmentation results of building_b ($h_r = 4, 5, 7, 10$).

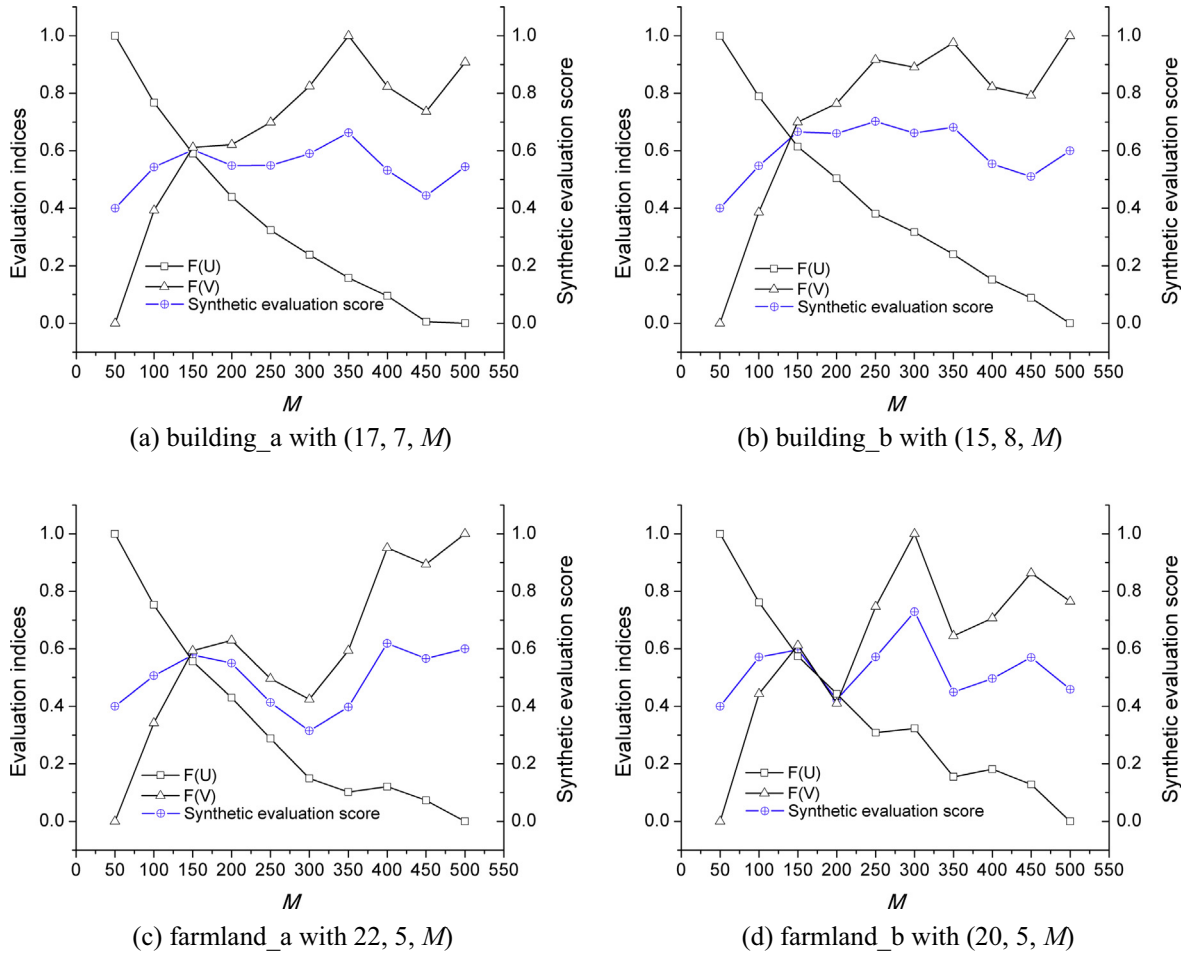


Fig. 13. Segmentation evaluations changing with M of the four experimental images.

Table 4
Peak ranges and optimal M for the four experimental images.

Image	building_a	building_b	farmland_a	farmland_b
Peak range (M)	150–200	100–200	150–200	100–200
Peak point (M)	150	150	150	150
Optimal M by pre-estimation	144	113	121	200

determined by quantitative evaluation is not really visually satisfied because the parcels are over-merged on the right part of the image.

5.2.3. Verification of optimal M

The evaluation of the segmentations is based on merging threshold M from 50 to 500 with step of 50. Other two scale parameters, h_s and h_r, are set as the selected optimal values. The evaluation results are demonstrated in Fig. 13.

Table 4 summarizes the peak ranges meeting the two requirements mentioned above and lists the optimal M determined by the proposed method. Table 4 shows that the optimal M selected by pre-estimation are within the peak ranges (the peak range means it not only contains the peak point but also covers the range within which both $F(U)$ and $F(V)$ are greater than 0.4) and they are very close to the peak points. As a pre-estimation of the optimal scale parameter, the performance of the scale parameter selection method is acceptable.

6. Conclusions

Scale selection in information extraction from high spatial resolution images is important for remote sensing applications. This paper applies spatial statistics to pattern recognition to select the optimal scale parameters in image segmentation for GEOBIA. There are two theoretical foundations involved in this research. The first one is that the essence of determining the optimal scale parameters is the statistical pre-estimation of global and local structure of the original image. The second one is that the range of semivariogram can be deemed as the measurement of similarity between variables, and it can indicate the size of a spatial object.

The scale parameters in multi-scale segmentation are generalized into three aspects: h_s (spatial bandwidth, spatial distance between classes), h_r (spectral bandwidth, spectral difference between classes) and M (merging threshold). Aiming at the selection of the optimal h_s, based on the semivariogram-based method proposed by Ming et al. (2012), this paper proposes that semivariance can be roughly replaced by Average Local Variance (ALV). The ALV-based optimal spatial bandwidth selection method performs as well as the semivariogram method. Moreover, the computation results of ALV are also useful in selecting the optimal spectral bandwidth h_r and merging threshold M. The selection of optimal h_r is based on histogram of ALV and the selection of optimal M is based on simple geometric computation. A series of experimental results verify that the selected scale parameters can achieve a

satisfactory segmentation result with high homogeneity and high heterogeneity. The proposed scale selection method is feasible and effective and it has characteristics as follows.

(1) Other than previous post-evaluation-based scale selection in many OBIA/GEOBIA studies, the proposed statistical method for scale parameter selection is a pre-estimation method. It employs the spatial statistical methods before segmentation to pre-estimate the optimal scale parameters. Comparing to the post-evaluation method, this proposed method can promisingly enhance the efficiency of GEOBIA.
(2) In the experiments, despite the selected optimal scale parameters do not absolutely achieve the highest segmentation evaluation scores in the verification, they quite approximate to the highest evaluation ones. The selected parameters perform very well because they can obtain high homogeneity and high heterogeneity in segmentations, which reduces necessary post-error enhancements and thus somewhat guarantees the accuracy of GEOBIA because high homogeneity and high heterogeneity is one of the necessary requirements for satisfactory accuracy in GEOBIA.
(3) The statistics-based scale parameter selection method is proposed based on mean-shift segmentation. However, this method is not only suitable for mean-shift segmentation. Spatial attribute and spectral attribute are general attributes for spatial data and they are used in most segmentation algorithms. The parameter *M*, merging threshold, is also used in many segmentation algorithms, such as watershed, region-growth and mean-shift algorithm. So this paper does not merely propose a scale parameter selection for mean-shift segmentation algorithm, but also conveys an scale selection idea for GeOBIA.
(4) This scale parameter selection method is independent of the spatial resolution of the image data. So the applicability is theoretically strong for different kind of sensor data.

In addition, there are still some following limitation, which are still subjects of future research.

(1) This paper only gives experiments and verifications for panchromatic images. However, this idea can be extended to the scale selection of multi-spectral image segmentation.
(2) Because the method is still in pilot stage and tremendous computing resources are needed to segment tremendous numbers of objects, the study areas or experimental data are of small sizes, otherwise the computers will breakdown and freeze during the segmentation. This should be considered a limitation especially when dealing with a large dataset (finer resolution data for a relatively large area), so enhancement of the algorithm efficiency is a problem to be solved in the future.
(3) The range of ALvariogram graph is determined by simple threshold control. This method is more suitable for local detailed information extraction from high spatial resolution images. In theory, for an image that contains a complicated scene or nested structure, an 'ideal' object scale does not exist. However, several optimal scales can be selected by a multiple range of multi-function fitting, and they are theoretically appropriate to different landscapes or different kinds of objects with different sizes. Objects from different levels of segmentation (spatially) and of different meanings (ecologically) have to be combined for many applications.
(4) Quantitative verification of this proposed method would actually refer to object based image classification, which inevitably involves a large amount of verification based on GeOBIA classification and thus falls beyond the scope of this paper. However, based on this scale parameter pre-estimation, to what extend could a GeOBIA classification be improved or the workload reduced is still a subject of further research.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (41371347) and "the Fundamental Research Funds for the Central Universities". Many thanks to Erdas Imagine for their providing with sample data. The authors specially appreciate the anonymous reviewers for their helpful comments.

References

Baatz, M., Schäpe, A., 2000. Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation. In: Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT-Symposium Salzburg 2000, Karlsruhe, Herbert Wichmann Verlag, pp. 12–23.

Baatz, M., Shäpe, A., 1999. Object-oriented and multi-scale image analysis in semantic networks. In: Proceeding of the 2nd International Symposium on Operationalization of Remote Sensing, August 16th–20th, Enschede. ITC, Enschede, The Netherlands.

Baatz, M., Benz, U.C., Dehghani, S., Heynen, M., Höltje, A., Hofmann, P., Lingenfelder, I., Mimler, M., Sohlbach, M., Weber, M., Willhauck, G., 2000. ECognition User guide. Definiens-Imaging, Munich, Germany.

Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M., 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS J. Photogramm. Remote Sens. 58, 239–258.

Blaschke, T., 2010. Object based image analysis for remote sensing. ISPRS J. Photogramm. Remote Sens. 65 (1), 2–16.

Blaschke, T., Burnett, C., Pekkarinen, A., 2004. Image segmentation methods for object-based analysis and classification. In: de Jong, S.M., van der Meer, F.D. (Eds.), Remote Sensing and Digital Image Analysis. Including the Spatial Domain. Book Series: Remote Sensing and Digital Image Processing, vol. 5. Kluwer Academic Publishers, Dordrecht, pp. 211–236 (Chapter 12).

Blaschke, T., Lang, S., Hay, G., 2008. Object-based Image Analysis: Spatial Concepts for Knowledge-driven Remote Sensing Applications. Springer-Verlag, Berlin.

Blaschke, T., Hay, G.J., Kelly, M., Lang, S., et al., 2014. Geographic object-based image analysis – towards a new paradigm. ISPRS J. Photogramm. Remote Sens. 87, 180–191.

Cheng, Y., 1995. Mean shift, mode seeking, and clustering. IEEE Trans. Pattern Anal. Mach. Intell. 17 (8), 790–799.

Claudio, R.J., 2007. Combining wavelets and watersheds for robust multiscale image segmentation. Image Vis. Comput. 25 (1), 24–33.

Comaniciu, D., 2003. An algorithm for data-driven bandwidth selection. IEEE Trans. Pattern Anal. Mach. Intell. 25 (2), 281–288.

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

Comaniciu, D., Ramesh, V., Meer, P., 2001. The variable bandwidth mean shift and data-driven scale selection. In: Proc. Eighth Int'l Conf. Computer Vision, vol. I, pp. 438-445.

Definiens AG, 2007. Definiens Developer 7 User Guide, Version 7.0.2.936, 23.

Drăgut, L. et al., 2009. Optimization of scale and parametrization for terrain segmentation: an application to soil-landscape modeling. Comput. Geosci. 35, 1875–1883.

Drăgut, L., Tiede, D., Levick, S.R., 2010. ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. Int. J. Geogr. Inform. Sci. 24 (6), 859–871.

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

Duro, D.C., Franklin, S.E., DubéM, G., 2012. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote Sens. Environ. 118, 259–272.

Eid, M.E., Khaled, M.E., Onsi, H.M., 2010. A proposed multi-scale approach with automatic scale selection for image change detection. Egypt. J. Remote Sens. Space Sci. 13, 1–10.

Einbeck, J., 2011. Bandwidth selection for mean-shift based unsupervised learning techniques: a unified approach via self-coverage. J. Pattern Recognit. Res. 2, 175–192.

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

Fukunaga, K., Hostetler, L., 1975. The estimation of the gradient of a density function, with applications in pattern recognition. IEEE Trans. Inf. Theory 21, 32–40.

Hay, G.J., Castilla, G., 2008. Geographic Object-Based Image Analysis (GEOBIA): a new name for a new discipline. Object-Based Image Anal. Lect. Notes Geoinform. Cartogr. 2008, 75–89. http://dx.doi.org/10.1007/978-3-540-77058-9_4.

Hay, G.J., Castilla, G., Wulder, M.A., Ruiz, J.R., 2005. An automated object-based approach for the multiscale image segmentation of forest scenes. Int. J. Appl. Earth Obs. Geoinf. 7 (4), 339–359.

He, M., Zhang, W., Wang, W., 2009. Optimal segmentation scale model based on object-oriented analysis method. J. Geodesy Geodyn. 29 (1), 106–109.

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

Karl, J.W., Maurer, B.A., 2010. Spatial dependency of predictions from image segmentation: a variogram-based method to determine appropriate scales for producing land-management information. Ecol. Inform. 5, 194–202.

Kim, M., Madden, M., Warner, T., 2008. Estimation of optimal image object size for the segmentation of forest stands with multispectral IKONOS imagery. In: Blaschke, T., Lang, S., Hay, G.J. (Eds.), Object-based Image Analysis–Spatial Concepts for Knowledge Driven Remote Sensing Applications. Springer, Berlin, pp. 291–307.

Li, P., Xiao, X., 2007. Multispectral image segmentation by a multichannel watershed-based approach. Int. J. Remote Sens. 28, 4429–4452.

Li, X.R., Wu, F.C., Hu, Z.Y., 2005. Convergence of a mean shift algorithm. J. Software 16 (3), 365–374.

Li, X.R., Hu, Z.Y., Wu, F.C., 2007. A note on the convergence of the mean shift. Pattern Recogn. 40, 1756–1762.

Ma, Y., 2014. Analysis of Object-Oriented Scale Selection and its Scale Effect Based on e-Cognition. Dissertation Bachelor's Degree in Geoinformatics. China University of Geosciences, Beijing.

Ming, D., Luo, J., Li, L., Song, Z., 2010. Modified local variance based method for selecting the optimal spatial resolution of remote sensing image. In: Proceeding of The 18th International Conference on Geoinformatics, IEEE. doi: http://dx.doi.org/10.1109/GEOINFORMATICS.2010.5567566.

Ming, D., Yang, J., Li, L., Song, Z., 2011. Modified ALV for selecting the optimal spatial resolution and its scale effect on image classification accuracy. Math. Computer Model. 54 (3–4), 1061–1068.

Ming, D., Ci, T., Cai, H., Li, L., Qiao, C., 2012. Semivariogram-based spatial bandwidth selection for remote sensing image segmentation with mean-shift algorithm. IEEE Geosci. Remote Sens. Lett. 9 (5), 813–817.

Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S., Weng, Q., 2011. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. Remote Sens. Environ. 115 (5), 1145–1161.

Anders, N.S., Seijmonsbergen, A.C., Bouten, W., 2011. Segmentation optimization and stratified object-based analysis for semi-automated geomorphological mapping. Remote Sens. Environ. 115 (12), 2976–2985.

Openshaw, S., 1984. The modifiable areal unit problem. Concepts and Techniques in Modern Geography, vol. 38. GeoBooks, Norwich, England.

Park, J.H., Lee, G.S., Park, S.Y., 2009. Color image segmentation using adaptive mean shift and statistical model-based methods. Comput. Math. Appl. 57, 970–980.

Peña-Barragán, J.M., Ngugi, M.K., Plant, R.E., Six, J., 2011. Object-based crop identification using multiple vegetation indices, textural features and crop phenology. Remote Sens. Environ. 115, 1301–1316.

Polat, E., Ozden, M., 2006. A nonparametric adaptive tracking algorithm based on multiple feature distributions. IEEE Trans. Multimedia 8 (6), 1156–1163.

Rao, S., Martins, A.D.M., Príncipe, J.C., 2009. Mean shift: an information theoretic perspective. Pattern Recogn. Lett. 30, 222–230.

Scheunders, P., Sijbers, J., 2002. Multiscale watershed segmentation of multivalued images. In: 16th Proceedings of International Conference on Pattern Recognition (ICPR'02), vol. 3, p. 30855.

Sun, J., Xu, Z., 2010. Scale selection for anisotropic diffusion filter by Markov random field model. Pattern Recogn. 43, 2630–2645.

Tan, Q., Liu, Z., Shen, W., 2007. An algorithm for object_oriented multi_scale remote sensing image segmentation. J. Beijing Jiaotong Univ. 31 (4), 11–115.

Tian, J., Chen, D.M., 2007. Optimization in multi-scale segmentation of high-resolution satellite images for artificial feature recognition. Int. J. Remote Sens. 28 (20), 4625–4644.

Vieira, M.A., Formaggio, A.R., Rennó, C.D., Atzberger, C., Aguiar, D.A., Mello, M.P., 2012. Object based image analysis and data mining applied to a remotely sensed landsat time-series to map sugarcane over large areas. Remote Sens. Environ. 123, 553–562.

Vincent, L., Soille, P., 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. IEEE Trans. Pattern Anal. Mach. Intell. 13 (6), 583–598.

Vojir, T., Noskova, J., Matas, J., 2014. Robust scale-adaptive mean-shift for tracking. Pattern Recognit. Lett. http://dx.doi.org/10.1016/j.patrec.2014.03.025.

Wang, G., Gertner, G., Xiao, X., et al., 2001. Appropriate plot size and spatial resolution for mapping multiple vegetation types. Photogramm. Eng. Remote Sens. 67 (5), 575–584.

Woodcock, C.E., Strahler, A.H., 1987. The factor of scale in remote sensing. Remote Sens. Environ. 21, 311–332.

Wu, K.-L., Yang, M.-S., 2007. Mean shift-based clustering. Pattern Recogn. 40, 3035–3052.

Yang, G., Pu, R., Zhang, J., Zhao, C., Feng, H., Wang, J., 2013. Remote sensing of seasonal variability of fractional vegetation cover and its object-based spatial pattern analysis over mountain areas. ISPRS J. Photogramm. Remote Sens. 77, 79–93.

Zhang, R., 2005. Spatial Variability Theory and its Applications. Science Press, Beijing, China.

Zhang, X., Cui, Y., Li, D., Liu, X., Zhang, F., 2012. An adaptive mean shift clustering algorithm based on locality-sensitive hashing. Optik 123, 1891–1894.

Zhao, M., Li, F., Tang, G., 2012. Optimal scale selection for DEM based slope segmentation in the loess plateau. Int. J. Geosci. 3, 37–43. http://dx.doi.org/10.4236/ijg.2012.31005.

Zheng, L., Shi, D., Zhang, J., 2010. Segmentation of green vegetation of crop canopy images based on mean shift and Fisher linear discriminant. Pattern Recogn. Lett. 31, 920–925.

Zhou, H., Li, X., Schaefer, G., Celebi, M.E., Miller, P., 2013. Mean shift based gradient vector flow for image segmentation. Comput. Vis. Image Underst. 117, 1004–1016.