Automatic Image Segmentation Optimized by Bilateral Filtering

Javier Sanchez, Estibaliz Martinez, Agueda Arquero, and Diego Renza

Polytechnic University of Madrid, DATSI, Informatics Fac. Campus de Montegancedo
28660 Boadilla del Monte, Madrid, Spain
jarsahe@hotmail.com,
{emartinez,aarquero}@fi.upm.es,
d.renza@alumnos.upm.es

Abstract. The object-based methodology is one of the most commonly used strategies for processing high spatial resolution images. A prerequisite to object-based image analysis is image segmentation, which is normally defined as the subdivision of an image into separated regions. This study proposes a new image segmentation methodology based on a self-calibrating multi-band region growing approach. Two multispectral aerial images were used in this study. The unsupervised image segmentation approach begins with a first step based on a bidirectional filtering, in order to eliminate noise, smooth the initial image and preserve edges. The results are compared with ones obtained from Definiens Developper software.

Keywords: Image segmentation, Bilateral filter, Self-calibrating framework.

1 Introduction

Remote sensing is an effective technology to acquire information about geographic objects. The automatic classification of remotely sensed data is an essential action within the process of generating or updating Geographical Information System (GIS) databases. The thematic mapping is a widely adopted method to obtain land cover information from satellite or aerial images and many classification algorithms have been extensively applied to. Increasing demands on the accuracy and thematic resolution of land cover maps from remote sensing imagery has created a need for novel image analysis techniques. During classical image classification, each pixel is assigned to a final class of the entire object according to their statistical properties, instead of determining the class label for each pixel separately (pixel-based methods). Pixel-based image classification encountered serious problem in dealing with high spatial resolution images. Therefore, object-based methods represent a good alternative [1] because the effect of the spectral variability, critical inconvenient present in these images, can be minimized.

One motivation for the object-oriented approach is the fact that, in many cases, the expected result of most image analysis tasks is the extraction of real world objects, proper in shape and proper in classification. The concept of "object" plays one of

central roles in image interpretation. This also has impact on the reliability of object recognition, which requires good whole-object segmentation. In this sense, image segmentation is critical for subsequent image analysis.

Segmentation subdivides an image into its constituent regions or objects. The level at which the subdivision is carried out depends on the problem being solved. That is, the segmentation should stop when the objects of interest in an application have been identified. Traditionally segmentation methods can be divided into methods based on pixel, on edge, and region-based methods. The last kind of segmentation methods includes region growing, region merging, region splitting and their combinations [2]. A world famous example of object oriented image analysis software is Definiens Developper [3], in which the multi-resolution image segmentation method [4] is key and a patented technology. But whatever the method, in remote sensed imagery, a critical task is the selection of segmentation parameters. In most cases, the parameters are selected by trial and error.

The primary focus of this paper is to introduce a new designed and implemented region-based image segmentation algorithm following self-calibrating idea proposed by Paglieroni [5], with an optimal initialization step based on a bidirectional filtering [6], which smoothes the initial image in order to reduce noise within regions while preserving edges between them. This method is validated by comparing the obtained results with ones obtained from Definiens Developper software.

2 Methodology

2.1 Study Scenes

Several remote sensed images with different spatial resolutions were used in this study, but only the obtained results with two selected multispectral aerial images (Scene 1 and Scene 2) are showed in this paper. The aerial image represents 8-bit RGB data with a high spatial resolution of 50 cm. In this case, the selection of segmentation parameters will be critical because you can end up with hundreds of thousands of objects.

The images cover an area of the order of 0.065 km^2. As with any remote sensing project, it is helpful if you have an idea of what you are looking at. In these scenes there are extensions of natural mediterranean forest partially or fully degraded, with the city-planning, industrial advance and road infrastructures that degrade natural zones in other times. The left corner is placed at 429581.98E, 4472991.21N and 428394.20E, 4474495.79 N (UTM geographic coordinates, h30) respectively.

2.2 Bilateral Filtering

This research proposes an initialization step where the image is filtered with a bilateral filter. Bilateral filtering is a non-linear filtering technique introduced by Tomasi [6], where the weight of each pixel is computed using a Gaussian in the spatial domain multiplied by an influence function in the intensity domain that decreases the weight of pixels with large intensity differences. Pixels that are very different in intensity from the central pixel are weighted less even though they may be in close proximity to the central pixel. Therefore, this is applied as two Gaussian

filters at a localized pixel neighborhood, one in the spatial domain, named the domain filter (Fig.1a) which smoothes homogeneous areas, and one in the intensity domain, named the range filter (Fig.1b) which controls the smoothing for preserving edges. Thus, the main advantage of using bilateral filter (Fig. 1c) is the growth of large and homogeneous regions.
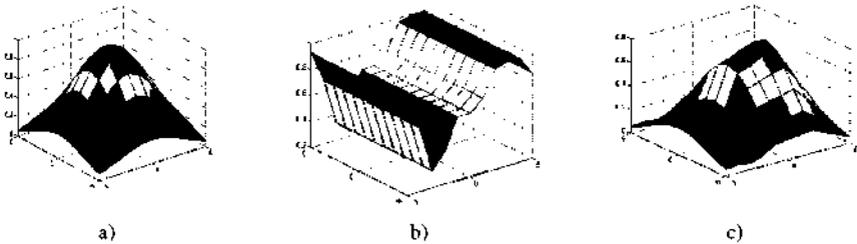

a)   b)   c)

Fig. 1. Domain (a), Range (b) and Bilateral (c) filters

2.3 Segmentation Process with Self-calibration Framework

Paglieroni [5] developed a self-calibration framework for automatically selecting parameter values that produce segmentations that most closely resemble a calibration edge map (derived separately using a simple edge detector), but there isn´t a commercial available implementation of this work. In order to implement our image segmentation software, it is proposed a new method based on its self-calibration framework for getting initial parameters but modifying some steps to try to optimize the quality of segmentation jointly with the computational efficiency.

After the initialization step where the image is processed with the bilateral filter described above, an adaptive Canny edge detection algorithm [7] is applied to obtain a calibration map. The Canny edge detector is employed to identify the mixed (edge) pixels in the image region and offers sub-pixel interpolation for detecting edges using a Gaussian filter. The algorithm is based on the computation of the image gradient. It works in a multistage process, which is summarized below. First, the image is smoothed by Gaussian convolution, then two-dimensional first derivatives are computed, the gradient magnitude (edge strength) and gradient direction are calculated. Traditionally, following steps consist of the non-maxima suppression, and the hysteresis thresholding by two parameters, low and high thresholds. These are used for estimating the population minimum parameter of region growing which is used to avoid small regions. These processes are applied with the objective of obtaining 1-pixel wide contours and to remove noisy maxima without breaking the contours, respectively.

In the methodology proposed, it is suggested erasing non-maxima suppression step because several experiments have showed that it is not critical for our goal, however this step would increase the execution time of the process. To obtain the region maps, the starting points of the segmentation, often referred to as 'seed' pixels, have to be identified. The regions are built around these pixels by joining the similar neighboring pixels to them. To compute the similarity the Euclidean distance is used. Once the regions have been obtained, a merging process is performed in order to eliminate

small regions. The complete procedure is summarized in the next program code (region-growing pseudo-code)

```
Begin Region-growing:
  1. find seed pixel (not labeled and following a row-order)
  2. label it:
     2.1 seed_pixel = i
     2.2 region_average = intensisty (seed_pixel)
  3. explore its eight neighbors (8-connected):
     3.1 list = neighbors(seed pixel)
     3.2 while (length(list) > 0):
             remove first pixel y from L
             if      (euclidean_distance(y,      region_average)<
             distance):
                    add y to L
                    y = i
                    update region average
Begin Region-merging:
  1. find region r with number of pixels < population_minimum
  2. list N = neighbors(r)
  3. explore its neighbors regions looking for the most
  similarity r_min:
     3.1 while (length(N) > 0):
             remove first region r' from N
             r_min = find_minimum(euclidean_distance(r, r'))
  4. merge(r, r_min)
End.
```

The automatically map generated prior to segmentation, is a calibration edge map that can be used to obtain the optimal region maps by means a measure of disparity (Figure 2). This disparity measure is calculated by comparing the maps of distances associated with edges maps of each region maps obtained during the segmentation process, with the map of distances from the calibration edge map. Note that this measure is obtained as described Paglieroni [5]. Consider a region map R and an associated binary border map B in which pixels of value 1 correspond to borders between different regions (specifically, $B(i,j) = 1$ if $R(i,j){\neq}R(i-1,j)$ or $R(i,j-1)$). Let E be the calibration edge map with edge pixels of value 1 on a background of zeros (R, B and E all have the same number of rows and columns). The disparity Δ_{BE} between R and E is given by

$$\Delta_{BE} = 0 \qquad\qquad n_B = n_E = 0$$
$$\Delta_{BE} = 1 \qquad\qquad n_B \text{ or } n_E = 0 \text{ but not both } 0 \qquad\qquad (1)$$
$$\Delta_{BE} = (n_{BE}+n_{EB})/n_B+n_E \qquad n_B , n_E \neq 0$$

In equation 1, n_B is the number of boundary pixels in B, n_E is the number of edge pixels in E, n_{BE} is the number of boundary pixels in B that are not associated with an edge pixels in E, and n_{EB} is the number of edge pixels in E not associated with a boundary pixel in B.

In order to evaluate this proposed segmentation methodology, the analysis of similarity factor (1-disparity factor) and the computation times has been carried out.
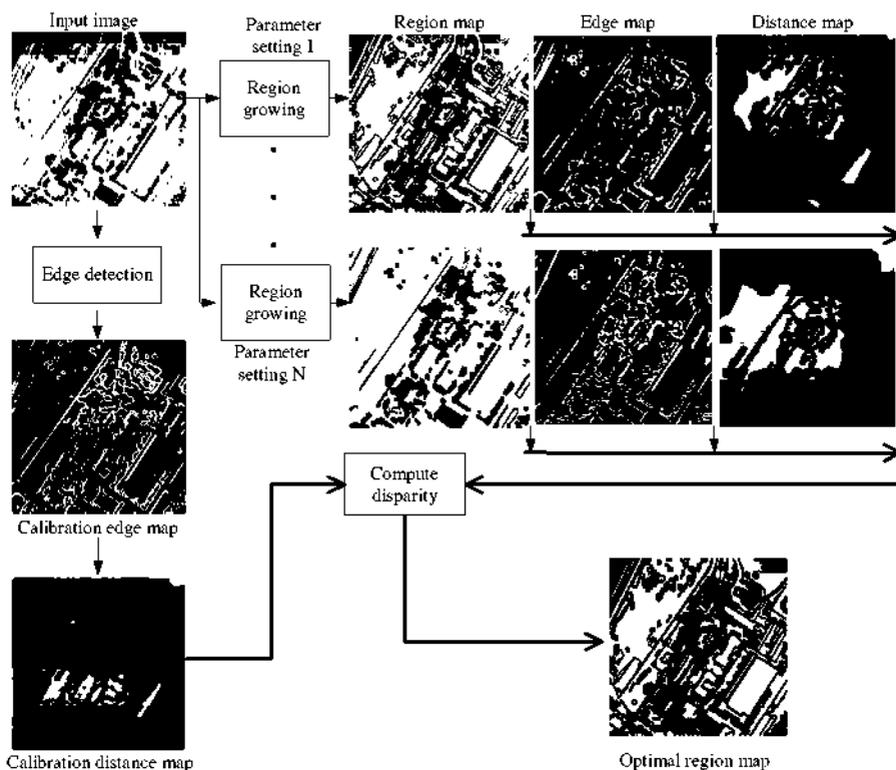
Fig. 2. Methodological scheme of self-calibration framework and segmentation

2.4 Segmentation Process with Definiens Developper Software

The segmentation technique used within Definiens Developer software for this investigation is the multi-resolution segmentation. This technique creates objects using an iterative algorithm, whereby objects (starting with individual pixels) are grouped until a threshold representing the upper object variance is reached. Critical parameters are used here to guide the segmentation result. The Scale parameter determines the maximum allowed heterogeneity for the resulting image objects. The Color criterion defines the weight with which the spectral values of the image layers contributes to image segmentation, as opposed to the weight of the Shape criterion. The relationship between Color and Shape criteria is: Color + Shape = 1. Maximum Color criterion 1.0 results in objects spectrally homogeneous; while with a value of less than 0.1, the created objects would not be related to the spectral information at all. Smoothness is used to optimize image objects with regard to smooth borders, and compactness allows optimizing image objects with regard to compactness [8]. The resulting objects also depend on the image data. For a given set of segmentation parameters, heterogeneous image data result in smaller image objects than homogeneous image data.

The first step in this image segmentation process has been to cluster the image and produce a low-level segmentation for further processing. This strategy permits to optimize the selection of parameters for the multi resolution segmentation (second step). The clustering step was carried out by means of the Iterative Self-Organizing Data Analysis Technique (ISODATA) [9]. Traditionally, ISODATA begins arbitrarily locating a given number of cluster centroids in the feature space. In this case, we have introduced in the process, the initial spectrally supervised centroids. Therefore, the algorithm classifies each pixel of an image into one of these initial clusters. The classification is carried out by assigning each pixel to the nearest cluster in the feature space. A Euclidean distance measure is applied. After the image has been processed, the cluster centers are recalculated as a mean vector of the observations of the cluster. This process continues iterating until one of the user-defined parameters; convergence threshold or the maximum number of iterations is reached. The convergence threshold is the percentage of those observations which remain in a same cluster during two successive iterations.

In the second step, image segmentation (multi resolution segmentation) was performed using the object-based image analysis software Definiens Developper for Earth Sciences [3]. The segmentation approach is a bottom-up region merging process based on heterogeneity of image objects, and controlled by three segmentation parameters: shape(S)/color, compactness/smoothness, and a scale parameter [10]. Shape/Color and compactness(C)/smoothness were set to 0.1/0.9 and 0.5/0.5 respectively, using as a guideline previous research with Supervised ISODATA Clustering (SIC) imagery in this study scenes. We segmented the images (each of the three layers, RGB has an equal weight) several times with different values for the scale parameter (SP) and choosing from 50 to 150 in increments of 50. At scales coarse than 150, some SIC individual class were being merged together, therefore 150 became the cut-off of the coarsest scale.

3 Results

First, the goodness of the new approach can be analyzed by comparing the visual interpretation of segmented images (Fig. 3). This method generates more real world objects or segments (RGB image supervision). The results from Definiens show a major number of segments than the obtained by the proposed methodology and thus an unreasonable segmentation of natural soil land cover is carried out (see Fig. 3c). If the Scale parameter is changed, the number of segments can decrease, but we will not get the proper segmentation of vegetation land covers (trees). This confirms that the selection of segmentation parameters is a critical task and in most cases, it has to be choosing by trial and error. However in this case, a previous low-level segmentation (ISODATA clustering) was carried out, in order to define the guidelines of the segmentation parameters.

On the other hand, the similarity factor obtained from disparity factor is shown as measure of images segmentation quality, in Table 1. These values indicate the goodness of the segmentation methodology proposed in this work.
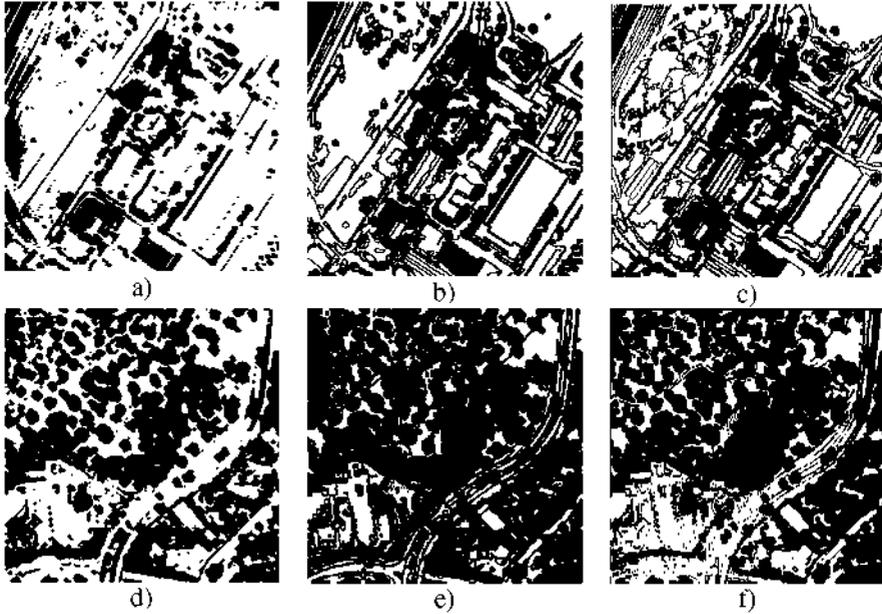
Fig. 3. Comparison of segmented images. a) and d) RGB images. b) and e) Segmented images by means the proposed methology. c) and f) Segmented images by multi resolution segmentation from Definiens Developper Software (SP = 50, S = 0.1 and C = 0.5).

Table 1. Evaluation parameters with a 2.1 GHz processor and 640 MB of RAM

	This work (Execution time / Similarity factor)	Definiens Developper (Execution time)
Scene 1	1.265 s / 88 %	5 s
Scene 2	1.297 s / 93 %	6 s

4 Conclusions

It was designed and implemented our image segmentation methodology for information extraction from remotely sensed imagery. We propose in this paper an unsupervised image segmentation approach with an optimal initialization step based on applying a bidirectional filter as a pre-processing step to eliminate noise, smooth the initial image and preserving edges.

Our segmentation method is validated with several experiments and the results show similar segmentation quality with ones obtained from Definiens Developper software, however in the self-calibration framework, the parameter values are automatically select, and this is a great advantage. Likewise the execution times for our segmentation algorithm are better than the obtained from multiresolution segmentation performed by Definiens Developper software.

In the future work, the possibility of adding new features such as color and texture as well as parallelization of the self-calibration process and the bilateral filter will be studied for optimizing the quality of segmentation and execution time on very large images.

References

1. Blashke, T.: Object based image analysis for remote sensing. ISPRS Journal of Photogrammetry and Remote Sensing 65, 2–16 (2010)
2. Pekkarinen, A.: A method for the segmentation of very high spatial resolution images of forested landscapes. International Journal of Remote Sensing 23, 2817–2836 (2002)
3. http://www.definiens.com/ definiens-developer-for-earth-sciences_180_7_9.html
4. Baatz, M., Schape, A.: Multiresolution Segmentation an Optimization Approach for High Quality Multi-scale Image Segmentation, http://www.ecognition.com
5. Paglieroni, D.: A self-calibrating multi-band region growing approach to segmentation of single and multi-band images. In: SPIE Photonics West, Optical Engineering at LLNL (2003)
6. Tomasi, R., Manduchi, R.: Bilateral Filtering for gray and color images. In: Sixth International Conference on Computer Vision, New Delhi, India (1998)
7. Canny, J.: A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 8, 679–698 (1986)
8. Baatz, M., Benz, U., Dehghani, S., Heynen, M., Holtje, A., Hofmann, P., Lingenfelder, I., Mimler, M., Sohlbach, M., Weber, M., Willhauck, G.: eCognition User's Guide. http://www2.definiens.com/central/default.asp
9. Tou, J.T., Gonzalez, R.C.: Pattern Recognition Principles. Addison-Wesley, Reading (1974)
10. Benz, U.C., Hofmann, P., Willhauck, G., Lingenfelder, I., Heynen, M.: Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry & Remote Sensing 58, 239–258 (2004)