



A supervised contextual classifier based on a region-growth algorithm

Jorge Lira^{a,*}, Gabriela Maletti^b

^a*Instituto de Geofísica-UNAM, Circuito Institutos, Cd. Universitaria, 04510 México D.F., Mexico*

^b*Informatics and Mathematical Modelling, Technical University of Denmark, DK-2800 Kgs Lyngby, Denmark*

Received 1 May 2001; received in revised form 15 February 2002; accepted 18 February 2002

Abstract

A supervised classification scheme to segment optical multi-spectral images has been developed. In this classifier, an automated region-growth algorithm delineates the training sets. This algorithm handles three parameters: an initial pixel seed, a window size and a threshold for each class. A suitable pixel seed is manually implanted through visual inspection of the image classes. The best value for the window and the threshold are obtained from a spectral distance and heuristic criteria. This distance is calculated from a mathematical model of spectral separability. A pixel is incorporated into a region if a spectral homogeneity criterion is satisfied in the pixel-centered window for a given threshold. The homogeneity criterion is obtained from the model of spectral distance. The set of pixels forming a region represents a statistically valid sample of a defined class signaled by the initial pixel seed. The grown regions therefore constitute suitable training sets for each class. Comparing the statistical behavior of the pixel population of a sliding window with that of each class performs the classification. For region-growth, a window size is employed for each class. For classification, the centered pixel of the sliding window is labeled as belonging to a class if its spectral distance is a minimum to the class. The window size used for classification is a function of the best separability between the classes. A series of examples, employing synthetic and satellite images are presented to show the value of this classifier. The goodness of the segmentation is evaluated by means of the κ coefficient and a visual inspection of the results.

© 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Region growing; Context classification; Homogeneity criterion; Spectral; Distance; Spectral separability

1. Introduction

Segmentation is a partition of the image into a number of regions (Cohen and Fam, 1992), each region related to class objects of the scene. The regions may be labeled as pertaining to a certain class of objects, hence generating a classification. The final product is therefore a thematic map useful for scene understanding. The first classifiers labeled the pixels of the image into classes using only their spectral properties and ignoring their context. This approach, named per-pixel classification, proved to be

limited in nature and applicable only to spectrally well-differentiated cases. In the last years, efforts have been devoted to develop contextual classifiers (Arai, 1993; Khazenie and Crawford, 1990; Gong and Howarth, 1992; Kontoes and Rokos, 1996). Classifiers that incorporate contextual information into the classification have been reported in the literature as well (Chica-Olmo and Abarca-Hernández, 2000; Atkinson and Lewis, 2000). A contextual classifier consistently produces higher classification accuracies than the per-pixel classifier (González Alonso et al., 1991; Kontoes and Rokos, 1996; Stuckens et al., 2000). In the approach of this paper, a pixel is labeled to a class by taking into account its spectral properties and the context of its location.

A supervised classifier employs a priori information of each determined class; this is usually done by means of

*Corresponding author. Tel.: +52-55-5622-4211; fax: +52-55-5550-2486.

E-mail addresses: lira@tonatiuh.igeofcu.unam.mx (J. Lira), gmm@imm.dtu.dk (G. Maletti).

training sets. These training sets are defined through closed polygons outlined on the image by some interactive procedure. In this definition, there is not a clear criterion to assume that the training sets are valid statistical samples of the classes (Cetin and Levandowsky, 1991). However, some work on the purification of training samples has been reported that improves the quality of classification (Buchheim and Lillesand, 1989; Bolstad and Lillesand, 1991; Arai, 1992). Region-growth-based algorithms for image segmentation and region classification have been reported in literature Raafat and Wong (1988) and Gahegan and Flack (1999). However, further work is required to use an optimized region-growth algorithm as a preliminary step in a contextual classification scheme.

As a basic premise to a classifier, a procedure should be established to assure that the training sets are representative samples of the classes. A second premise is that the classification of a pixel should be performed by direct comparison between the statistical behavior of the classes and that of the pixel-centered window. A third premise is to make no assumption on the particular statistical behavior of the density function of the classes. This model-based approach yields the automated determination of the set of parameters handled by the classifier; the only exception is the locations of the initial seeds.

In this work, a new contextual classifier is proposed that determines statistical samples of defined classes as a result of an automated region-growth algorithm. A pixel is then classified by comparing the density function of the pixel-centered window population to those of the classes. The comparison is done by means of a measure of similarity between such density functions. The growth of each region is performed by employing a window size and a threshold value suitable for each class. The measure of similarity is utilized to calculate best values for such parameters. The window used for classification is a function of the best separability between the classes. As explained in the next sections, this scheme of classification is valid for a spectral contextual classifier, even though extension to a textural classifier is easy. In brief the goals of this research are the following: (a) develop a new supervised contextual classifier to segment an image into a number of classes, (b) derive the statistics of the classes from a region-growth algorithm, (c) determine the goodness of such classifier.

2. Contextual classifier

2.1. Region-growth scheme

The contextual classifier uses the training sets determined by an automated region-growth algorithm

(Lira and Frulla, 1998). This algorithm begins by seeding pixels in suitable places of the image where the existence of a class is known. This task is done manually by visual inspection of the image with the support of ancillary data. Once the seeds are determined, one per class, the growth of the class regions starts. The growth is performed by pixel aggregation satisfying a homogeneity criterion. The criterion is evaluated in a window with the best size for each class. A pixel is aggregated into the region provided the difference between the homogeneity value of the seed- and the pixel-centered window does not exceed a certain threshold. The growth of a region is terminated when this homogeneity criterion is no longer satisfied. The homogeneity criterion and the threshold are both derived from a measure of separability and heuristic criteria; details of this are given in the next paragraphs.

Let $g(\mathbf{r})$ be a multi-spectral image and $\mathbf{p}_{ij}^0 \in g$, and let $\{R_0 = \mathbf{p}_{ij}^0\}$ be the initial sub-region signaling a given class. The pixel \mathbf{p}_{ij}^0 is known as the seed related to R_0 . The vector \mathbf{r} defines the coordinates of the pixel \mathbf{p}_{ij}^0 in the image, i.e. $\mathbf{r} = (i, j)$. In general, the pixel \mathbf{p}_{ij}^0 is formed by the tuple $\{p_1, p_2, \dots, p_\gamma\}_{ij}$, where γ is the number of bands of the multi-spectral image. Let R'_0 be the set of pixels that do not belong to R_0 but having at least a neighbor with R_0 under certain connectivity. Let $E(R_0)_v$ be the value of the homogeneity criterion applied to the window v of R_0 . The set R_1 is the region jointly formed by R_0 and the pixels $\mathbf{p}_{kl}^1 \in R'_0$ for which $E(\mathbf{p}_{kl}^1)_v$ differs from $E(\mathbf{p}_{ij}^0)_v$ in less than a threshold ε . In other words, R_1 is the following set:

$$R_1 \equiv \{\mathbf{p}_{kl}^1 : |E(\mathbf{p}_{kl}^1)_v - E(\mathbf{p}_{ij}^0)_v| \leq \varepsilon\}. \quad (1)$$

The real number ε is known as the parameter of uniformity, directly related to the homogeneity criterion. Once R_1 has been determined the previous step is repeated; so, in general the region R_m is given by

$$R_m \equiv \{\mathbf{p}_{kl}^m : |E(\mathbf{p}_{kl}^m)_v - E(\mathbf{p}_{ij}^0)_v| \leq \varepsilon\}. \quad (2)$$

The homogeneity criterion E is always tested facing the original sub-region window of R_0 . The growth of a region continues until no change occurs from one step to the next: $R_{k+1} = R_k$. The above is easily generalized for a number of initial regions, i.e. for a set of classes. Thus, aggregation of a tested pixel \mathbf{p}_{kl} is carried out as

$$\mathbf{p}_{kl} \rightarrow \text{region } t : |E(\mathbf{p}_{kl})_v^t - E(\mathbf{p}_{ij}^0)_v^t| \leq \varepsilon, \quad (3)$$

where $E(\mathbf{p}_{kl})_v^t$ is the homogeneity criterion applied to the tested pixel, and the region t is identified as the training set of class t .

2.2. Estimation of region-growth parameters

The best values for the window v and the threshold ε are obtained as follows. An odd-sized window is assumed for each seeded pixel. Beginning from $v = 3$

pixels, the windows are systematically incremented in size. Only squared odd windows are considered. The density function is estimated. Let $\mathcal{S}_v^a(\mathbf{i})$ be the density function of the pixel population in a window v for spectral class a . Such a function is normalized to unity according to

$$\sum_{\mathbf{i}} \mathcal{S}_v^a(\mathbf{i}) = 1. \tag{4}$$

The above is assuming 256 quantization levels in the image, and \mathbf{i} is a vector with dimension equal to the number of bands of the multi-spectral image. Let a be any spectral class in the image, then

$$\delta_{\Delta v}^a = \sum_{\mathbf{i}} |\mathcal{S}_v^a(\mathbf{i}) - \mathcal{S}_{v+2}^a(\mathbf{i})|. \tag{5}$$

On the grounds of extensive tests with several images, the best window for class a is set whenever $\delta_{\Delta v}^a \leq 0.3$. A similar condition is established for the set of classes. The size of the window must satisfy a two-fold condition: large enough for statistical validity, and small enough for maximum spatial resolution. In this sense, the window size is adjusted according to class heterogeneity (Hodgson, 1998). A homogeneous class requires a small window size; a heterogeneous class requires a larger window size. Therefore, a window size is determined for each class. As v increments, the density function of a class experiments a great variation, until the number of pixels is enough to represent a valid statistical sample of the class. While the window size is further incremented, the density function is approximately constant as long as the window is entirely included in the class. If the window population is a mixture of two or more classes, the density function changes again. The condition $\delta_{\Delta v}^a \leq 0.3$ is a heuristic rule to measure such change.

The second parameter handled in the region-growth algorithm is a threshold ε , named the uniformity parameter. This parameter is set in the following way. Let d_v^{ab} be the distance between any two classes a and b for a given window size v :

$$d_v^{ab} = \sum_{\mathbf{i}} |\mathcal{S}_v^a(\mathbf{i}) - \mathcal{S}_v^b(\mathbf{i})|, \quad \forall a \neq b. \tag{6}$$

The distance d_v^{ab} is of spectral separability. Let a be a spectral class in the image; thus, given the set $\{v_a, v_b, \dots, v_u\}$ of windows of the classes, the following minimum is determined:

$$d^a = \min \left[d^{ab} = \sum_{\mathbf{i}} |\mathcal{S}_{v_a}^a(\mathbf{i}) - \mathcal{S}_{v_b}^b(\mathbf{i})| \right], \quad \forall a \neq b. \tag{7}$$

On the grounds of extensive tests with several images, the threshold corresponding to the class j is given by the following criterion:

$$e^j = 0.75d^j \tag{8}$$

and so on for the set of window classes: $\{v_a, v_b, \dots, v_u\}$.

The region-growth is performed on the grounds of a homogeneity criterion defined in terms of the distance determined by Eq. (6). Thus, given the best window for class j :

$$\mathbf{p}_{kl} \rightarrow \text{class } j : \sum_{\mathbf{i}} \left| \mathcal{S}_{p_{kl}}^j(\mathbf{i}) - \mathcal{S}_{p_0}^j(\mathbf{i}) \right| \leq \varepsilon, \tag{9}$$

where \mathbf{p}_0 is the pixel seed for class j . In brief, Eqs. (5), (8) and (9) determine the best window, the threshold and the homogeneity criterion, respectively. This means that, once the seed pixels are determined, the classification process is fully automated.

2.3. Rationale of spectral classifier

The contextual classifier uses a pixel-centered window to estimate the density function associated to a pixel. This function is compared to the density function of the classes determined by the training classes in the region-growth process. A pixel is classified to that class where the comparison is the best. However, a tested pixel window may share pixels from two or more classes. In this situation the density function is a composite of various classes. The pixels having windows located in more than one class are named border pixels. To avoid misclassification, the border pixels must be treated separately, labeled as such, and classified to the nearest class in a separate process. Once classified, the border pixels are encoded with the rest of the pixels to form the final classification of the image. Border pixels are identified as follows: The pixel-centered window is displaced on every pixel of the image. In each position of the window, the density function is obtained. This function is compared with the density function of the classes and to the density function of pair of classes. The result of this comparison leads to the labeling of a border pixel; details of this are in the classifier algorithm steps (v) and (viii).

The rationale of the spectral classifier consists of three stages: (1) construct a model of spectral separability to calculate the best window to perform the classification, (2) design an overall block diagram of the classifier, and (3) list the details of the algorithm for classification. The construction of such a model now follows.

A set of 36 images with 6 classes is generated. The density function of the classes is Rayleigh-like (Dougherty, 1999). The size of the images is 256×192 pixels. The difference of the mean among these classes ranges in the interval: $[1, 2, \dots, 32]$. The difference of the standard deviation ranges in the same interval. This produces a set of 216 classes with varying spectral separability among them. The normalized density function of each class is obtained according to Eq. (4). In each class, a pixel-centered window is considered. This window is systematically incremented from 1×1 to 23×23 pixels. The distance between the density function of this

Table 1
Separability as a function of window size

Window	Separability	Window	Separability	Window	Separability
1	1.9997–2.0000	7	1.1606–1.6299	17	0.3406–0.3746
3	1.9187–1.9997	11	0.6772–1.1606	21	0.2293–0.3405
5	1.6299–1.9186	15	0.3746–0.6771	23	0.2193–0.2293

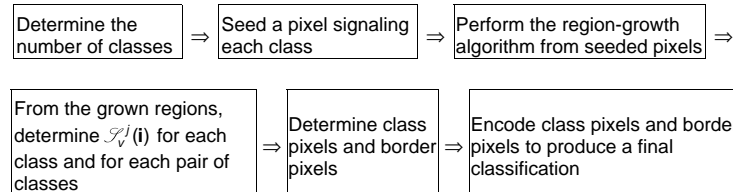


Fig. 1. Block diagram of contextual classifier.

window and the one from each other class is derived (Eq. (6)). A pixel-centered window is classified into a class where this distance is the least. The precision of classification is calculated for each window size. The best window is the one yielding the highest precision. Table 1 condenses this model of spectral separability. In this model, the following is observed:

- (i) Whenever the separability is complete, the maximum precision is obtained using a per-pixel classification, i.e. a 1×1 pixel window.
- (ii) Classes with similar separability have a similar window size for classification.
- (iii) An inverse relationship exists between the separability of the classes and the best window size.

The overall block diagram of the classifier is given in Fig. 1. According to this, the details of the classifier algorithm are the following:

- (i) Let $g(\mathbf{r})$ be the image to be classified. The bands selected for classification are loaded into RAM memory. Decorrelated bands resulting from principal component analysis are usually employed in this step.
- (ii) Pixels are seeded in selected places of each spectral class defined for segmentation. Let $\xi \equiv \{a, b, \dots, u\}$ be the set of spectral classes.
- (iii) The best window is derived for each class according to Eq. (5). Let $\eta \equiv \{v_a, v_b, \dots, v_u\}$ be the set of windows.
- (iv) On the grounds of the region-growth algorithm, the normalized density function $\mathcal{S}_{v_j}^j(\mathbf{i})$ for each training set is obtained $\forall j \in \xi$.
- (v) A matrix of the average of density functions for a pair of classes is constructed. Let $\mathcal{S}^{ts}(\mathbf{i})$ be the

elements of such a matrix for any two spectral classes t and s , $\forall t \neq s$. The function $\mathcal{S}^{ts}(\mathbf{i})$ is named as the density function of a border formed by a pair of distinct classes.

- (vi) The distance $d^{ab} = \sum_{\mathbf{i}} |\mathcal{S}_{v_a}^a(\mathbf{i}) - \mathcal{S}_{v_b}^b(\mathbf{i})|$, $\forall a \neq b \in \xi$ is obtained as a function of class window v_i , $\forall v_i \in \eta$. For classification, the best window v is obtained from Table 1 using $d = \min[d^{ab}]$, $\forall a \neq b \in \xi$.
- (vii) For every pixel \mathbf{p}_{kl} to be classified, the normalized density function $\mathcal{S}_v^{kl}(\mathbf{i})$ is generated, where (k, l) are the coordinates of the pixel in the image: $\mathbf{r} = (k, l)$, and v is the best window obtained in step (vi).
- (viii) A pixel \mathbf{p}_{kl} of the multi-spectral image is classified according to the following procedure:

- (a) The distance between the density function $\mathcal{S}_v^{kl}(\mathbf{i})$ of the pixel-centered window and the density function of the classes $\mathcal{S}_{v_j}^j(\mathbf{i})$ is calculated

$$d_{kl}^j = \sum_{\mathbf{i}} |\mathcal{S}_v^{kl}(\mathbf{i}) - \mathcal{S}_{v_j}^j(\mathbf{i})|, \quad \forall k, l \in \mathbf{g}, l = \forall j \in \xi. \tag{10}$$

- (b) The distance between the density function $\mathcal{S}_v^{kl}(\mathbf{i})$ of the pixel-centered window and the density function of the pair of classes $\mathcal{S}^{ts}(\mathbf{i})$ is calculated

$$d_{kl}^{ts} = \sum_{\mathbf{i}} |\mathcal{S}_v^{kl}(\mathbf{i}) - \mathcal{S}^{ts}(\mathbf{i})|, \quad \forall k, l \in \mathbf{g}, \forall (t, s) \in \xi. \tag{11}$$

- (c) The distances d_{kl}^j and d_{kl}^{ts} are sorted out by value. Let $d = \{d_{kl}^j, d_{kl}^{ts}\}_v$ be the ordered set of such distances.

- (d) If $\min \{d_{kl}^l, d_{kl}^{ts}\}_v$ is a distance to a border, a new window v' for \mathbf{p}_{kl} is considered. The size of this window is the one providing the best separability (Table 1) between the density function of such border and the density function corresponding to the second smallest distance of δ .
- (e) A new set of distances $\delta' = \{d_{kl}^l, d_{kl}^{ts}\}_{v'}$ is calculated for v' . If $\min \{d_{kl}^l, d_{kl}^{ts}\}_{v'}$ is a distance to the border found in step (d), then $\mathbf{p}_{kl} \rightarrow \text{border}$.
- (f) If $\min \{d_{kl}^l, d_{kl}^{ts}\}_v$ is the distance to a class t , then $\mathbf{p}_{kl} \rightarrow \text{class } t$.
- (g) If $\mathbf{p}_{kl} \rightarrow \text{border}$, its average density function is considered. Let this border be the one formed by classes q and r . The following distances are calculated:

$$d_{kl}^l = \sum_i |\mathcal{S}_v^{kl}(\mathbf{i}) - \mathcal{S}_v^l(\mathbf{i})|, \quad \forall k, l \in g, \quad l = (q, r) \in \xi. \quad (12)$$

Then \mathbf{p}_{kl} is labeled as border-class q if $d_{kl}^q = \min \{d_{kl}^l\}$, $l = (q, r) \in \xi$.

This algorithm generates a segmentation of the image in classes and borders. The encoding of class pixels and border pixels produces a final classification of the image as shown in Section 4. This algorithm is not computationally demanding: the density functions and the images are loaded into the RAM memory, the whole procedure for the multi-spectral classification presented in this research takes approximately 2 min on a Pentium

IV @ 700 MHz. The following section describes the test images used in the classifier.

3. Test images

Two types of examples are presented in this paper. These examples are worked out on the grounds of:

- (a) A set of two synthetic images (Figs. 2a and 3a) with well-known statistical parameters for each class. These synthetic images are single band and contain five (Fig. 2a) and three (Fig. 3a) barely discernible classes. The class separation, in both images, is five gray levels between means. In Fig. 2a, each class has a variance of 11.26. In Fig. 3a, each class has a variance of 24.92. The dimension of these images is 128×128 and 96×96 pixels, respectively. The density function of the classes in both the images is Gaussian-like. Fig. 2a shows an image with low content of noise: $\text{snr} = 7.9586 \text{ dB}$. Fig. 3a depicts a relatively high level of noise: $\text{snr} = 0.3719 \text{ dB}$. The formula to calculate the noise is

$$\text{snr} = \frac{1}{N_C - 1} \sum_{i=1}^{N_C} 20 \log \left[\frac{\mu_{i+1} - \mu_i}{\sigma} \right], \quad (13)$$

where N_C is the number of classes, μ_i and σ are the mean and the standard deviation of the classes,

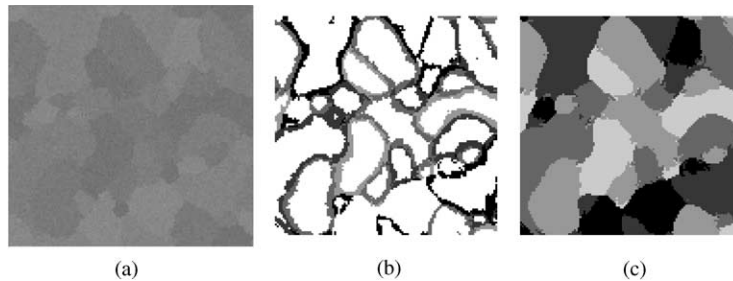


Fig. 2. (a) Synthetic image with five classes and signal-to-noise ratio of 7.9586 dB, (b) border pixels of synthetic image depicted in Fig. 2a, (c) encoding of border pixel and class pixels of the synthetic image depicted in Fig. 2a.

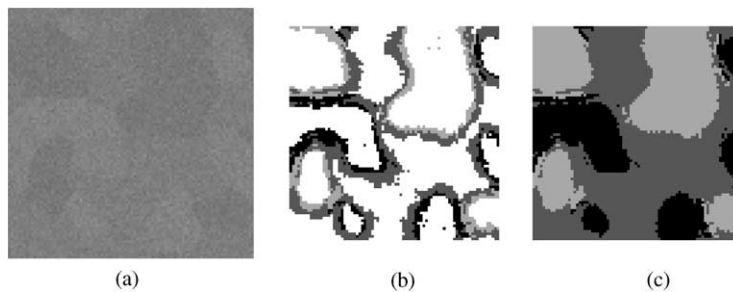


Fig. 3. (a) Synthetic image with three classes and signal-to-noise ratio of 0.3719 dB, (b) border pixels of the synthetic image depicted in Fig. 3a, (3) encoding of border pixel and class pixels of synthetic image depicted in Fig. 3a.

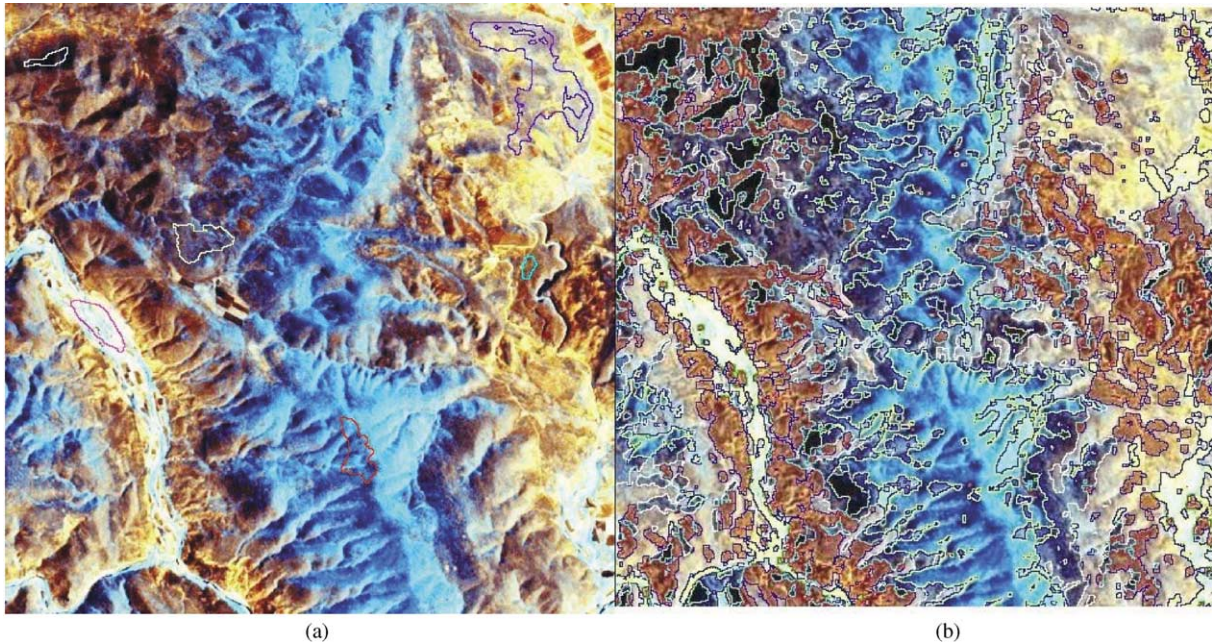


Fig. 4. (a) RGB composite of SPOT image overlain with training fields defined by region-growth, (b) RGB composite of the SPOT image overlain with borders of spectral classes defined in classification process.

respectively. The means are sorted by the value:
 $\mu_{i+1} \geq \mu_i$.

- (b) A satellite SPOT multi-spectral (Fig. 4a) image. The dimension of this image is 512×512 pixels, covering a region of Central México. This image is not geocoded.

4. Results and discussion

Results are organized in two categories: for the synthetic images final results are provided. For the SPOT image full details are given.

The synthetic images are considered an extreme case to test the validity of the classification scheme. Fig. 2b shows the border pixels of Fig. 2a, while Fig. 2c shows the encoding of class and border pixels with the final segmentation achieved. The finite size of the window employed in this process reduces the size of the images by $(v-1)/2$ pixels per side. The border pixels and the segmentation achieved for Fig. 3a are presented in Fig. 3b and 3c, respectively. The high content of noise in Fig. 3a produces a less precise segmentation than in Fig. 2a. The synthetic nature of the images used in this work allows a quick quantification of the classification accuracy. The region of the image occupied by each class is well known. The confusion matrices related to the segmentation of Figs. 2a and 3a are given in Tables 2

and 3. From these matrices, the κ coefficient is easily calculated (Landis and Koch, 1977; Smits et al., 1999). For Fig. 2a, κ is 0.8586, and for Fig. 3a, it is 0.7937.

The second example deals with a multi-spectral SPOT image covering a region of Central México. A principal component analysis was applied to this image, only the first two components were retained for the classification process.

By visual inspection on this image, and with the help of ancillary data, six pixels were seeded signaling six spectral classes. Fig. 4a shows the region growth for the six classes overlain with an RGB color composite of the image. The color composite is, [RGB]=[Principal component 1, Principal component 2, Band 3]. These regions define the training fields for the spectral classes. From upper left in Fig. 4a, and going in clockwise direction, the classes defined in the classification process are: shadows, soil/herbage, microphylum thicket, oak/pine woods, river bed/herbage, and submontane thicket/oak. The distance (Eq. (6)) among the density functions of the regions measures the separability matrix of such classes. This separability matrix is shown in Table 4. The least separability in Table 4 is 1.8028; therefore, based on the values of Table 1, this classification was performed using a window of 5×5 pixels. To classify the border pixels, the distance among the pair of classes is calculated. The least separability in Table 5 is 0.9014; therefore, based on the values of Table 1, this classification is performed using a window of 11×11 pixels.

Table 2
Confusion matrix of classification for synthetic image 1

Class	1	2	3	4	5	Region size (pixels)
1	1862	175	73	43	49	2202
2	76	3410	133	104	41	3764
3	62	178	2710	86	20	3056
4	33	48	85	2556	58	2780
5	13	41	80	94	1426	1654
Total	2046	3852	3081	2883	1594	13456

κ Coefficient: 0.8586.

Table 3
Confusion matrix of classification for synthetic image 2

Class	1	2	3	4	5	6	Region size (pixels)
1	6436	158	189	0	0	0	6783
2	253	6367	47	42	0	0	6709
3	100	57	7189	222	81	0	7649
4	0	227	127	7045	161	75	7635
5	0	0	94	23	6543	167	6827
6	0	0	0	43	239	6461	6743

κ Coefficient: 0.7973.

Table 4
Distance among the class density functions in SPOT image

Class	1	2	3	4	5	6
1	0.0000	2.0000	2.0000	2.0000	1.9895	1.8028
2	2.0000	0.0000	1.8359	1.9873	2.0000	2.0000
3	2.0000	1.8359	0.0000	1.8931	1.9499	1.9919
4	2.0000	1.9873	1.8931	0.0000	2.0000	2.0000
5	1.9895	2.0000	1.9499	2.0000	0.0000	1.9925
6	1.8028	2.0000	1.9919	2.0000	1.9925	0.0000

The classification procedure generates a segmentation of the multi-spectral image. Fig. 4b shows the result of the segmentation once the class and border pixels are encoded. The overlay of the RGB composite plus the border of the regions (Fig. 4b) shows a precise segmentation of the spectral classes defined in this process. The evaluation of the goodness of this segmentation is as follows:

- (i) A visual inspection of the overlay of the class borders to the RGB composite. This inspection shows a precise segmentation of the image according to the classes defined.
- (ii) A verification of each region growth with ancillary data. The ancillary data was used to corroborate that each region contains the class signaled by the seed pixels. The resulting classification map was also corroborated with ancillary data.
- (iii) A comparison of the κ coefficient for the contextual vs. a per-pixel classification (Richards and Jia,

1999). The κ coefficient is calculated from the confusion matrix resulting from the contextual classification process (Table 6). The value of the κ coefficient is 0.98823. On the other hand, using the grown regions as training sets as input to a per-pixel classifier, the first two principal components of the SPOT image were classified into six classes. A mode filter employing windows of 5×5 and 7×7 pixels was applied to this classification. The resulting κ coefficient for this set of images is: 0.81198 for no filter, 0.91169 for the 5×5 filter, and 0.93382 for the 7×7 filter. Increasing the size of the window in the mode filter did not bring a significant improvement to the κ coefficient. The contextual classification was performed with a window of 5×5 pixels. The spatial detail achieved in this classification is comparable with that obtained from the per-pixel classifier employing a mode filter of 5×5 pixels.

Table 5
Distance among pair of classes and class density functions in SPOT image

Class, Class pair	1	2	3	4	5	6
1,2	1.0000	1.0000	1.8887	1.9875	1.9948	1.8773
1,3	1.0000	1.8913	1.0000	1.9204	1.9697	1.8701
1,4	1.0000	1.9936	1.9262	1.0000	1.9948	1.8773
1,5	0.9948	2.0000	1.9504	2.0000	0.9948	1.8710
1,6	0.9014	2.0000	1.9960	1.9078	1.9910	0.9014
2,3	2.0000	0.9180	0.9180	0.9936	1.9749	1.9928
2,4	2.0000	0.9936	1.8148	1.9875	2.0000	2.0000
2,5	1.9895	1.0000	1.8391	1.9875	1.0000	1.9925
2,6	1.8507	1.0000	1.8846	1.9875	1.9963	1.0000
3,4	2.0000	1.8849	0.9466	0.9466	1.9749	1.9928
3,5	1.9895	1.8913	0.9749	1.9204	0.9749	1.9882
3,6	1.8507	1.8913	0.9960	1.9204	1.9712	0.9960
4,5	1.9895	1.9936	1.8766	1.0000	1.0000	1.9925
4,6	1.8507	1.9936	1.9222	1.0000	1.9963	1.0000
5,6	1.8407	2.0000	1.9478	2.0000	0.9963	0.9963

Table 6
Confusion matrix for classification of SPOT image

Class	1	2	3	4	5	6	Region size (pixels)
1	730	0	0	0	18	15	763
2	0	199	6	0	0	0	205
3	0	0	918	0	0	0	918
4	0	0	13	772	0	0	785
5	0	0	0	0	122	0	122
6	0	0	0	0	1	3475	3476
Total	730	199	937	772	141	3490	6269

κ Coefficient: 0.98823

κ Coefficient: 0.98823.

5. Conclusions

A new contextual classifier based upon an automated region-growth algorithm has been developed and tested. This algorithm provides valid statistical samples of defined classes as input into a contextual spectral classifier. Even though the grown regions do not encircle pure spectral objects, the resulting classification is better compared to the one from a per-pixel classifier. The contextual classifier is semi-automatic, embracing only two parameters derived from heuristic criteria. More work is needed regarding the estimation of these parameters though. In order to avoid the heuristic determination of their value, a model-based approach may be adequate. The best location to seed the initial pixels for region growth requires some attention too.

The classification and growth of the regions are performed employing best windows for each class. No a priori assumptions are made concerning the density functions of the classes. This is a basic premise since, based on experimentation, in a multi-spectral image

some classes show a Gaussian behavior and some a Rayleigh behavior. A classifier is prone to produce wrong classification results whenever a particular density function is assumed and the image classes show a different statistical behavior. The synthetic images have a set of classes barely discernible due to the noise content. However, the contextual classifier produces good results, generating segmentation where the classes are clearly differentiated. The classification results concerning the SPOT image are good as well. The κ coefficient shows a greater spectral separability of the contextual classifier compared to a per-pixel classifier. Our contextual classifier performs well for multi-spectral images; a direct application might be in the classification of satellite images for remote sensing of the environment. The classification results in our classifier show a κ coefficient competitive with any other contextual classifier. Finally, the rationale of classification presented in this work is of general nature and might be adapted to new models of texture and spectral separability.

Acknowledgements

This work is supported by project CONACYT 27838A.

References

- Arai, K., 1992. A supervised thematic mapper classification with a purification of training samples. *International Journal of Remote Sensing* 13 (11), 2039–2049.
- Arai, K., 1993. A classification method with a spatial-spectral variability. *International Journal of Remote Sensing* 14 (14), 699–709.
- Atkinson, P.M., Lewis, P., 2000. Geostatistical classification for remote sensing: an introduction. *Computers & Geosciences* 26 (4), 361–371.
- Bolstad, P.V., Lillesand, T.M., 1991. Semi-automated training approaches for spectral class definition. *International Journal of Remote Sensing* 13 (16), 3157–3166.
- Buchheim, M.P., Lillesand, T.M., 1989. Semi-automated training field extraction and analysis for efficient digital image classification. *Photogrammetric Engineering & Remote Sensing* 55 (9), 1347–1355.
- Cetin, H., Levandowsky, D.W., 1991. Interactive classification and mapping of multi-dimensional remotely sensed data using n -dimensional probability density functions (n pdf). *Photogrammetric Engineering & Remote Sensing* 57 (12), 1579–1587.
- Chica-Olmo, M., Abarca-Hernández, F., 2000. Computing geostatistical image texture for remotely sensed data classification. *Computers & Geosciences* 26 (4), 373–383.
- Cohen, F.S., Fam, Z., 1992. Maximum likelihood unsupervised textured image segmentation. *CVGIP: Graphical Models and Image Processing* 54 (3), 239–251.
- Dougherty, E.R., 1999. *Random Processes for Image and Signal Processing*. SPIE Optical Engineering Press, Bellingham, WA, 592pp.
- Gahegan, M., Flack, J.C., 1999. Recent developments towards integrating scene understanding within a geographic information system for agricultural applications. *Transactions in GIS* 3 (1), 31–50.
- Gong, P., Howarth, P.J., 1992. Frequency-based contextual classification and gray-level vector reduction for land-use identification. *Photogrammetric Engineering & Remote Sensing* 58 (4), 423–437.
- González Alonso, F., Lopez Soria, S., 1991. Using contextual information to improve land use classification of satellite images in Central Spain. *International Journal of Remote Sensing* 12 (11), 2227–2235.
- Hodgson, M.E., 1998. What size window for image classification? A cognitive perspective. *Photogrammetric Engineering & Remote Sensing* 64 (8), 797–807.
- Khazenie, N., Crawford, M.M., 1990. Spatial-temporal autocorrelated model for contextual classification. *IEEE Transactions on Geoscience and Remote Sensing* 28 (4), 529–539.
- Kontoos, C.C., Rokos, D., 1996. The integration of spatial context information in an experimental knowledge-based system and the supervised relaxation algorithm—two successful approaches to improving spot- x_s classification. *International Journal of Remote Sensing* 17 (16), 3093–3106.
- Landis, J., Koch, G.G., 1977. The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174.
- Lira, J., Frulla, L., 1998. An automated region-growing algorithm for segmentation of texture regions in SAR images. *International Journal of Remote Sensing* 19 (18), 3595–3606.
- Raafat H, .M., Wong, A.K., 1988. A texture information-directed region-growth algorithm for image segmentation and region classification. *Computer Vision, Graphics and Image Processing* 43 (1), 1–21.
- Richards, J.A., Jia, X., 1999. *Remote Sensing Digital Image Analysis, An Introduction*, 3rd Edition. Springer, Berlin, 363pp.
- Smits, P.C., Dellepiane, S.G., Schowengerdt, R.A., 1999. Quality assesment of image classification algorithms for land-cover mapping: a review and a proposal for a cost-based approach. *International Journal of Remote Sensing* 20 (8), 1461–1486.
- Stuckens, J., Coppin, P.R., Bauer, M.E., 2000. Integrating contextual information with per-pixel classification for improved land cover classification. *Remote Sensing of Environment* 71 (3), 282–296.