Unsupervised multiscale ROIs determination for supervised thematic classification

Consuelo Gonzalo-Martín[1], Mario Lillo-Saavedra[2], Juan Gómez-Carpintero[3], Ángel M. García-Pedrero[3] and Ernestina Menasalvas[1]

[1]*Universidad Politécnica de Madrid, Center for Biomedical Technology (CTB), Pozuelo de Alarcón, Spain; consuelo.gonzalo/ernestina.menasalvas@upm.es*
[2]*Universidad de Concepción, School of Agricultural Engineering, Chillán, Chile; malillo@udec.cl*
[3]*Universidad Politécnica de Madrid, Computer School, Boadilla del Monte, Spain; juan.gomez-carpintero.garcia&am.garcia@alumnos.upm.es*

Abstract. In this paper, it is proposed an unsupervised methodology based on the Object Image Based Analysis (OBIA) paradigm, for the determination of multiscale training sets (ROIs). This methodology is based on the following hypothesis: the objects selected in an unsupervised way and characterized by certain attributes provide meaningful and reliable training sets for supervised classification. The proposed methodology allows the determination of regions at different scales, adapting the training set (size and number) to land cover characteristics. In order to show the potential and validity of this methodology, regions of interest have been used as input patterns to a nonparametric classifier (Decision Tree) and the results have been compared with classification results obtained when the classifier is trained with a set of training patterns obtained manually. Some initial experiments show that the proposed methodology provides classification results with comparable quality, and in most cases better, than that obtained when the ROIs are manually selected. Moreover, this methodology eliminates routine tasks and operator involvement is limited to make decisions, reducing time, cost and subjectivity in the selection of the ROIs. Therefore, it is expected that a more thorough study of the selection criteria and attributes used for ROIs characterization, will improve the quality of ROIs in terms of the accuracy of the classification results, maintaining the advantages already mentioned.

Keywords. Segmentation, Quickshift, OBIA, training patterns, multiscale, supervised classification, Decision Trees.

1. Introduction

Nowadays, it is not questionable the interest and usefulness of remote sensing data in a large number of applications. Thus, one of the objectives of Horizon 2020 space is the exploitation of space data, in general, and of Earth Observation (EO), in particular. An efficient and effective exploitation of these data (EO) involves the availability of methodologies and tools adapted to the data characteristics. In the last decades, there has been emerging an important amount of satellite imagery with high spatial and/or spectral resolution, providing a huge useful data volume for Earth Observation. Consequently, many papers proposing novel techniques for processing and analyzing high-resolution remote sensed images have arose in the literature, most of them under the umbrella of the OBIA paradigm (Object Image Based Analysis). OBIA methodologies interpret the images based on objects in a similar way than Human Visual System (HVS) does [1]. The HVS capabilities are derived from the perception of objects characterized by different attributes (size, shape, texture, color, among others) and also from the use of structural knowledge and contextual information as well as information about the object shape and the spatial relations between the image regions at different scales [2]. That is specially interesting for the analysis of land covers which require a multiscale analysis which matches the different scales with the sizes of the interest objects in the image

[3, 4, 5]. In this context, the multiscale concept is referred to different spatial dimensions of entities, patterns and processes that can be observed and measured in the image [6]. Some of the advantages of image processing based on the OBIA paradigm against the traditional paradigm based on pixels are: the reduction of the so called "salt and pepper effect" [7], moreover these methods use spatial and contextual information avoiding the low spectral separability problem and finally, they provide an important reduction in the number of data to be processed, since the number of objects is always considerably smaller than the number of pixels [8, 9]. This last aspect implies a reduction of the costs of the processing delay. Moreover, since high-resolution images provide a huge amount of data to be processed, unsupervised techniques that reduce human operator participation in the data exploitation are required.

Perhaps one of the most tedious tasks in supervised classification is the selection of training patterns or regions of interest (ROI). It is well known that the number and quality of a particular training set has a strong influence on the quality of supervised classification results [10]. This task presents some additional difficulties in the case of high spatial resolution imagery. On one hand, it is necessary to capture the whole high variance of the image, considering that the high spatial variance provokes that the training patterns present high heterogeneity intra-class and low spectral separation inter-class [11]. Thus, it supposes a reduction in the statistical separability of the different land-cover classes in the spectral domain, causing poor classification results (“salt and pepper effect”); to reduce these effects, a large number of samples have to be selected, increasing training time.

In this work, it is proposed an unsupervised methodology for the determination of multiscale training sets (ROIs) based on the OBIA paradigm. The fundamental hypothesis is that the objects selected in an unsupervised way and characterized by attributes provide meaningful and reliable training sets for supervised classification. In addition, the proposed methodology allows the determination of regions at different scales, adapting the training set (size and number) to land cover characteristics; and there is a noteworthy data volume decreasing which implied a reduction in computational cost. This methodology is described in the next section.

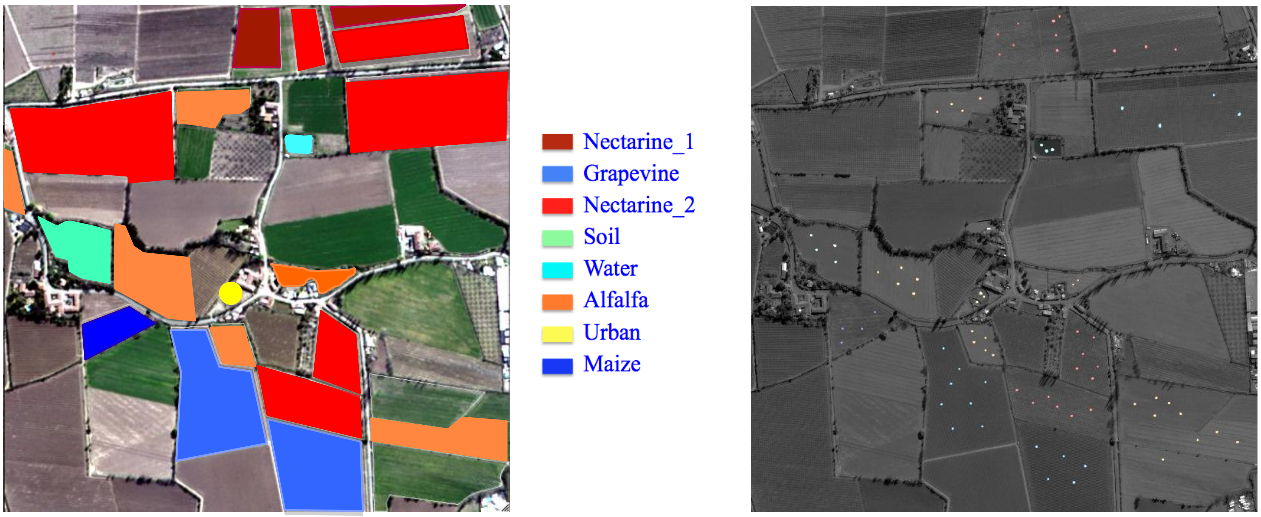
2. Data Set and Methodology

The proposed methodology uses multispectral and panchromatic images, as well as, in situ data utilized to label the ROIs.

2.1. Satellite imagery

Sensor on board of the satellite Worldview-2 has registered the satellite images used in this work. This sensor provides a high-resolution panchromatic band (0.46 m) and eight multispectral bands (1.8 m), at nadir. The multispectral image includes four standard bands (red, blue, green, near-IR) and four new bands (red-edge, coastal, yellow and near-IR2) [12]. A color composition of the scene for the study area has been included in figure 1 a) and the corresponding panchromatic scene in figure 1 b).

The panchromatic image size is 2048x2048 pixels, equivalent to an area of 10,485 ha. (upper left corner coordinates 32º 51’ 7.91” E, 70º 39’ 5.10”W). The record date was September 11, 2011. The area corresponds to a rural zone located at Valparaiso region in Comuna de Los Andes, Chile. No pre-processing of the image was undertaken with actions such as atmospheric correction viewed as unnecessary for the classification of a single image [13]. The use of small training sets) and the agricultural fields were readily identifiable in the imagery removing the need for a geometric correction.



a) b)

Figure 1. a) Color composition of scene under studio with labelled shape vector overlaid. b) Panchromatic image with manual ROIs overlaid.

2.2. In situ data

A supervised visual inspection by Google Earth was carried out previously to the field campaign, for planning it. During the field campaign additional information to the type of coverage was collected: phenological status of the different covers, canopy height and type of planting (frame, row and complete). The covers identified in the study area were 7: generic agricultural land, water bodies, buildings and urban construction, and four different types of crops in different phenological states (nectarine, grapevine, alfalfa and maize). The nectarine has been separated in two different land cover: nectarin_1 and nectarin_2. A survey of the existing land cover types in the study site was conducted by mean recorded data. This information was included in a vector file and used for labelling automatically the ROIs and later as ground-truth data for validating the regions of interest and the classification results. The labelled polygons included in the vector file are overlaid to the original image in figure 1 a).

2.3. Regions of interest definition

Figure 2 shows a scheme of the ROIs determination process at each scale. The segmentation of satellite images is an especially difficult task; due to that the natural land covers exhibit a high shape complexity, randomness and irregularity. In this sense, the determination of the features or attributes in which it is based segmentation is a particularly critical issue. The proposed methodology supports the use of many input datasets, from different sources, as deemed necessary for the analysis and interpretation of the scene. Since each dataset is defined by the spatial and/or temporal distribution of any attribute that can provide useful information about this scene. In this work, input datasets for the segmentation process are all spectral bands of the multispectral image plus the panchromatic image. The multiscale and hierarchical algorithm Quickshift (QS) has been used to segment the multiband data set [14]. QS is a mode seeking algorithm like mean shift that can be used to partition an image into a set of segments. Unlike mean shift, QS does not iteratively shift each point towards a local mode; instead, it simply moves each point to the nearest neighbor for which there is an increment of the density. It has been proved that QS can balance under- and over-fragmentation of the clusters by the choice of a real parameter, moreover, this algorithm is very competitive, resulting in good (and sometimes better) segmentations compared to mean shift, at a fraction of the computation time. The QS implementation provided by [15] depends on three parameters. The ratio (r), which establishes the tradeoff between the weight of color and spatial information; as far as our experience, the segmentation sensitivity of this last parameter is very low. The kernel size (ks) that is the bandwidth of the Parzen window estimator of the density. And the distance (d), it controls the maximum distance between neighbors that should be linked, avoiding the search for the nearest neighbor that increases the density over all pixels. This parameter (d) can be interpreted as a scale parameter. A set of md segments is obtained for each scale (j) or each value of d. A multiscale and hierarchy segmentation of the image is obtained for a fixed ks value, and different d values.

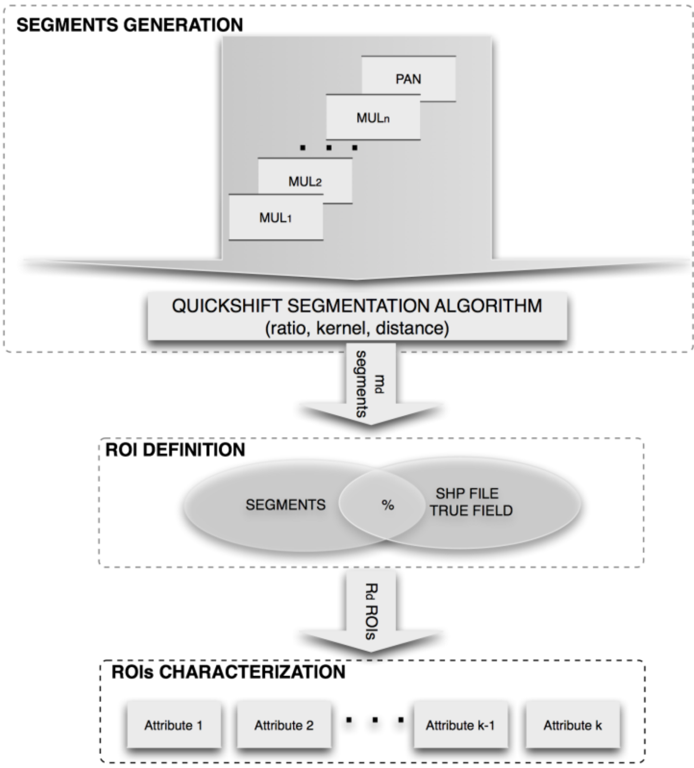


Figure 2. Methodology schematics for the determination of the ROIs at each scale.

A shape file labeled with field data has been used to select the segments employed as ROIs at each scale. Each segment that belongs to a particular shape in a ratio upper a threshold is validated as a ROI of the kind of the land cover determined by the corresponding shape. The validated segments will be characterized by a set of features or attributes selected for discriminating the different land covers (spectral bands, spectral indices, textural, among others). This set depends on the end-use of the information. In this way, multiscale-labelled training patterns are obtained, whose homogeneity the segmentation process ensures. These training patterns will be called ROIs_OBIA_d, where d represents the value of parameter d. From here, it will be considered that each value of d corresponds with a different scale (j).

2.4. Scale determination

Once the ROIs at each scale j (ROIj, j=1… m) have been determined, the next step is to find “the best scale” for each kind of land cover presented in the scene under study. In this paper, the best scale is considered to be that which provides the higher accuracy in a classification process. This process starts with the ROIs at the finest scale (j=1). The ROIs at all scales are randomly separated in a training set and a test set. Experiments with different percentages for the separation have been carried out. A Decision Tree classifier for each scale is training with the corresponding training set. The goodness of the classifiers are assessed by using the test set for each scale and obtaining a different value of the indicator I_i^j for each land cover (i=1, … number of land covers) at each scale. Two conditions are evaluated to determine if a coarse scale is better than the previous finer scale.

$$\text{If } I_i^j > th \text{ or } I_i^j > I_i^{j-1} \text{ then scale}=j \qquad (1)$$

Where th represents a threshold value of this indicator that has to be previously established by the user. For those land covers that meet one of the two conditions, it is considered that the current scale j is better than the previous one. When the process is over, we will have different scales for every land cover such that we will keep a high rate of quality while we keep the number of ROIs to the minimum. The first condition will assure that we will use the lower number of ROIs in order to improve the performance of the classification.

Then, the multi-scale ROIs set is obtained by the union of the ROIs at different scales for each land cover. In that cases, in which a particular land cover cannot be associated with a scale, that means finer scales are needed.

2.5. Validation of methodology

With the aim of validate the methodology proposed, the multiscale ROIs obtained has been used for training Decision Tree classifiers. The goodness of the classifiers has been assessed by a standard methodology [16]. For that the confusion matrix has been calculated for each classification experiment, as well as, the mean producers accuracy (PR Mean= # of pixels correctly classified as one particular land cover) / (# ground reference pixels in this land cover), the mean user accuracy (US Mean=# of pixels correctly classified as one particular land cover) / (total # of pixels classified as this particular land cover) and overall accuracy (# pixels correctly classified) / (total # of pixels).

The results have been compared with the outcomes provided by classifiers trained with ROIs selected manually and ROIs selected by the proposed methodology at different scales.

3. Results

The input to the segmentation algorithm (QS) has been the 8 spectral bands of the Worldview-2 image, as well as the corresponding panchromatic scene. Segmentation results were obtained for the following values of the QS algorithm parameters: r=0.7 and ks and d = 5, 10, 15, 20, 25 and 30. After the analysis of different segmentations, the ks parameter has been fixed to the value 10, which provides a mean segment size for the lower value of d adequate for the scene under study. The percentage established to determine the belonging of each segment to a particular shape has been 90%. Even though a large number of experiments have been performed trying to determine the best set of attributes with some kind of physical sense, no conclusive results have been obtained yet. Therefore it has been used as attributes the Worldview-2 spectral bands recommended for vegetation analysis (Blue, Green, Yellow, Red Edge, NIR2), as well as the NDVI defined for this kind of images. As result of the methodology described in section 2.3, ROIs_OBIA for 6 different scales (r=0.7, k=10, d=5, 10, 15, 20, 25 and 30) have been obtained. The number of ROIs determined for each scale is displayed at table 1 (only three values of d have been included in the table, 5, 15 and 30, from column 2 to 4). Each set of ROIs_OBIA_d has been randomly separated in a training set and a test set. Experiments with different percentage for the separation have been carried out. The results included in this work correspond to 70% of the samples for the training set and 30% for the test set, since these values provide a good trade-off between the number of samples used for each process. Fifteen Decision Trees have been trained for each training set and the average behavior considered. Finally the confusion matrix has been obtained for each experiment, as well as the mean producers, the mean user accuracy and the overall accuracy.

A set of multiscale ROIs has been obtained by the methodology described in section 2.4. For this case de number of ROIs has been 2164, as it is displayed at table 1. Table 2 includes the d value determined for this methodology for each kind of land cover. Fifteen Decision Trees have been trained with these ROIs, in the same conditions than for the ROIs_OBIA case. And the classification accuracy has been assessed.

Table 1. Data number to be processed in the different classification experiments.

Manual ROIS (pixels)	ROIS_OBIA_5	ROIS_OBIA_15	ROIS_OBIA_30	Multi-scale ROIs
6125	9238	517	133	2164

Using the ROIs tool provided by ENVI 5.0 has carried out the manual selection of the ROIs. After the initial selection, a refinement process has been performed in order to improve ROIs statistical values. A total of 6125 data have been used for training in this case (figure 1 a). For the sake of comparison, these ROIs have been characterized with the same attributes that the ROIs_OBIA and the multiscale ROIs; and they have been used as training and set patterns for Decision Trees in the same way that in the previous experiments. The reduction in the data number to be processed implies a decreasing of computation time.

Table 2. Values of *d* (“best scale”) determined for each land cover

Nectarine_1	Parronal	Nectarine_2	Agricultural land	Water	Alfalfa	Urban	Maize
15	25	25	20	5	5	5	10

Table 3 summarizes the values obtained for the mean producers accuracy, the mean user accuracy and overall accuracy for classifiers trained with ROIs selected manually (first column), with ROIs_OBIA_d (d=5, 15 and 30) and with multiscale ROIs (last column).

Table 3. Decision Tree Classification assessment obtained for different type of ROIs.

	Manual ROIs	ROIs_OBIA_5	ROIs_OBIA_15	ROIs_OBIA_30	Multi-scale ROIs
PR Mean (%)	87	76	63	51	64
US Mean (%)	87	80	60	55	66
Overall Accuracy (%)	89	96	94	95	97

The overall accuracy obtained for the distinct sets of ROIs indicates that the automatic selection of the ROIs, always provided better classification results that the ROIs selected manually, even though the quality of these is proved by the high values obtained in the three indicators. It can be appreciated that the overall accuracy decreases as the value of d increases. This fact is due that higher values of d provide coarser ROIs, decreasing the homogeneity; however, when different scales (or values of d) are used to sample different land covers, the overall accuracy gets the higher value. A similar behavior is showed for producer and user accuracy, except for the manual ROIs.

After the determination of the accuracy indicators, the different classifiers have been used to classify the whole images, pixel by pixel. Classified images are included in figure 3 with the aim to carry out a visual comparison. It can appreciate the similitude between the classified image using manual ROIs (figure 3 a) and using ROI_OBIA_5 (figure 3 b). In both the 8 classes are discriminated; however, as the value of d increases (figure 3 c and d) some classes are almost losing, as for example urban, and water, that is the classes that cannot be represented at the corresponding scales. These two classes are discriminate adequately when ROIS_MS are using for training. Another important aspect to be considered is the noise reduction for higher values of d and especially for ROIS_MS.

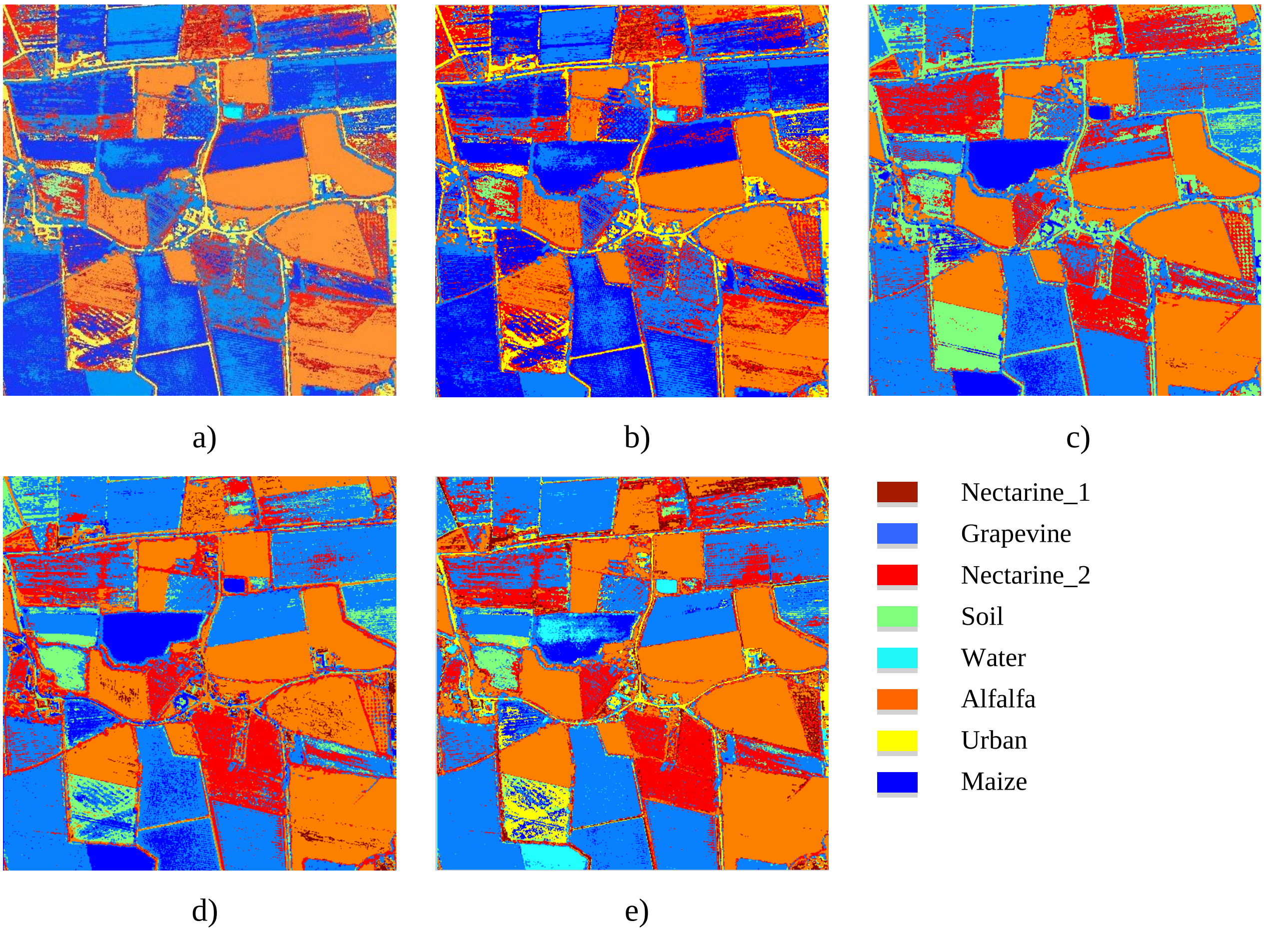


Figure 3. Images classified by Decision Trees with different sets of ROIs. a) Manual ROIs. b) ROIs OBIA_5. c) ROIS_OBIA_15. d) ROIs_OBIA_30. e) Multiscale ROIs

4. Conclusions

The hypothesis of this work was that the objects selected in an unsupervised way and characterized by certain attributes provide meaningful and reliable training sets for supervised classification.

Some initial experiments, with simple criteria for the selection of the multiscale ROIs and different sets of spectral attributes for their characterization have supported this hypothesis. The proposed methodology allows the determination of ROIs at different scales, adapting the training set (size and number) to the characteristics of each of the land cover presented in the scene under study. Moreover, eliminates routine tasks and operator involvement is limited to make decisions, reducing time, cost and subjectivity in the selection of the ROIs. Classification results with comparable quality, and in some cases better than that obtained when ROIs are manually selected, has been obtained with multiscale ROIs. Therefore, it is expected that a more thorough study of the selection criteria and attributes used for ROIs characterization, will improve the quality of ROIs in terms of the accuracy of the classification results, maintaining the advantages already mentioned. Moreover, the methodology should be validated for different type of images.

Acknowledgements

This work has been funded by the Universidad Politécnica de Madrid (AL13-PID-07) and Fondo de Fomento al Desarrollo Científico y Tecnológico (FONDEF- D09I1069) de Chile.

References

[1] Blaschke, T., Lang, S. y Hay, G. J. (editors), 2008. *Object-Based Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*, Springer Publishing Company, Incorporated, 1 edition.

[2] Blaschke, T. 2011. *Object based image analysis for remote sensing*, ISPRS Journal of Photogrammetry and Remote Sensing, 65, 2-16.

[3] Baatz, M. and Schäpe, A., 2000. *Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation*. Journal of Photogrammetry and Remote Sensing, 58,12–23.

[4] J. L. Dungan, J. N. Perry, M. R. T. Dale, P. Legendre, S. Citron-Pousty, M.-J. Fortin, A. Jakomulska, M. Miriti and M. S. Rosenberg, 2002. *A balanced view of scale in spatial statistical analysis*, Ecography, 25, 626-640.

[5] Hay, G. J., and Marceau, D. J., 2004. *Multiscale Object-Specific Analysis (MOSA): An integrative approach for multiscale landscape analysis*. In S. M. de Jong and F. D. van der Meer (Eds). Remote Sensing and Digital Image Processing. Volume 4. Chapter 3. Kluwer Academics.

[6] Hay, G. J., Castilla, G., Wulder, M. A. and Ruiz, J. R., 2005. *An automated object-based approach for the multiscale image segmentation of forest scenes*. International Journal of Applied Earth Observation and Geoinformation, 7, 339-359.

[7] Blaschke, T. and Strobl, J. 2001. *What's wrong with pixels? Some recent developments interfacing remote sensing and GIS*. GIS, 6, 12-17.

[8] Castilla, G. and Hay, G.J., 2008. *Image objects and geographic objects. In Object-Based Image Analysis*, T. Blaschke, S. Lang and G.J. Hay (Eds.), 91–110 (Berlin: Springer).

[9] Blaschke, T. and Lang, S., 2006. *Object based image analysis for automated information extraction−a synthesis*. In Measuring the Earth II ASPRS Fall Conference (San Antonio, Texas), 6−10.

[10] Lillesand, T. M., Kiefer, R. W. and Chipman, J. W., 2008. *Remote Sensing and Image. Interpretation*. 6th ed. (Hoboken: John Wiley&Sons).

[11] Peña-Barragan, J. M., Ngugi, M. K., Richard, E. y Six, J. 2011. *Object-based crop identification using multiple vegetation index, textural features and crop phenology*, Remote Sensing of Environment, 115, 1301-1316

[12] Digital Globe, http://worldview2.digitalglobe.com /docs/WorldView-2_8-Band_Applications_ Whitepaper.pdf

[13] Song, C., Woodcock, C. E., Seto, K. C., Pax Lenney, M., & Macomber, S. A., 2001. *Classification and Change Detection Using Landsat TM Data. When and how to correct for atmospheric effects?*, Remote Sensing of Environment, 75, 230−244.

[14] Vedaldi, A. and Soatto S., 2008. *Quick Shift and Kernel Methods for Mode Seeking*. ECCV, Part IV, LNCS 5305, pp. 705–718, Forsyth, P. Torr, and A. Zisserman (Eds.): Springer-Verlag Berlin Heidelberg.

[15] Vedaldi, A. and Fulkerson, B., 2008. VLFeat: An Open and Portable Library of Computer Vision Algorithms (http://www.vlfeat.org/)

[16] Conalton, R. G., 1991. *A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data*, Remote Sensing Environment, 37, 35-46.