Line-based image segmentation method: a new approach to segment VHSR remote sensing images automatically

Jaime Lopez , John W. Branch & Gang Chen

Published online: 06 Dec 2019.

Submit your article to this journal ⌕

View related articles ⌕

View Crossmark data ⌕

Taylor & Francis
Taylor & Francis Group

Line-based image segmentation method: a new approach to segment VHSR remote sensing images automatically

Jaime Lopez [iD][a], John W. Branch [iD][b] and Gang Chen [iD][c]

[a]Department of Engineering, Universidad del Tolima, Ibague, Colombia; [b]Department of Computer and Decision Sciences, Faculty of Mines, Universidad Nacional de Colombia, Medellin, Colombia; [c]Department of Geography and Earth Sciences, University of North Carolina at Charlotte, Charlotte, NC, USA

ABSTRACT

There exist different approaches for segmenting Very High Spatial Resolution (VHSR) remote sensing imagery with competitive performance, including object-based (e.g. Multiresolution), gradient-based (e.g. Watershed), and clustering-based (e.g. k-means) segmentation. However, they have a strong dependence on human assistance for tuning the required parameters (e.g. scale value, clusters number or tolerance thresholds), usually following a trial-and-error methodology that becomes tedious, hardly reproducible or transferable to other images, affecting negatively the methods' robustness and efficiency. In this communication, we propose a novel method denominated Line-based segmentation (LBS) that automatically segments VHSR remote sensing imagery through a data-driven approach, bypassing the parameters' definition by experts (i.e. region growing´s seeds and thresholds). The proposed algorithm offers flexibility and accuracy to segment regions with varying sizes and shapes, tested on different VHSR images, including multispectral images (WorldView-3, GeoEywe-1, Ikonos, QuickBird and SkySat), RGB aerial image (NAIP) and panchromatic image (Ikonos). The results revealed the LBS method shows a competitive performance compared against two well-known segmentation approaches, but without user intervention and generating consistent and repeatable segmentation results following an automatic fashion.

Introduction

Efficient image segmentation is one of the most critical steps in automated remote sensing image processing (Dey, Zhang, & Zhong, 2010; Han, Zhang, Cheng, Guo & Ren, 2015), because it aims to find meaningful regions of interest (i.e. discernible objects) that becomes fundamental for particular analysis in later stages, such as object-based image classification (Blaschke, 2005; Lang, 2008; Zhang, 2015). Technically, segmentation is defined as an image partition process, where pixels (i.e. picture elements) are assigned to segments or regions (groups of pixels) (Pal & Pal, 1993). Over the past decades, this topic has received considerable attention using low or medium spatial resolution imagery. Recently, with the higher availability of Very High Spatial Resolution (VHSR) remote sensing imagery, a variety of algorithms have been developed to extract fine-scale regions, representing some discernible objects such as individual trees (Hirschmugl, Ofner, Raggam, & Schardt, 2007), shadows (Adeline, Chen, Briottet, Pang, & Paparoditis, 2013), buildings (Baltsavias, 2004; Belgiu & Drǎguţ, 2014; Hofmann, 2001), roads (Kumar, Singh, Raju, & Krishnamurthy, 2014; Nobrega, O'Hara, & Quintanilha, 2008) and vehicles (Razakarivony & Jurie, 2016) to name but a few.

Automatic segmentation

While many steps of remote sensing image processing can be carried out with minimum human intervention, automation of image segmentation still presents several problematic aspects. First of all, the process usually requires human participation (i.e. experts) regarding the most appropriate parameter settings based on the specific image and ground characteristics, demanding the use of a trial-and-error methodology (Hay, Blaschke, Marceau, & Bouchard, 2003; Myint, Gober, Brazel, Grossman-Clarke, & Weng, 2011; Yu et al., 2006). This results in a time-consuming process and a bottleneck for the analysis of large volumes of data (Cheng & Han, 2016). Second, although each VHSR image scene usually contains regions with distinct characteristics (e.g. size, shape or texture), also exists different regions comparting similar appearances making it difficult to properly determine their boundaries (Beneš & Zitová, 2015; Chen, Hay, & St-Onge, 2012; Cheng & Han, 2016; Gurcan et al., 2009).

Considerable research efforts have been devoted to automatizing image segmentation, many of which were originated from medical image analysis or computer vision, such as extracting retinal vessels (Zhao, Wang, Wang, & Shih, 2014), breast tumor

CONTACT Jaime Lopez ✉ jlopez@ut.edu.co 🖃 Department of Engineering, Universidad del Tolima, Ibague, Ibague 730006299, Colombia

segmentation (Rouhi, Jafari, Kasaei, & Keshavarzian, 2015) or handwritten text recognition (Saeed & Albakoor, 2009). Given the unique characteristics of remotely sensed imagery, researchers have developed many methods following a semi-automatic approach aiming to delineate image objects with minimum human intervention. For example, the Region-based Image Segmentation (RISA) method relies on clustering k-means algorithm to define seeds for growing regions (Wang, Jensen, & Im, 2010), but it still requires the definition of clusters' number and stopping criteria by a human user. Zhang and Maxwell (2006) proposed a fuzzy segmentation, which applies fuzzy theory to assist parameter selection. This algorithm requires an initial image segmentation to train a fuzzy system, and extensive testing to improve its robustness working on different types of imagery. Byun, Kim, Lee, and Kim (2011) proposed an approach to segment VHSR images using a region growing algorithm that automatically selects seeds following a block-based method, where block size needs to be defined a priori. To facilitate automatic segmentation, Akcay and Aksoy (2008) combined image's spectral and structural information along with morphological operations to build a hierarchical tree for the segmentation process. Although this algorithm showed promising results, it requires experts to define structural elements and clusters based on the objects of interest. To date, user intervention is needed for most of the algorithms to properly define segmentation parameters.

Regions' scale

Actually, multiscale analysis is a hot topic in image segmentation (Blaschke et al., 2014; Zhang, Xiao, Feng, Feng, & Ye, 2015; Zhou, Li, Feng, Zhang, & Hu, 2017). Since the proposal of Witkin (1984) who defined a concise description of a signal scale through a tree involving hierarchical relations between all observation scales; and Marr and Hildreth (1980) proposed a two-step approach to detect intensity changes at different scales, and afterwards localize them spatially, stating that identified segments are not independent, allowing to deduce rules to describe an image (raw primal sketch), resulting in a hierarchy of descriptions covering a range of scales (full primal sketch). Tabb and Ahuja (1997) presented a new concept of scale that represents image structures at different scales, combining scale and structure explicitly, making it possible to detect edges and regions. Baatz and Schape (2000) proposed its iconic and now widely used multiresolution analysis, that opened a new way to analyse images following an object-based approach (Benz, Hofmann, Willhauck, Lingenfelder, & Heynen, 2004), that is evolving into a paradigm (Blaschke et al., 2014), allowing the addition of spatial relations between objects and expert knowledge into the process (Hay & Castilla, 2006). Although the scale selection still remains a big challenge to overcome, it often depends on subjective trial-and-error methods (Blaschke, 2010; Zhou et al., 2017).

Arbelaez, Maire, Fowlkes, and Malik (2011) combined contours and segmentation capturing multiscale information (three scales), without showing any significant improvement on the datasets used. Their approach consists of transforming contour detector into a hierarchical region tree, getting an image segmentation as a result. They recognized that contour detection and segmentation are related, but not identical because contour detectors offer no guarantee to produce closed contours and hence do not necessarily provide a partition of the image into regions, as image segmentation does.

Drăguţ, Tiede, and Levick (2010) presented a very promising method to automate the selection of the best scale parameter value for multiresolution segmentation (MRS) involving the intrasegment homogeneity defining a three-level hierarchy to segment regions of different sizes through a tool called ESP (Estimation of Scale Parameters). This represents an improvement for the original approach based on a user-defined scale determination (Baatz & Schape, 2000), that avoid the automatization of this process. Also, Yang, He, and Weng (2015) made a refinement of local peak method (implemented by ESP tool), adding the intersegment heterogeneity that improves the scale determination by 17%. Later, Dragut, Csillik, Eisank, and Tiede (2014) extend the ESP tool functionality, to determine three scale levels using multiple image bands.

This research aims to propose a new segmentation method that departs from previously described multi-scale approaches able for monitoring, modeling, and managing complex landscapes through linking different scales hierarchically (Burnett & Blaschke, 2003) present in VHSR remote sensing images. Although multiscale analysis allows different windows of perception, it depends heavily on user assistance, which motivates us to focus on automatic region segmentation, that could benefit other segmentation approaches.

Specifically, our approach employs an entropy function to select a particular line from the image under analysis. Afterwards, an optimization process is applied to this line to determine the best scale value necessary for the automatic segmentation of the image, avoiding the intervention of experts for the parameters' tuning while still achieving a competitive performance. Although this method could be considered as single-scale, it is important to note the fact it allows the segmentation of different sized regions, starting from the small discernible regions in the image until large regions with different sizes, shapes and spatial locations.

This communication starts with a detailed description of the proposed method (Section 2). Then, its

performance was tested on different VHSR images comparing it against two well-known segmentation algorithms, under a qualitative and quantitative evaluation (Section 3). The obtained results are explained, compared and discussed in Section 4, and the conclusions are presented at the end (Section 5).

Materials and methods

Sample imagery

All selected images have very high spatial resolutions, ranging between 0.8 and 3.20 m on multispectral, visible and panchromatic modes. They were acquired from different sensors with proven performance, including images from Maxar's Satellite Constellation (Ikonos, GeoEye, WorldView), Planet (SkySat) and the U.S. National Agriculture Imagery Program (NAIP) (Figure 1). These images present areas with urban/rural mixed landscapes and a reasonable diversity of region types and sizes, such as roads, buildings, lakes, parks, etc. Table 1 highlights the images' characteristics. A complete VHSR remote sensing scene usually has a large size, so one often employed strategy is to test new methods on a small section of it (Troya-Galvis, Gançarski, & Berti-Équille, 2003) to reduce processing times and to alleviate the high labor intensity required for the manual segmentation during the extraction of ground truth necessary for comparison and evaluation purposes.

Image processing

The proposed image segmentation method is based on a region growing (RG) idea, but adopting a data-driven approach that works applying an entropy function on transversal lines over the image under analysis, reason which motivated its name as Line-based segmentation (LBS). It is composed of three main steps: (a) image preprocessing; (b) identification of segmentation parameters – seeds and thresholds; and (c) line (LRG) and random region growing (RRG) segmentation. The flowchart is shown in Figure 2.

Actually, recent advances in satellite and sensor technologies have enabled considerable improvements in the radiometric resolution of the VHSR remote sensing images, from 8bit to 16 or 32 bit, but images with higher radiometric resolution require more radiometric values to record, spreading over more values (Ose, Corpetti, & Demagistri, 2016), which means larger image size. Usually, the most common radiometric image resolution is 8-bit, because the extraction of image information not always adheres to the highest possible radiometric resolution (Rama-Rao, Garg, & Ghosh, 2006; Verde, Mallinis, Tsakiri-Strati, Georgiadis, & Patias, 2018), having this into account, to keep a balance between computational efficiency and results accuracy, the multispectral and visible images were 8-bit, while panchromatic was 16-bit.

The proposed algorithm operates on grayscale imagery, meaning a color to grayscale conversion was needed for multispectral and visible images, an

Figure 1. Selected VHSR images showing natural and man-made discernible objects: (a) Ikonos-Pan, (b) NAIP, (c) WorldView3, (d) GeoEye1, (d) Ikonos, (e) QuickBird and (f) SkySat.

Table 1. Description of VHSR satellite and aerial imagery used.

| Image	Type	Spatial resolution (metres/pixel)	Radiometric resolution	Dimensions (pixel × pixel)	Format
Ikonos-Pan	Panchromatic	0.81	16-bit	1000 × 1000	GeoTIFF
NAIP	RGB	1.00	8-bit	500 × 500	Mr.SID
WorldView3	Multispectral	1.24	8-bit	500 × 500	GeoTIFF
GeoEye1	Multispectral	1.65	8-bit	500 × 500	GeoTIFF
SkySat	Multispectral	2.00	8-bit	500 × 500	GeoTIFF
QuickBird	Multispectral	2.62	8-bit	500 × 500	GeoTIFF
Ikonos-MS	Multispectral	3.20	8-bit	500 × 500	GeoTIFF
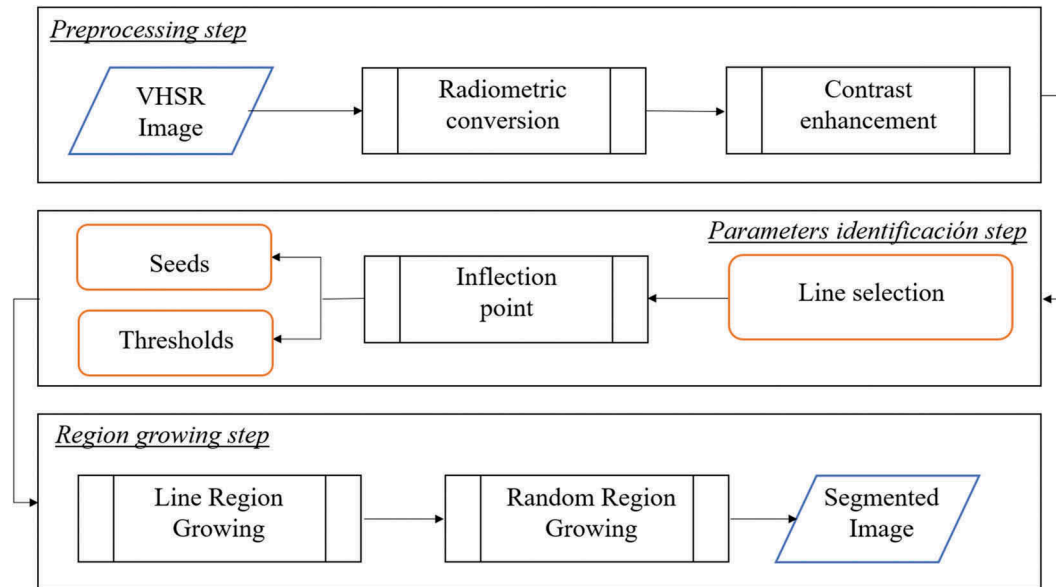
Figure 2. Main components of our line-based image segmentation approach.

operation widely used as a preprocessing step in image processing in different domains such as medical imaging, satellite imaging and agriculture/environment real-scene (Ma, Zhao, Zeng, & Wang, 2015; Nafchi, Shahkolaei, Hedjam, & Cheriet, 2017), mainly to reduce processing times and computer requirements. Also, the color often introduces some redundancies (Acharya & Ray, 2005) or unnecessary information into the process (Kanan & Cottrell, 2012). There exist several procedures to convert color to grayscale images, that can be categorized into global, local, and hybrid (Nafchi et al., 2017). The method we chose belongs to the global category, known as luma, that forms a weighted average, being more sophisticated than the average method, known as Intensity, because the former takes account that humans do not perceive all colors equally. Usually, this linear luminance is calculated as a weighted sum of the three linear-intensity values, representing the brightness of a panchromatic monochrome image. It combines the red, green, and blue signals in proportion to the human eye's sensitivity to them (Pohl & Van Genderen, 1998). The conversion applied here was a version of luminance through the next formula (Kanan & Cottrell, 2012):

$$g(x, y) = 0.2125 [R(x, y)] + 0.7154 [G(x, y)] + 0.0721 [B(x, y)] \quad (1)$$

where R(x,y), G(x,y), and B(x,y) represents gray level (intensity) values at red, green and blue spectral bands, respectively, and the coefficients represent specific spectral weighting to compute the true CIE luminance of any RGB image following the Rec.709 Color space (ITU-R, 1990). This is a default method implemented by well-known image processing software such as GIMP (https://www.gimp.org), Matlab (https://www.mathworks.com) and ImageJ (https://

imagej.net), sometimes with small coefficients changes (Bosch, Zisserman, & Munoz, 2007, Dec).

This color-to-grayscale conversion is required to preserve the salient features of the color images, such as brightness, contrast and structure of the color image (Günes, Kalkan, & Durmu, 2015), and at the same time, retaining as much information about the original color image as possible, while simultaneously producing perceptually plausible grayscale results (Cadík, 2008).

Although the majority of the VHSR images presents four bands, including the near-infrared band, this band was not used because of: (a) the conversion function from color to grayscale image does not include NIR band, (b) RGB images showed satisfactory performance compared with CIR (Composite Infrared) and MS (multispectral) images (Zheng et al., 2018) and (c) some authors indicate NIR band tends to produce larger regions in comparison with RGB images, as is described below.

Torres-Sánchez, López-Granados and Peña (2015) compared 8-bit UAV images from RGB and multispectral camera, applying MRS segmentation with the same weight for each band. They considered scale as the most important factor for controlling region size, finding that for each scale, RGBNir tends to produce larger regions than RGB images, on three different crops, as scale value increases. It is important to note, their main objective was to accurately discriminate between vegetation and bare soil coverage types. It is known that NIR band is very useful to differentiate vegetation, but it was shown to affect regions' sizes, which could reduce the benefit of better image spatial resolution. Another research (Drăguţ et al., 2010), corroborated this when their ESP tool found larger scales for NIR image than Red channel favoring the detection of larger regions, which could be harmful when used on VHSR images. Yang, Li, and He (2014)

indicated that man-made materials, such as metal, asphalt and concrete, which have minimal variability in the NIR region, may feature strong variability in the visible light region on account of painting, special coatings or even age on VHSR images (QuickBird and WorldView3), allowing a better recognition for diverse land cover regions in different urban landscapes; while Van Etten (2018) compared grayscale, RGB and MS images without finding any significant difference between them, all yielding high F1 scores (F1 > 0.80) for detection of regions at different scales (0.15–3.0 m) on VHSR satellite images.

An image contrast enhancement is further applied to improve region perceptibility, improving brightness differences among regions, which was completed through a remapping of grayscale values to cover the complete images' dynamic range. This is a linear transformation (Gonzalez & Woods, 2008) using the following function:

$$g'(x, y) = [g(x, y) - GL(min)/GL(max) - GL(min)] \times 255 \qquad (2)$$

where GL(max) and GL(min) represent the maximum and minimum gray values of the image, respectively, and g(x,y) is the pixel's gray value at coordinate (x,y).

Identification of segmentation parameters – seeds and stopping threshold

The traditional region growing (RG) method requires the identification of a seeds' set that correspond to starting points of growing regions, and a threshold as the stopping criterion for every region. Liu, Li, and Wang (2015) indicate those seeds can be determined using manual selection, gradient-based or random approaches. Given the fact that the first two methods require a certain level of human assistance, we decided to choose the random-option to pursue our objective of automatic image processing. Sánchez, Martínez, and Arquero (2015) evaluated four strategies for automatic seed selection for a region growing algorithm, including random selection, edge-based selection, distance-based selection and a combination of edge and distance-based. They found the edge-based seed selection was the better strategy. We proposed a new method based on the entropy theory in order to define the stopping thresholds. This permits to identify the maximum variability between pixels' values in the image. Shannon cited by (MacKay, 2003) formulated his entropy formula with the function defined as below:

$$H(X) = - \sum [P(x) \log_2 P(x)] \qquad (3)$$

where H is used to denote entropy, P is the probability, and x represents each image grayscale value.

First, the image under analysis is fully scanned using three different straight line patterns, such as rotational, horizontal and vertical, but for illustration purposes, only the main lines are shown (Figure 3). The entropy value from each line is stored. Then, the line with maximum entropy value is selected as this line contains the highest variability in gray levels of the VHSR image under study (Singh & Singh, 2008) likely representing the largest amount of regions vital for later analysis. Figure 4 shows the selected lines in each image.

Next, a vector is created with all the gray values contained inside the selected line keeping the same order. Then, a merging process begins between neighboring pixels depending on if the gray values fall inside a particular threshold, starting from one and increasing up to a maximum predefined value. This maximum value was chosen according to the image's radiometric type (8-bit or 16-bit) to avoid values which could give too few regions (two or one final regions). With every unitary increment, the new gray levels are stored (labels) and a graph is built using both arrays (increments vs. gray levels), applying an optimization process, specifically the iterative L-method approach proposed by Salvador and Chan (2004) to find its respective inflection point. This point represents the turning point where the transition occurs from small to large regions that could be interpreted as the point where the smaller and discernible regions of the image under study can be detected. Following this strategy, allowed us to bypass the two most critical elements of RG algorithm (Fan, Zeng, Body, & Hacid, 2005), such as the determination of thresholds and selection of seeds.
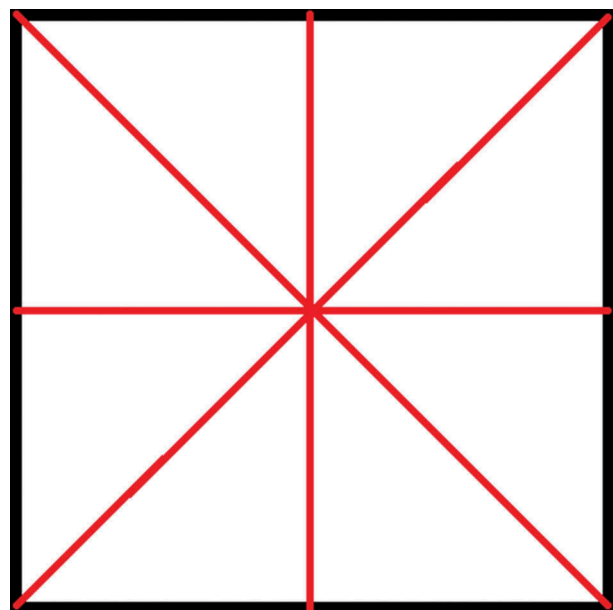

Figure 3. Three different scanning patterns (horizontally, vertically, and diagonally) used to calculate entropy values.
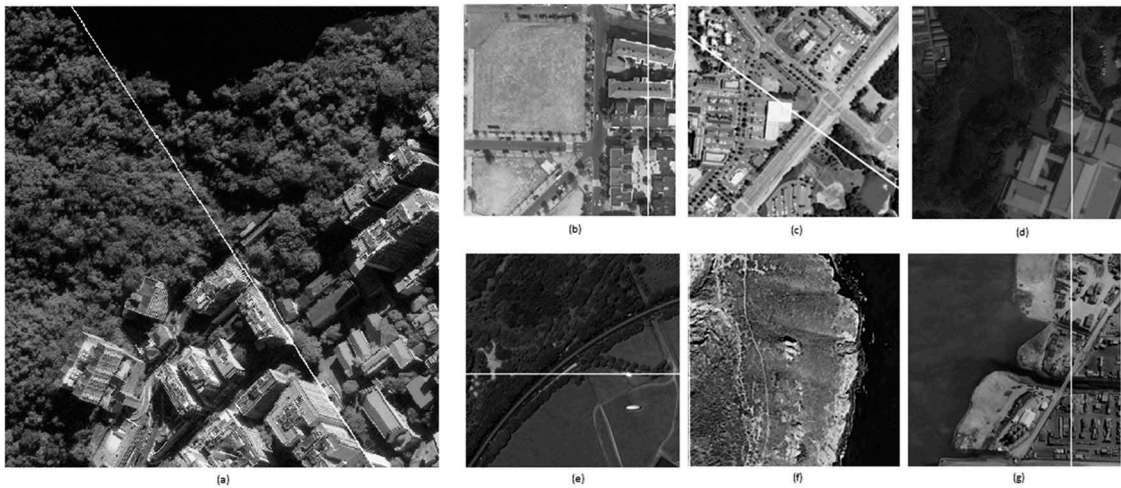
Figure 4. Illustration of selected lines in each test image showing different directions: (a) Ikonos-Pan, (b) NAIP, (c) WorldView3, (d) GeoEye1, (e) QuickBird, (f) SkySat and (g) Ikonos.

Figure 5 illustrates the aforementioned steps with a simple example. The process starts with the highest entropy line extracted from a 3-bit toy image (8 different gray levels), with ascending gray levels from 0 (black) to 7 (white). To be more specific, these numeric values represent the initial vector (Figure 5(a)) and illustrates how the gray values change through merging operations according to the actual threshold (i.e. numbers on the left), and the gray levels that were stored (numbers on the right). Figure 5(b) depicts the graph built using both number arrays, locating the inflection point at coordinates (3,2), corresponding to the selected threshold value and final regions, respectively (Figure 5(c)).

The regions' number identified by the previously described optimization process was used to define the scale in every image under study, through the arithmetic division between the vector length and the number of regions detected. Figure 6, shows the scale values obtained for each image. There are some important aspects to note from it. First, it could be inferred that a better image's spatial resolution tends to produce higher entropy values, but it does not

always seem to produce larger scale values, because scale depends strongly on the existence of large areas with similar gray values inside the image, such as grasslands or waterbodies. Second, the images' size seems not to influence the scale values, as it can be verified with the Ikonos-Pan image, which has a double size compared to others, getting a very small-scale value (1.0), but not too far from the small-sized images; by the contrary, the GeoEye image has the largest scale value (7.14), mainly obtained for the influence of a large region (roof) with a constant pixel value, reducing the number of regions detected (only 70). Third, it could be found an inverse relationship between entropy values and scale, because a greater entropy value means more regions detected, causing the scale decrease.

Region growing

The region growing process is split into two components: Line Region Growing (LRG) and Random Region Growing (RRG). Both components work
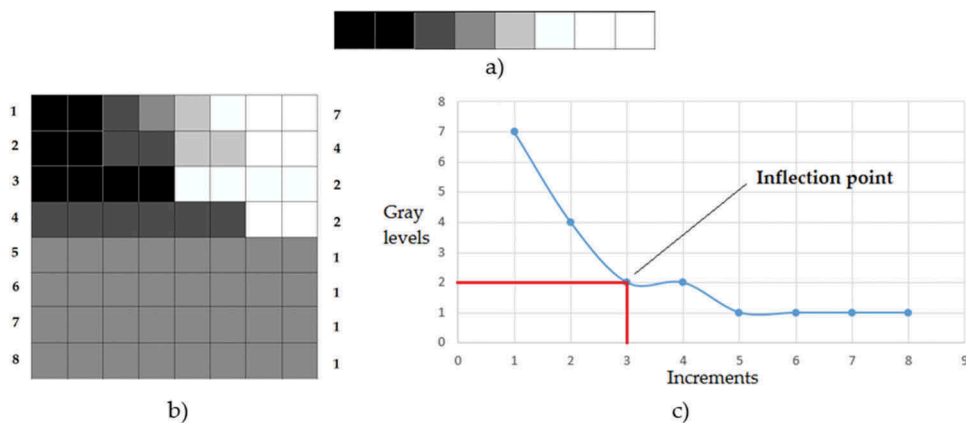

Figure 5. Sequential steps from (a) to (c), starting with gray values taken from selected line to conform the starting vector that changes through merging operations to find the corresponding threshold and seeds required for the region growing process.
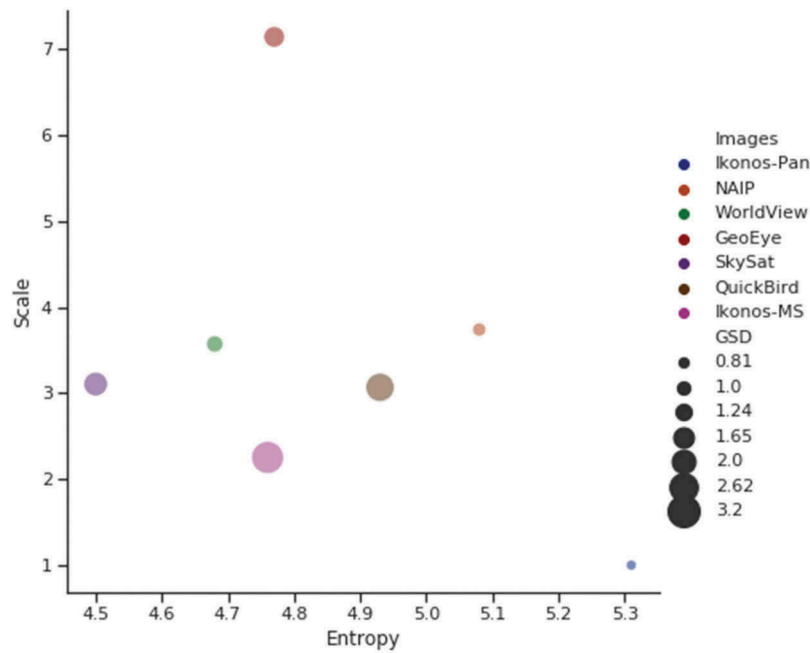
Figure 6. The entropy values and their respective scales for each VHSR image along with their spatial resolutions (GSD).

similarly, adding the pixel from its immediate neighbours (up, down, left and right), where its gray value is closest to the seed value. However, the first component (LRG) grow regions along the selected line (Section 3.2), while the second one (RRG) grow regions randomly.

The LRG starts growing regions following a sequential pattern using the threshold and seeds found in the previous section. The process starts with the first left pixel on the line, growing until the first region is extracted. Next, it continues with the second seed, growing until the second region is obtained. The same process continues until reaching the line's last pixel (Figure 7(a)). Once all the regions on the line are delineated, the algorithm adopts a random pattern (RRG) selecting seeds randomly for growing regions (Figure 7(b)), until no more regions are found (Figure 7(c)). All regions' boundaries are displayed in Figure 7 (d), indicating the completion of segmentation. All previous steps are illustrated in Figure 7

Figure 8 describes the main steps of the proposed region growing algorithms for LRG and RRG, which differs only in the way they generate seeds.

Results

Segmentation metrics

The metrics employed to assess the segmentation performance and make an objective comparison between our method against other segmentation approaches are known as supervised metrics (Beneš & Zitová, 2015; Zhang, Fritts, & Goldman, 2008), also called discrepancy methods by Cardoso and Corte-Real (2005) which require a reference segmentation (also known as ground truth or gold standard) as the basis. These metrics include different approaches such as statistical, positional and geometrical (Table 2), providing a solid and wide evaluation criteria to reduce
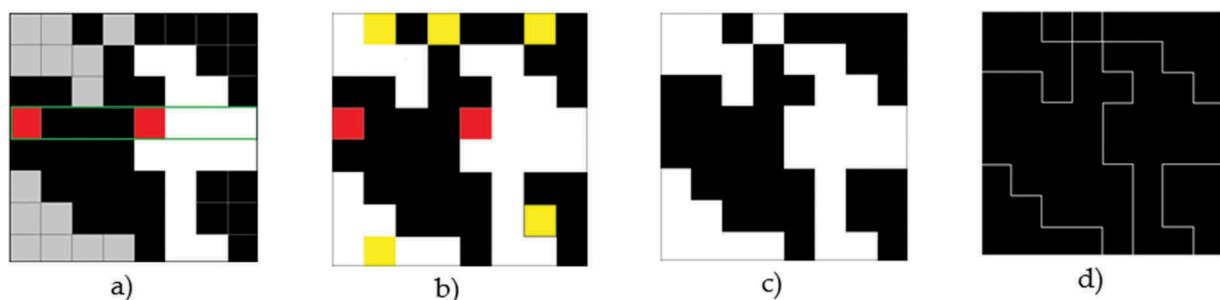

Figure 7. (a) LRG starting from red seeds inside chosen line (green rectangle), creating two objects (black and white); (b) RRG, grows using yellow seeds; (c) All created objects belong to only two classes (black and white); (d) Final segmentation result (white boundaries).

Algorithm 1 Region growing algorithm

Require: *image, contours, visited, label, seed, threshold*

> gray_value = *image* [*seed*]
>
> actual_position = *seed*
>
> while gray_value> *threshold* .Low and gray_value <*threshold* .High
>> for i in four_orientations:
>>> if *contours* [i] == 0:
>>>> append(contours, actual_position)
>>>>
>>>> *contours* [actual_position] = *label*
>>>>
>>>> *visited* [actual_position] = True
>>
>> gray_value = minimum_distance(*contours*)
>>
>> current_position = *contours* [gray_value.index]
>>
>> delete(*contours* [gray_value.index])
>
> **return** *contours, visited*

Figure 8. Description of region growing algorithm implemented.

Table 2. General description of chosen metrics.

Metric	Range	Optimum value	Author	Year
Rand index	[0,1]	1	Rand	1971
Position discrepancy	[0, …]	0	Montagui et al.	2013
Area fit index	[−1,0,+1]	0	Lucieer	2004
Hamming distance	[0,1]	0	Hamming	1950
Jaccard distance	[0,1]	0	Jaccard-Needham	1908
Fowlkes-Mallows index	[0,1]	1	Fowlkes-Mallows	1983
Overall Accuracy	[0,1]	1	Congalton	1991

any possibility to obtain distorted results caused by insufficient metrics (Beneš & Zitová, 2015).

In our case, the reference segmentation was obtained by averaging segmentation results realized by five human experts, to get a ground truth closer to an ideal segmentation. This ground truth was not created just for one particular type of regions (like buildings or trees), but for all regions located along the selected line. This approach allows us to use a larger amount and diversity of regions without any subjective regions' selection.

Figure 9 provides an example of performance evaluation, where an image includes a rounded region (having a black boundary) over a white background with an arbitrary horizontal maximum entropy line (green rectangle) (Figure 9(a)), red points represent a particular segmentation of the proposed algorithm, and yellow points represent the reference segmentation (Figure 9(b)). Each segmentation is transformed into a binary vector, where '1′ means "boundary" and '0′ is the region area (Figure 9(c)), the final binary

vector resulting from an AND logic operation (intersection) between two previous vectors, with 1 and 0 representing agreements and non-agreements, respectively (Figure 9(d)).

Qualitative evaluation

The qualitative evaluation process involved the analysis of those visual features that could be easily recognized by an expert human. In this case, we selected three different attributes; boundary fidelity, area similarity and presence of artifacts inside every region. For these tests, we choose meaningful regions representing discernible regions, such as roof, street or trees.

Ideal or perfect segmentations rarely occur because any segmentation process usually produces artifacts that do not belong to some particular region, or sometimes, the regions' boundary presents a minor displacement from real border which could not be appreciated visually. To avoid a biased evaluation it was established three region sizes' categories for region delineation according to their boundary fidelity, area similarity and presence of artifacts, using large (lake and building), medium (grass field and forested area) and small regions (trees and streets) (Figure 10).

We found the next peculiarities within the large regions (lake and building). The LBS offered a better area similarity of both, while the WS method obtained the best boundary fidelity of the building region and the MRS and LBS methods manifested some artifacts over it. Within medium regions (forested area and grass field), all methods oversegmented the forested area possibly by its rough texture, but not showed with
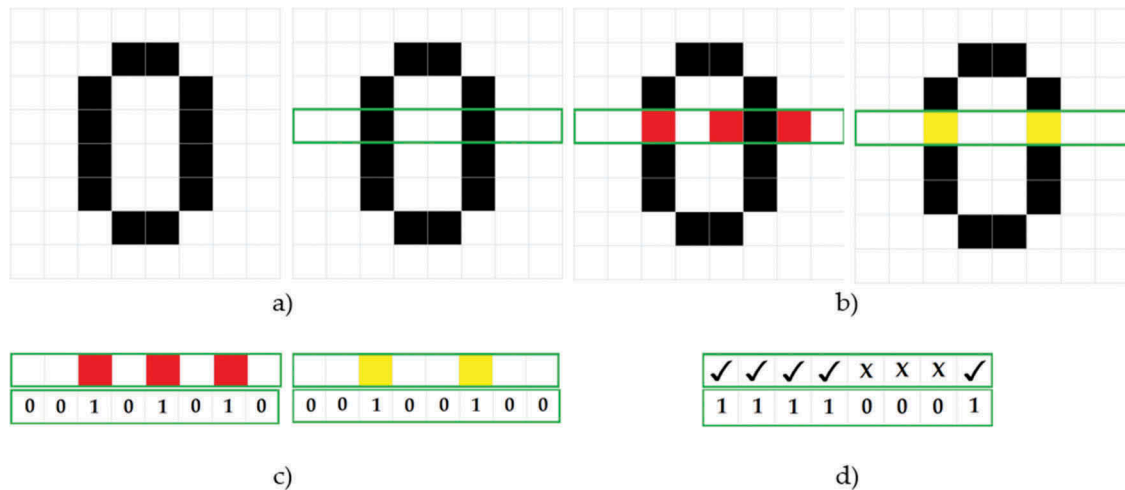
Figure 9. Performance evaluation through vector intersection using AND logic operation on a small toy image (8x8 pixels), with a final performance of 62.5%.

the grass field, where all methods gave a very good boundary delineation, with no presence of artifacts and a good area similarity. About the small regions (trees and car), the LBS merge the trees with its shallows, as one region, and also created a few artifacts inside the segments. With the car region, all methods segmented it well, although with a little difference in the boundary fidelity and area similarity.

In general, all three segmentation methods produced a high boundary fidelity, although it was appreciated that sometimes the boundaries were traced by the external border, and other times by internal borders, but without detecting any particular pattern for any of the methods tested, mostly produced by the particular parameters setting employed for each image. In relation with the area similarity feature, it was found that MRS behave better, especially when taking into account the different categories of regions' sizes, indicating that the ESP tool was a good scale selector. Although the proposed LBS method is single-scale, it showed good performance segmenting regions of any size. The presence of artifacts did not show any particular difference between these three methods, appearing in any place, as in the boundary or inside/outside of regions.

Also, to check if the proposed method behaves differently delineating regions on the panchromatic band, it was chosen three regions with different sizes, such as small (balconies), medium (street) and large (roof), as can be seen in Figure 11. The segmentations let us affirm the proposed method can recognize very small regions like balconies and can delineate them well, although it tends to produce some artifacts inside as the region gets larger. The MRS behaves very well under all regions' sizes tested, giving a well-defined boundary, while the WS method has some trouble to

find a parameters' set that can obtain a good delineation of the regions under study, leading to a regions' oversegmentation (Figure 11).

Next, it will be described the quantitative evaluation to avoid any subjective evaluation that could be biased by an experts' own judgement, and that could be controversial for some other evaluators.

Quantitative evaluation

Seven metrics were chosen to make a quantitative evaluation involving a wide spectrum of metrics (Table 3), including Rand index (Rand, 1971), Positional discrepancy (Montaghi, Larsen, & Greve, 2013), Area fit index (Lucieer, 2004), Hamming distance (Norouzi, Fleet, & Salakhutdinov, 2012), Jaccard distance (Real & Vargas, 1996), Fowlkes-Mallows index (Fowlkes & Mallows, 1983) and overall accuracy, one of the most commonly used measures for performance evaluation (Congalton, 1991). Also, we performed a *t*-test for means of two independent samples from descriptive statistics, which is a two-sided test for the null hypothesis that two independent samples have identical average values.

Specifically, Rand index is related to accuracy or similarity between clusters, using agreements and disagreements to calculate its value. It penalizes both false positive and false negative during clustering allowing us to measure the correspondence between edges, but at the same time, showing the correspondence between non-edges. This metric did not show a good performance for the LBS method, this could be partially explained because that under large and medium scales have less coincidences between non-boundary pixels of regions producing higher values for false positives and false negatives, and consequently reducing its Rand index substantially.
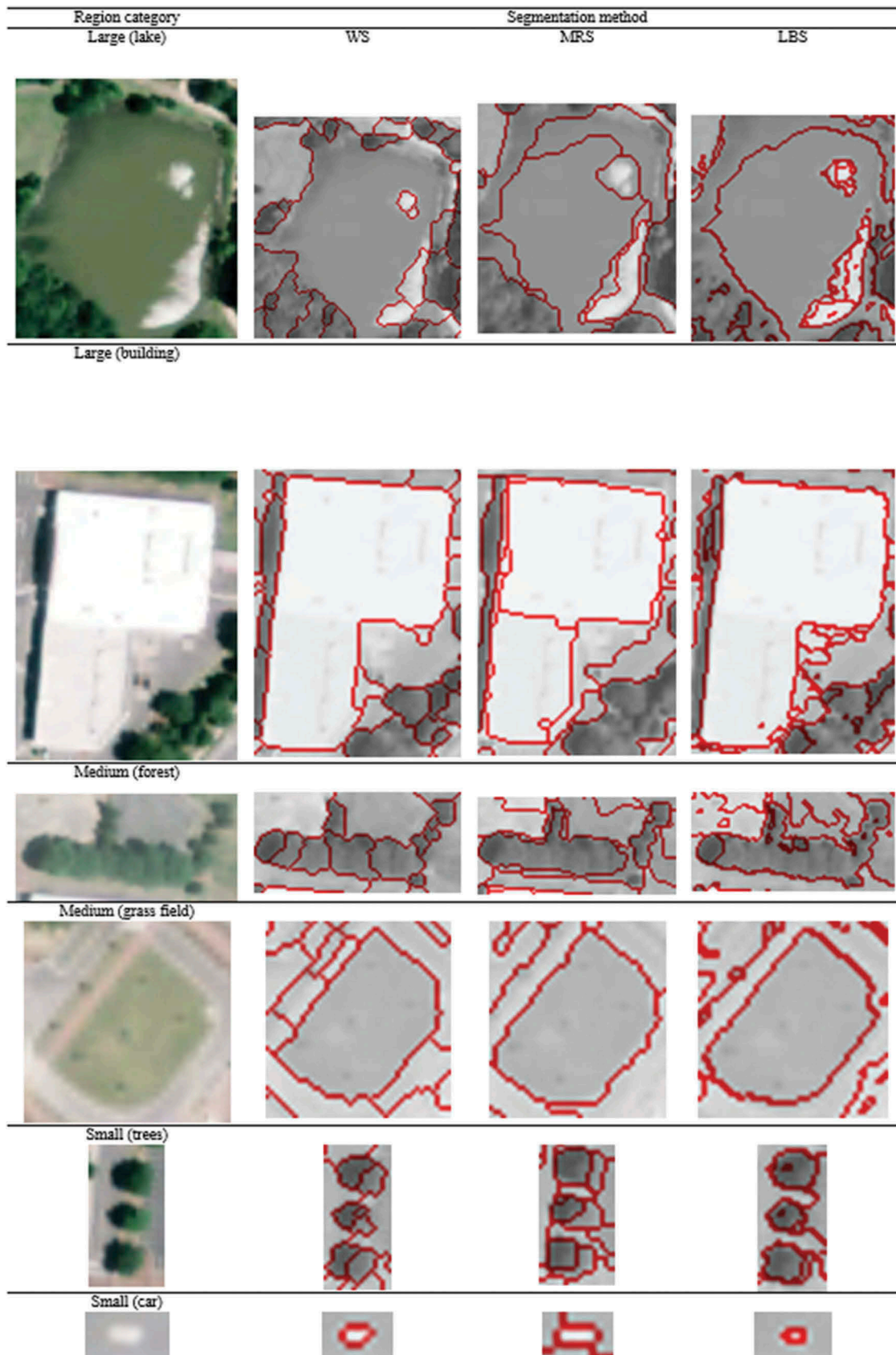
Figure 10. Overlapping of segmentation results produced on some discernible regions.

The Hamming and Jaccard distances measure dissimilarity between sample sets, although it is known that Jaccard distance is usually more sensitive to small sample sizes, but that was not a problem here, because our sample size was equal or superior to 500 samples. Any of the three segmentation methods showed no superior performance under these two distances, getting values between 22% and 26%, with no statistical differences between them.

The Positional discrepancy (PD) and Area fit index (AFI) metrics, both involve spatial coincidences

Figure 11. Segmentations by category sizes (a) small, (b) medium and (c) large for panchromatic image (Ikonos-Pan).

Table 3. Metric values (Bold numbers mean statistically different values and numbers in parenthesis correspond to panchromatic band).

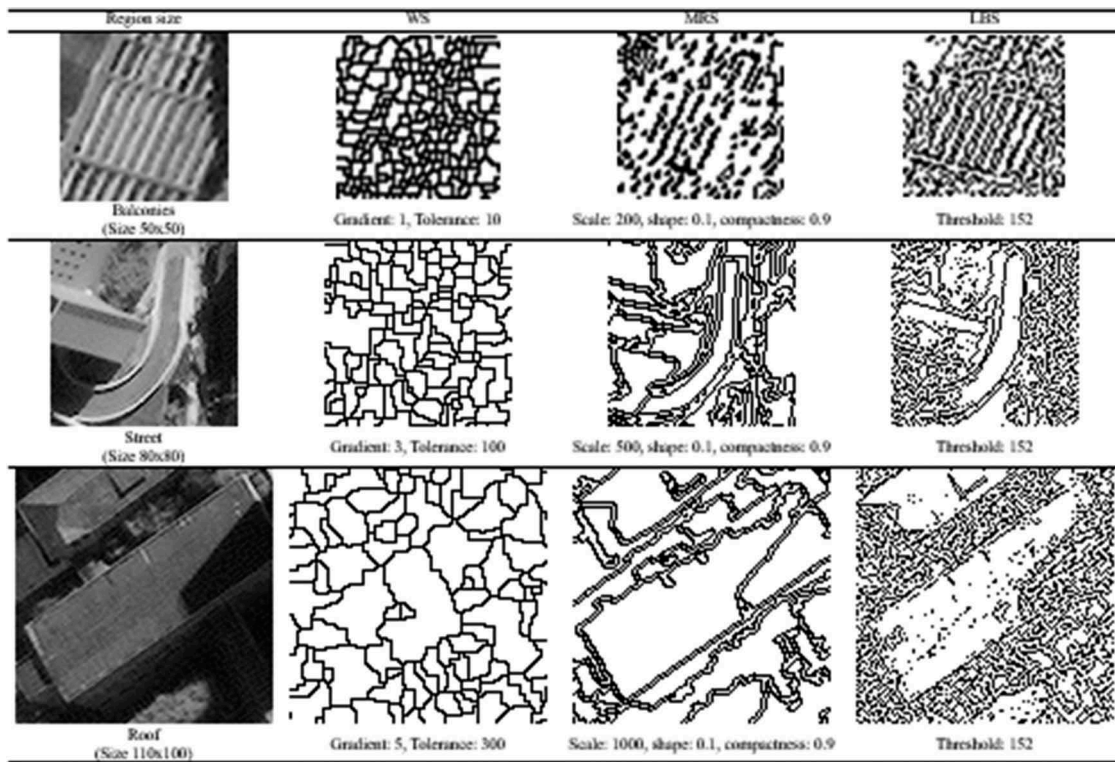|	Segmentation methods		
Evaluation metric	Watershed (WS)	Multiresolution (MRS)	Proposed (LBS)
Rand Index	**0.72 ± 0.04*** (*0.59*)	**0.71 ± 0.05*** (*0.65*)	0.64 ± 0.06 (0.62)
Positional discrepancy	4.60 ± 1.47 (*20.38*)	4.49 ± 1.81 (*18.23*)	**3.81 ± 2.49*** (2.47)
Area Fit Index	0.32 ± 0.24 (*0.75*)	**0.15 ± 0.68*** (0.80)	0.30 ± 0.22 (0.50)
Hamming distance	0.22 ± 0.10 (0.29)	0.22 ± 0.10 (0.23)	0.24 ± 0.07 (0.25)
Jaccard distance	0.23 ± 0.11 (0.29)	0.24 ± 0.10 (0.23)	0.26 ± 0.08 (0.27)
Fowlkes-Mallows Index	**0.82 ± 0.03*** (*0.76*)	**0.82 ± 0.04*** (*0.77*)	0.75 ± 0.07 (0.74)
Overall Accuracy	0.78 ± 0.10 (0.71)	0.77 ± 0.10 (0.77)	0.76 ± 0.07 (0.75)

between regions but behave differently, because the positional discrepancy takes into account only the centroid position, while Area fit index uses the regions' boundaries. The positional discrepancy metric for the proposed method showed statistical difference, indicating their superior performance, getting positional discrepancy of 3.81 pixels only, meaning it can detect the regions' centroids easily. In relation to the AFI metric, with its values varying in the range [−1,1], showed the LBS method behaves satisfactorily (0.30), giving some space to improve.

With the Fowlkes-Mallows (FM) index, defined as the geometric mean of the pairwise precision and recall extracted from contingency table, indicated the proposed LBS method needs to reduce the interference of false negatives (FN) values, which affected statistically significant its performance.

Last, but not least, the overall accuracy (OA) metric, maybe one the most used metric to evaluate segmentation performance, showed similar performance for all the three segmentation methods, with no statistical differences between them, with values varying between 76% and 78%. It is worthy to point out that the segmentation parameters were found following a trial-and-error methodology in the case for the WS method, while for MRS method, its scale parameter was defined using the ESP tool (Dragut et al., 2014) and its shape parameter through a visual examination, meaning that possibly a different parameter setting could let them improve their performances . Figure 12 illustrates the segmentations produced for each method organized by category size, including the proposed method (LBS) and the two selected methods (WS and MRS) used for comparison purposes through the calculated metrics.

Fig. 11. Small Scale

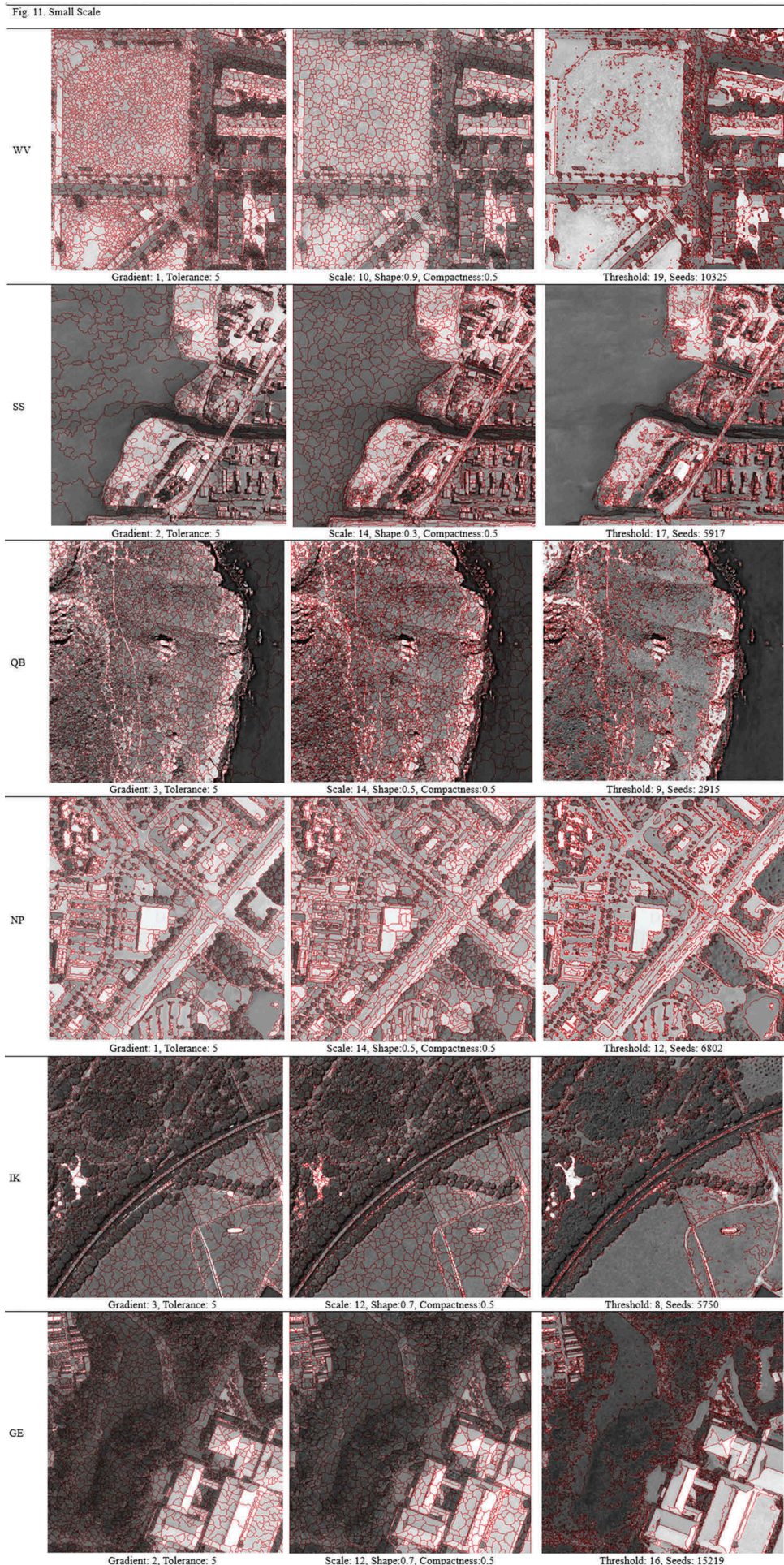| WV | | |
Gradient: 1, Tolerance: 5	Scale: 10, Shape:0.9, Compactness:0.5	Threshold: 19, Seeds: 10325
SS		
Gradient: 2, Tolerance: 5	Scale: 14, Shape:0.3, Compactness:0.5	Threshold: 17, Seeds: 5917
QB		
Gradient: 3, Tolerance: 5	Scale: 14, Shape:0.5, Compactness:0.5	Threshold: 9, Seeds: 2915
NP		
Gradient: 1, Tolerance: 5	Scale: 14, Shape:0.5, Compactness:0.5	Threshold: 12, Seeds: 6802
IK		
Gradient: 3, Tolerance: 5	Scale: 12, Shape:0.7, Compactness:0.5	Threshold: 8, Seeds: 5750
GE		
Gradient: 2, Tolerance: 5	Scale: 12, Shape:0.7, Compactness:0.5	Threshold: 16, Seeds: 15219

Figure 12. Image segmentation results (red lines) over each original images (grayscale version), organized for category sizes, (a) small, (b) medium and (c) large.
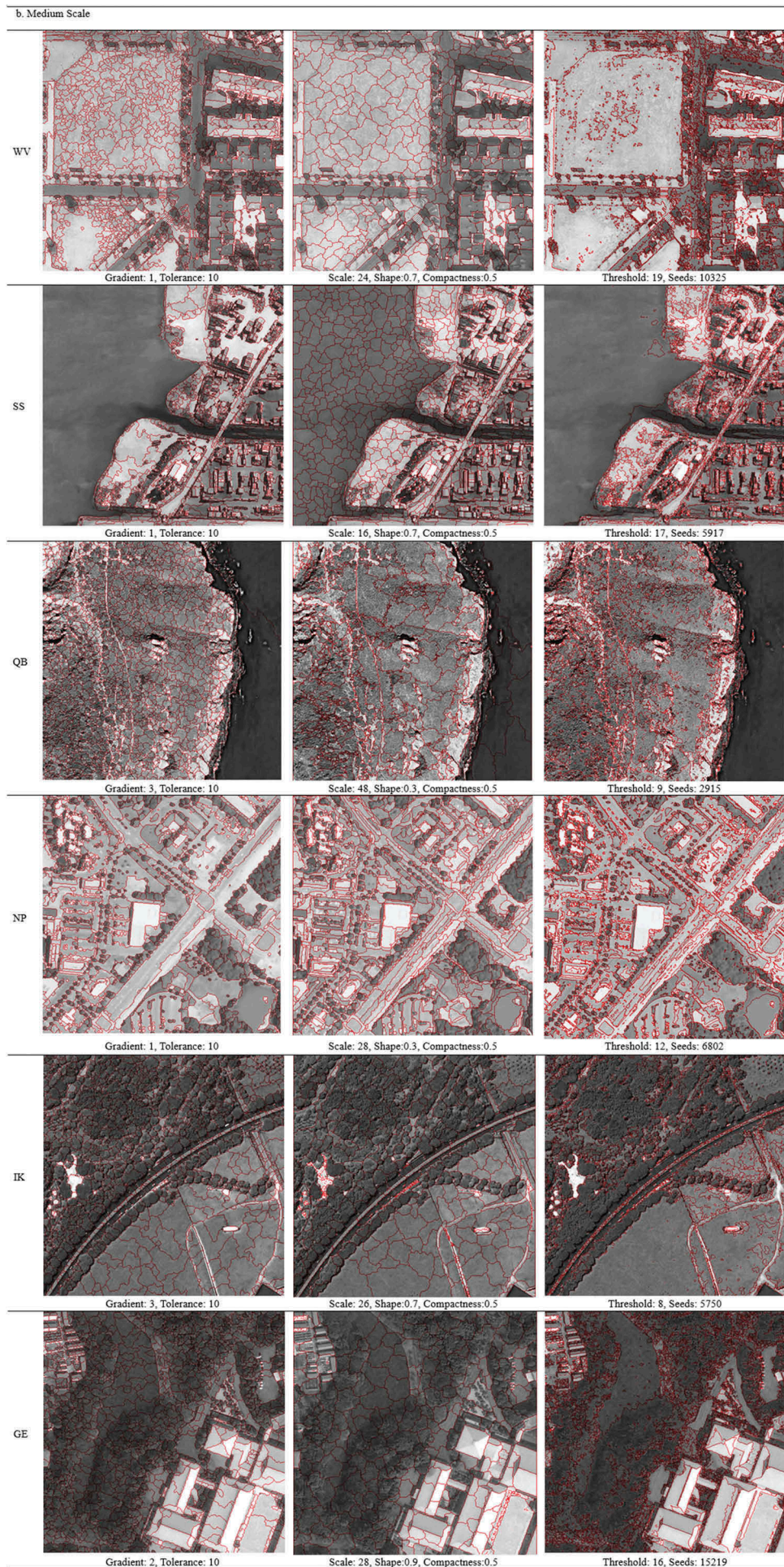
b. Medium Scale

Figure 12. Continued.

c. Large scale

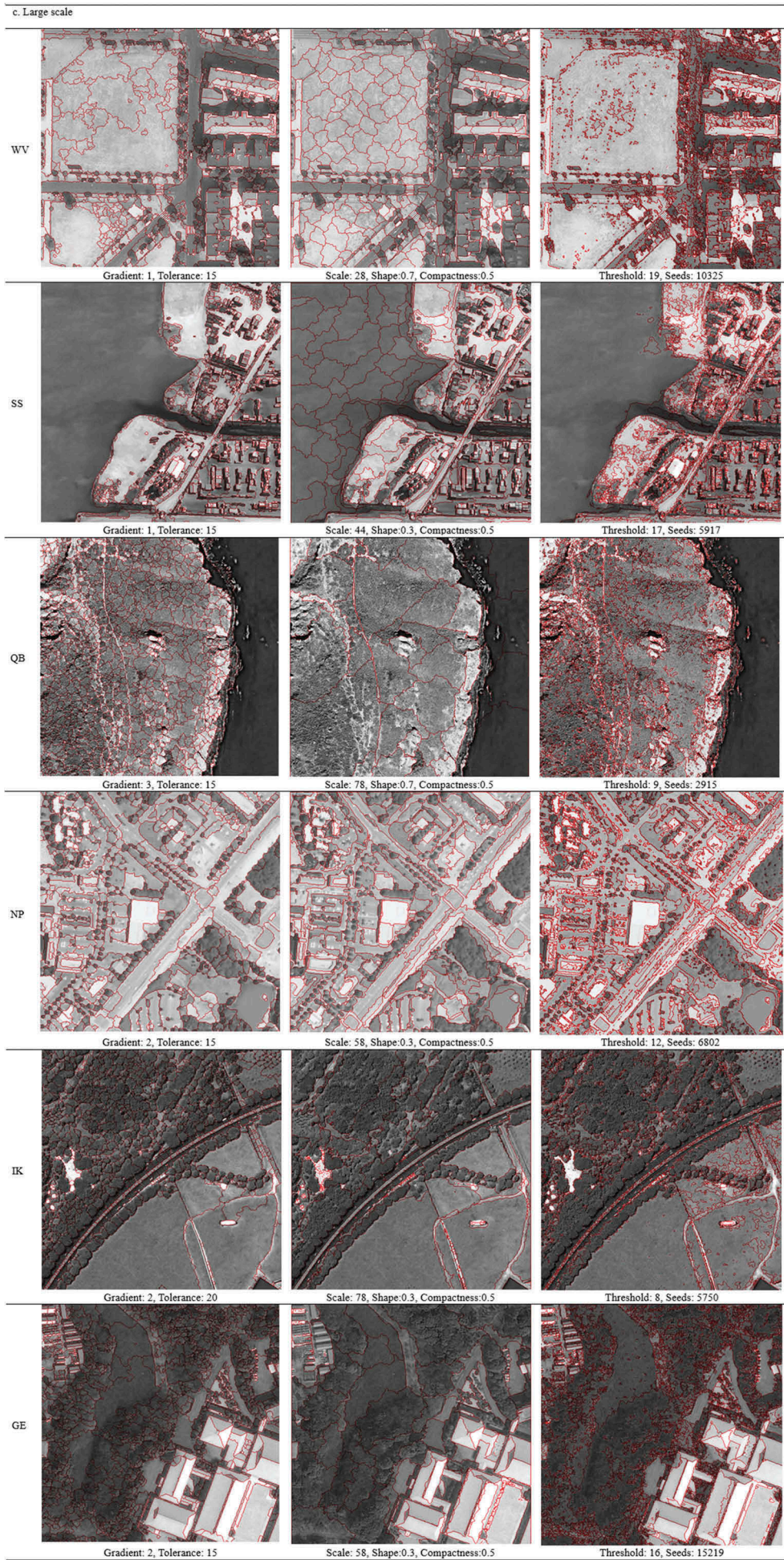| | Gradient: 1, Tolerance: 15 | Scale: 28, Shape:0.7, Compactness:0.5 | Threshold: 19, Seeds: 10325 |
WV			
SS	Gradient: 1, Tolerance: 15	Scale: 44, Shape:0.3, Compactness:0.5	Threshold: 17, Seeds: 5917
QB	Gradient: 3, Tolerance: 15	Scale: 78, Shape:0.7, Compactness:0.5	Threshold: 9, Seeds: 2915
NP	Gradient: 2, Tolerance: 15	Scale: 58, Shape:0.3, Compactness:0.5	Threshold: 12, Seeds: 6802
IK	Gradient: 2, Tolerance: 20	Scale: 78, Shape:0.3, Compactness:0.5	Threshold: 8, Seeds: 5750
GE	Gradient: 2, Tolerance: 15	Scale: 58, Shape:0.3, Compactness:0.5	Threshold: 16, Seeds: 15219

Figure 12. Continued.

The proposed segmentation method (LBS) is based on region growing algorithm, whose parameters, thresholds and seeds, were determined completely automatic, without any user assistance, a very relevant behavior because it reached a comparable performance with two well-known methods for segmentation of remote sensing images, such as watershed and multiresolution, both following a user-assisted approach, although the latter could be described as a semi-automatic approach, because it was used in conjunction with a tool for scale determination, as was mentioned before.

Discussion

Automatic image processing

As it was mentioned before, the Line-based segmentation method (LBS) reached a similar performance with watershed (WS) and multiresolution (MRS) methods according to overall accuracy index, Hamming and Jaccard distances, and a superior performance according to the positional discrepancy metric. We consider this result very relevant because it reached following a fully automatic approach, especially with overall accuracy which evaluates the true positives cases (i.e. real regions' borders) overall samples used. While the Hamming and Jaccard distances confirmed that have not found statistical differences between the proposed method compared against the other two supervised methods, knowing these two distance metrics represent the false negatives cases (i.e. real regions' borders not identified), that can be interpreted as a complementary measure of overall accuracy discussed earlier.

The best performance for the positional discrepancy metric obtained by the proposed method, showing statistical differences against the supervised methods, indicated a superior capability for centroids' detection of segments, which is very important for forward image analysis, such as points clustering or weighting by centroids density, that could be used for estimation of region agglomeration or identification of trends in urban growth. To achieve this metric value without using any prior knowledge about image content during the segmentation process represents a real advantage point supporting its high reproducibility, eliminating the need of the user's visual image interpretation or being influenced by posterior image analysis.

It is important to mention the proposed method (LBS) does not represent a universal solution for segmentation of VHSR remote sensing images, because it is known that a good image segmentation strongly depends on the particular information the user is looking for, but it could be considered as a fair initial step when the user does not have any knowledge about the image content, as a preliminary segmentation result to compare against other image segmentation results or when it is necessary to segment a high volume of images automatically.

Scale selection

As we have explained before, the proposed segmentation (LBS) is a single-scale method, departing from the multi-scale approaches that actually are highly appreciated for the analysis of very complex natural phenomena captured by remote sensing images, using hierarchical linking to represent relationships between regions that belong to different scales. Although our approach is single-scale, it has the capability to grow simultaneously different sized regions (e.g. small and large regions), because it does not have a limiting size parameter to prevent it to generate regions larger than a specific size, although it becomes fundamental to generate hierarchical levels under other segmentation approaches.

The image segmentation methods chosen for comparison purposes were the Watershed and Multiresolution segmentation, known as supervised methods, which require most of the times a trial-and-error methodology for the parameters' tuning step. The Watershed segmentation (https://imagej.net/Morphological_Segmentation) has two parameters, tolerance and gradient, varying their values according to the category region size under analysis. For the tolerance parameter, three values were chosen: 5, 10 and 15 corresponding to small, medium and large regions. As the tolerance value represents an intensity parameter, the chosen values were 10 (8-bit images) and 10, 100, 300 (16-bit pachromatic image) as starting points, knowing that lower and greater values increase or decrease the number of regions, respectively. The gradient type parameter corresponds to the difference between dilation and erosion operations within a particular neighbourhood. The selected gradient parameter value varied between 1 and 5, that were chosen through a visual examination.

The multiresolution segmentation has three main parameters, scale, shape and compactness. To define their values, we followed a mixed strategy, combining an automatic methodology (for scale parameter), a trial-and-error methodology (for shape parameter) and fixed value (for compactness parameter). For the scale definition, we used the ESP (Estimation of Scale Parameter) tool (Dragut et al., 2014), based on Local Variance (LV) of region heterogeneity within a scene, choosing a hierarchy of scale from thresholds in rates of changes. The selected scales for all images varied between these ranges for 8-bit images (10–14, 16–48 and 28–78 for small, medium and large regions, respectively) and for 16-bit image (200, 500 and 100 for small, medium and large regions, respectively). The shape parameter was selected through a visual examination, assigning values in the range 0.3–07, while compactness value was defined as 0.5

Table 4. Change detected in segmentations after applying the LBS method for each image (5 repetitions).

Criteria	Value	Change
Structural Similarity Index (SSIM)	0.80 ± 0.04	0.67%
Mean Square Error (MSE)	298.80 ± 27.16	2.67%
Number of objects	$13,041 \pm 84$	0.01%

(Kavzoglu & Yildiz, 2014) for 8-bit images, while 0.1 and 0.9 for shape and compactness parameters were used for the 16-bit image. We consider very important and necessary to overcome the use of the trial-and-error methodology for parameters' definition, such as scale, to enable the automatic implementation of the image segmentation process. Usually, it involves multiple testing converting this into a subjective and time-consuming task. Here, it was used as an optimization process to define parameters' values automatically allowing the identification of the smaller and discernible regions in the image. Although at first, it was not thought to use it on multiscale approaches, the original image could be splitted in small fractions to segment regions with larger sizes, possibly enabling its hierarchical behavior. The approaches proposed by Dragut et al. (2014) and Yang et al. (2014, 2015) represents a new and promising way to identify the scales for the images under study, but they still require the user assistance to define additional parameters such as the number of iterations, ranges or increments.

Seeds selection

The process of seeds selection sometimes impacts negatively the quality of image segmentation when region-growing algorithms are used (Fan et al., 2005). Under the present approach, these seeds were selected automatically using a new method based on the highest entropy line. Also, as the regions start to grow from that line initially, it reduces the intrinsic variability between segmentations from the same image, as is explained next.

After performing several segmentations (five in total for each image) using the proposed method (LBS), we found very small differences (0.01%) between obtained regions, as can be seen in Table 4.

To evaluate further the impact of seeds selection on the behavior of the proposed segmentation method, we selected two statistical criteria; Mean Square Error (MSE), which is a global measure that estimates the perceived error between two images with a value of zero (0) indicating perfect similarity, and Structural Similarity Index (SSIM), which tries to model the perceived changes in images' structural information using small windows, giving values between −1 (different) and 1 (perfect similarity) (Wang, Bovik, Sheikh, & Simoncelli, 2004) . Both criteria confirmed a high similarity and small changes between repeated

segmentations confirming the high consistency and reproducibility of the proposed LBS image segmentation method (Table 4).

Specifically, the MSE obtained was close to 300, which represents a very low value compared to the total of regions used (13,000 approx.), meaning it is just 2% away from the optimal MSE value (zero value), also with a very low variability (2.67%). In relation with the SSIM metric, its optimal value is one, and the results obtained from LBS segmentations gave an average value of 0.80, near to the optimal value, and also with extremely low variability (0.67%).

Conclusions

In this manuscript, it is proposed a novel remote sensing image segmentation method named Line-based segmentation (LBS) adopting a region growing approach with the aim to eliminate human intervention in segmenting VHSR remote sensing imagery to make possible the automatic image segmentation. This was realized through the definition of a proper segmentation parameter without the use of a trial-and-error methodology, and using the image's highest entropy line instead, getting the required thresholds and seeds through an optimization procedure. This method was tested on several VSHR images and evaluated qualitatively and quantitatively, showing comparable segmentation performance in comparison with two well-known segmentation methods in remote sensing.

The LBS method is able to delineate simultaneously both small and large regions because it does not have any limiting size parameter, becoming a very important feature knowing that every single remote sensing image usually presents regions with high variability in size and shape. Also, it has shown the capacity to reduce or eliminate the need of prior knowledge about the scene allowing the automatic processing of large volumes of remote sensing images. Because of these qualities, we consider this line-based approach could open new ways to process and analyse VHSR remote sensing imagery effectively.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Jaime Lopez 　 http://orcid.org/0000-0002-4267-1596

John W. Branch http://orcid.org/0000-0002-0378-028X
Gang Chen http://orcid.org/0000-0002-7469-3650

References

Acharya, T., & Ray, A.K. (2005). *Image processing. Principles and applications* (pp. 37). Wiley.

Adeline, K.R.M., Chen, M., Briottet, X., Pang, S.K., & Paparoditis, N. (2013). Shadow detection in very high spatial resolution aerial images: A comparative study. *ISPRS Journal of Photogrammetry and Remote Sensing, 80,* 21–38. doi:10.1016/j.isprsjprs.2013.02.003

Akcay, H.G., & Aksoy, S. (2008). Automatic detection of geospatial objects using multiple hierarchical segmentations. *IEEE Transactions on Geoscience and Remote Sensing, 46* (7), 2097–2111. doi:10.1109/TGRS.2008.916644

Arbelaez, P., Maire, M., Fowlkes, C., & Malik, J. (2011). Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 33*(5), 898–916. doi:10.1109/TPAMI.2010.161

Baatz, M., & Schape, A. (2000). Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation. *Angewandte Geographische Informationsverarbeitung, 12,* 12–23.

Baltsavias, E.P. (2004). Object extraction and revision by image analysis using existing geodata and knowledge: Current status and steps towards operational systems. *ISPRS Journal of Photogrammetry and Remote Sensing, 58*(3), 129–151. doi:10.1016/j.isprsjprs.2003.09.002

Belgiu, M., & Drăguţ, L. (2014). Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. *ISPRS Journal of Photogrammetry and Remote Sensing, 96,* 67–75. doi:10.1016/j.isprsjprs.2014.07.002

Beneš, M., & Zitová, B. (2015). Performance evaluation of image segmentation algorithms on microscopic image data: Performance evaluation of image segmentation algorithms. *Journal of Microscopy, 257*(1), 65–85. doi:10.1111/jmi.12186

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 Journal of Photogrammetry and Remote Sensing, 58*(3/4), 239–258. doi:10.1016/j.isprsjprs.2003.10.002

Blaschke, T. (2005). A framework for change detection based on image objects. In S. Erasmi, B. Cyffka, & M. Kappas (Eds.), *Remote sensing & gis for environmental studies: applications in geography* (Vol. 113, pp. 1–9). Göttingen: Goltze.

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing, 65*(1), 2–16. doi:10.1016/j.isprsjprs.2009.06.004

Blaschke, T., Hay, G.J., Kelly, M., Lang, S., Hofmann, P., Addink, E., . . . Tiede, D. (2014). Geographic object based image analysis-towards a new paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing, 87,* 180–191. doi:10.1016/j.isprsjprs.2013.09.014

Bosch, A., Zisserman, A., & Munoz, X. (2007, December). Image classification using random forests and ferns. *International Conference on Computer Vision (ICCV-2007).* doi:10.1109/ICCV.2007.4409066.

Burnett, C., & Blaschke, T. (2003). A multi-scale segmentation/object relationship modeling methodology for landscape analysis. *Ecological Modelling, 168*(3), 233–249. doi:10.1016/S0304-3800(03)00139-X

Byun, Y., Kim, D., Lee, J., & Kim, Y. (2011). A framework for the segmentation of high-resolution satellite imagery using modified seeded-region growing and region merging. *International Journal of Remote Sensing, 32* (16), 4589–4609. doi:10.1080/01431161.2010.489066

Cadík, M. (2008). Perceptual evaluation of color-to-grayscale image conversions. *Pacific Graphics, 27*(7), 1745–1754. doi:10.1111/j.1467-8659.2008.01319.x

Cardoso, J.S., & Corte-Real, L. (2005). Toward a generic evaluation of image segmentation. *IEEE Transactions on Image Processing, 14*(11), 1773–1782. doi:10.1109/TIP.2005.854491

Chen, G., Hay, G.J., & St-Onge, B. (2012). A GEOBIA framework to estimate forest parameters from lidar transects, Quickbird imagery and machine learning: A case study in Quebec, Canada. *International Journal of Applied Earth Observation and Geoinformation, 15,* 28–37. doi:10.1016/j.jag.2011.05.010

Cheng, G., & Han, J. (2016). A survey on object detection in optical remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing, 117,* 11–28. doi:10.1016/j.isprsjprs.2016.03.014

Congalton, R.G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment, 37,* 35–46. doi:10.1016/0034-4257(91)90048-B

Dey, V., Zhang, Y., & Zhong, M. (2010). *A review on image segmentation techniques with remote sensing perspective.* W. Wagner & B. Székely (Eds.), (Vol. XXXVIII, Part 7A). Vienna, Austria: na. Retrieved from http://www.isprs.org/proceedings/XXXVIII/part7/a/pdf/31_XXXVIII-part7A.pdf

Drăguţ, L., Tiede, D., & Levick, S.R. (2010). ESP: A tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *International Journal of Geographical Information Science, 24*(6), 859–871. doi:10.1080/13658810903174803

Dragut, L., Csillik, O., Eisank, C., & Tiede, D. (2014). Automated parameterisation for multi-scale image segmentation on multiple layers. *ISPRS Journal of Photogrammetry and Remote Sensing, 88,* 119–127. doi:10.1016/j.isprsjprs.2013.11.018

Fan, J., Zeng, G., Body, M., & Hacid, M.-S. (2005). Seeded region growing: An extensive and comparative study. *Pattern Recognition Letters, 26*(8), 1139–1156. doi:10.1016/j.patrec.2004.10.010

Fowlkes, E.B., & Mallows, C.L. (1983). A method for comparing two hierarchical clusterings. *Journal of the American Statistical Association, 78*(383), 553. doi:10.2307/2288117

Gonzalez, R.C., & Woods, R.E. (2008). *Digital image processing* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Günes, A., Kalkan, H., & Durmu, E. (2015). Optimizing the color-to-grayscale conversion for image classification. *Signal Image and Video Processing.* doi:10.1007/s11760-015-0828-7

Gurcan, M.N., Boucheron, L.E., Can, A., Madabhushi, A., Rajpoot, N.M., & Yener, B. (2009). Histopathological image analysis: A review. *IEEE Reviews in Biomedical Engineering, 2,* 147–171. doi:10.1109/RBME.2009.2034865

Han, J., Zhang, D., Cheng, G., Guo, L., & Ren, J. (2015). Object detection in optical remote sensing images based on weakly supervised learning and high-level feature learning. *Ieee Transactions on Geoscience and Remote Sensing, 53*(6):3325-3337. doi: 10.1109/TGRS.2014.2374218

Hay, G.J., Blaschke, T., Marceau, D.J., & Bouchard, A. (2003). A comparison of three image-object methods for

the multiscale analysis of landscape structure. *ISPRS Journal of Photogrammetry and Remote Sensing*, 57(5), 327–345. doi:10.1016/S0924-2716(02)00162-4

Hay, G.J., & Castilla, G. (2006, May). Object-based image analysis: Strengths, opportunities and threats (SWOT). In OBIA (Ed.), *The international archives of the photogrammetry, remote sensing and spatial information sciences,* ISPRS, Salzburg, Austria (pp. 4–5).

Hirschmugl, M., Ofner, M., Raggam, J., & Schardt, M. (2007). Single tree detection in very high resolution remote sensing data. *Remote Sensing of Environment*, 110(4), 533–544. doi:10.1016/j.rse.2007.02.029

Hofmann, P. (2001). Detecting buildings and roads from IKONOS data using additional elevation information. *GeoBIT/GIS*, 6(1), 28–33.

ITU-R. (1990). *Recommendation BT.709 basic parameter values for the hdtv standard for the studio and for international programme exchange*. Geneva, Switzerland: International Telecommunication Union.

Kanan, C., & Cottrell, G.W. (2012). Color-to-Grayscale: Does the method matter in image recognition? *PloS One*, 7(1), e29740. doi:10.1371/journal.pone.0029740

Kavzoglu, T., & Yildiz, M. (2014). Parameter-based performance analysis of object-based image analysis using aerial and quikbird-2 images. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. II*, 7, 31–37. doi:10.5194/isprsannals-II-7-31-2014

Kumar, M., Singh, R.K., Raju, P.L.N., & Krishnamurthy, Y.V.N. (2014). Road network extraction from high resolution multispectral satellite imagery based on object oriented techniques. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, II*, 8, 107–110. doi:10.5194/isprsannals-II-8-107-2014

Lang, S. (2008). Object-based image analysis for remote sensing applications: Modeling reality – Dealing with complexity. In T. Blaschke, S. Lang, & G.J. Hay (Eds.), *Object-based image analysis* (pp. 3–27). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-540-77058-9_1

Liu, J., Li, P., & Wang, X. (2015). A new segmentation method for very high resolution imagery using spectral and morphological information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 101, 145–162. doi:10.1016/j.isprsjprs.2014.11.009

López-Granados, T.-S.J., & Peña, J.M. (2015). An automatic object-based method for optimal thresholding in UAV images: Application for vegetation detection in herbaceous crops. *Computers and Electronics in Agriculture*, 114, 43–52. doi:10.1016/j.compag.2015.03.019

Lucieer, A. (2004). *Uncertainties in segmentation and their visualisation* (PhD). Utrecht University and International Institute for Geo-Information Science and Earth Observation (ITC), Utrecht. Retrieved from http://eprints.utas.edu.au/3164/2/lucieer_thesis_www.pdf

Ma, K., Zhao, T., Zeng, K., & Wang, Z. (2015). Objective quality assessment for color-to-gray image conversion. *IEEE Transactions on Image Processing*, 24(12), 4673–4685. doi:10.1109/TIP.2015.2460015

MacKay, D.J.C. (2003). *Information theory, inference, and learning algorithms*. Cambridge, UK: Cambridge University Press.

Marr, D., & Hildreth, E. (1980, February). Theory of edge detection. *Proceedings of the Royal Society of London*, London, GB, 207(1167), 187–217. doi:10.1098/rspb.1980.0020.

Montaghi, A., Larsen, R., & Greve, M.H. (2013). Accuracy assessment measures for image segmentation goodness of

the land parcel identification system (lpis) in Denmark. *Remote Sensing Letters*, 4(10), 946–955. doi:10.1080/2150704X.2013.817709

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 Sensing of Environment*, 115(5), 1145–1161. doi:10.1016/j.rse.2010.12.017

Nafchi, H.Z., Shahkolaei, A., Hedjam, R., & Cheriet, M. (2017). CorrC2G: Color to gray conversion by correlation. *IEEE Signal Processing Letters*, 24(11), 1651–1655. doi:10.1109/LSP.2017.2755077

Nobrega, R.A.A., O'Hara, C.G., & Quintanilha, J.A. (2008). An object-based approach to detect road features for informal settlements near Sao Paulo, Brazil. In T. Blaschke, S. Lang, & G.J. Hay (Eds.), *Object-based image analysis* (pp. 589–607). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-540-77058-9_32

Norouzi, M., Fleet, D.J., & Salakhutdinov, R. (2012). Hamming distance metric learning. In *NIPS'12 Proceedings of the 25th international conference on neural information processing systems* (pp. 1061–1069), Lake Tahoe, Nevada.

Ose, K., Corpetti, T., & Demagistri, L. (2016). Multispectral Satellite Image Processing. In N. Baghdadi & M. Zribi (Eds.), *Optical remote sensing of land surface techniques and methods* (pp. pp.57–124). Elsevier. doi:10.1016/B978-1-78548-102-4.50002-8

Pal, N.R., & Pal, S.K. (1993). A review on image segmentation techniques. *Pattern Recognition*, 26(9), 1277–1294. doi:10.1016/0031-3203(93)90135-J

Pohl, C., & Van Genderen, J.L. (1998). Multisensor image fusion in remote sensing: Concepts, methods and applications. *International Journal of Remote Sensing*, 19(5), 823–854. doi:10.1080/014311698215748

Rama-Rao, N.R., Garg, P.K., & Ghosh, S.K. (2006). The effect of radiometric resolution on the retrieval of leaf area index from agricultural crops. *GIScience & Remote Sensing*, 43(4), 377–387. doi:10.2747/1548-1603.43.4.377

Rand, W.M. (1971). Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*, 66(336), 846. doi:10.1080/01621459.1971.10482356

Razakarivony, S., & Jurie, F. (2016). Vehicle detection in aerial imagery : A small target detection benchmark. *Journal of Visual Communication and Image Representation*, 34, 187–203. doi:10.1016/j.jvcir.2015.11.002

Real, R., & Vargas, J.M. (1996). The probabilistic basis of jaccard's index of similarity. *Systematic Biology*, 45(3), 380–385. doi:10.1093/sysbio/45.3.380

Rouhi, R., Jafari, M., Kasaei, S., & Keshavarzian, P. (2015). Benign and malignant breast tumors classification based on region growing and CNN segmentation. *Expert Systems with Applications*, 42(3), 990–1002. doi:10.1016/j.eswa.2014.09.020

Saeed, K., & Albakoor, M. (2009). Region growing based segmentation algorithm for typewritten and handwritten text recognition. *Applied Soft Computing*, 9(2), 608–617. doi:10.1016/j.asoc.2008.08.006

Salvador, S., & Chan, P. (2004). *Determining the number of clusters/segments in hierarchical clustering/ segmentation algorithms* (pp. 576–584). IEEE Computer Society. doi:10.1109/ICTAI.2004.50

Sánchez, J., Martínez, E., & Arquero, A. (2015). Improving parameters selection of a seeded region growing method for multiband image segmentation. *IEEE Latin America Transactions*, 13(3), 843–849. doi:10.1109/TLA.2015.7069113

Singh, B., & Singh, A.P. (2008). Edge detection in gray level images based on the shannon entropy. *Journal of Computer Science*, *4*(3), 186–191. doi:10.3844/jcssp.2008.186.191

Tabb, M., & Ahuja, N. (1997). Unsupervised multiscale image segmentation by integrated edge and region detection. *IEEE Transactions on Image Processing*, *6*(5), 642–655. doi:10.1109/83.568922

Troya-Galvis, A., Gançarski, P., & Berti-Équille, L. (2003). Remote sensing image analysis by aggregation of segmentation-classification collaborative agents. *Ecological Modelling*, *168*(3), 233–249. doi:10.1016/j.patcog.2017.08.030

Van Etten, A. (2018). You only look twice: Rapid multi-scale object detection in satellite imagery. *arXiv:1805.09512v1 [cs.CV]*.

Verde, N., Mallinis, G., Tsakiri-Strati, M., Georgiadis, C., & Patias, P. (2018). Assessment of radiometric resolution impact on remote sensing data classification accuracy. *Remote Sensing*, *10*(8), 1267–1284. doi:10.3390/rs10081267

Wang, Z., Bovik, A.C., Sheikh, H.R., & Simoncelli, E.P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13*(4), 600–612. doi:doi:10.1109/TIP.2003.819861

Wang, Z., Jensen, J.R., & Im, J. (2010). An automatic region-based image segmentation algorithm for remote sensing applications. *Environmental Modelling & Software*, *25*(10), 1149–1165. doi:10.1016/j.envsoft.2010.03.019

Witkin, A. (1984, March). Scale-space filtering: A new approach to multi-scale description. In *ICASSP '84. IEEE International Conference on Acoustics, Speech, and Signal Processing* (pp.19–21). San Diego, CA doi:10.1109/ICASSP.1984.1172729

Yang, J., He, Y., & Weng, Q. (2015). An automated method to parameterize segmentation scale by enhancing intra-segment homogeneity and intersegment heterogeneity. *IEEE Geoscience and Remote Sensing Letters*, *12*(6), 1282–1286. doi:10.1109/LGRS.2015.2393255

Yang, J., Li, P., & He, Y. (2014). A multi-band approach to unsupervised scale parameter selection for multi-scale image segmentation. *ISPRS Journal of Photogrammetry and Remote Sensing*, *94*, 13–24. doi:10.1016/j.isprsjprs.2014.04.008

Yu, Q., Gong, P., Clinton, N., Biging, G., Kelly, M., & Schirokauer, D. (2006). Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. *Photogrammetric Engineering & Remote Sensing*, *72*(7), 799–811. doi:10.14358/PERS.72.7.799

Zhang, H., Fritts, J.E., & Goldman, S.A. (2008). Image segmentation evaluation: A survey of unsupervised methods. *Computer Vision and Image Understanding*, *110*(2), 260–280. doi:10.1016/j.cviu.2007.08.003

Zhang, X., Xiao, P., Feng, X., Feng, L., & Ye, N. (2015). Toward evaluating multiscale segmentations of high spatial resolution remote sensing images. *IEEE Transactions of Geoscience and Remote Sensing*, *53*(7), 3694–3706. doi:10.1109/TGRS.2014.2381632

Zhang, Y., & Maxwell, T. (2006, May). A fuzzy logic approach to supervised segmentation for object-oriented classification. In ASPRS (ed.), *Annual conference* (pp. 1–11). Reno, Nevada: Author.

Zhang, Y.-J. (2015). Half century for image segmentation. http://Services.Igi-Global.Com/Resolvedoi/Resolve.Aspx?Doi=10.4018/978-1-4666-5888-2.Ch584, 5906–5915

Zhao, Y.Q., Wang, X.H., Wang, X., & Shih, F.Y. (2014). Retinal vessels segmentation based on level set and region growing. *Pattern Recognition*, *47*(7), 2437–2446. doi:10.1016/j.patcog.2014.01.006

Zheng, H., Cheng, T., Li, D., Zhou, X., Yao, X., Tian, Y., . . . Zhu, Y. (2018). Evaluation of RGB, color-infrared and multispectral images acquired from unmanned aerial systems for the estimation of nitrogen accumulation in rice. *Remote Sensing*, *10*(6), 824. doi:doi:10.3390/rs10060824

Zhou, Y., Li, J., Feng, L., Zhang, X., & Hu, X. (2017). Adaptive scale selection for multiscale segmentation of satellite images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *10*(8), 3641–3651. doi:10.1109/JSTARS.2017.2693993