Anais IX Simpósio Brasileiro de Sensoriamento Remoto, Santos, Brasil, 11-18 setembro 1998, INPE, p. 1087-1095.

A Supervised Classifier for Multispectral and Textured Images Based on an Automated Region Growing Algorithm

Jorge Lira and Gabriela Maletti

Instituto de Geofísica-UNAM
Circuito Institutos, Cd. Universitaria
04510 México D.F., México
labpr@tonatiuh.igeofcu.unam.mx

Abstract. A couple of supervised classifiers to segment optical multispectral images and textured radar images has been developed. In both classifiers, an automated region-growing algorithm delineates the training sets. Optimum statistics for defined classes are derived from the training sets. This algorithm handles three parameters: an initial pixel seed, a window and a threshold for each class. A suitable pixel seed is manually implanted through visual inspection of the image classes. The optimum value for the window and the threshold are obtained from spectral or texture distances. These distances are calculated from mathematical models of spectral and textural separabilities. A pixel is incorporated into a region if a spectral or texture homogeneity criterion is satisfied in the pixel-centered window for a given threshold. In this scheme, a region grows as much as possible but maintains the overlap with other regions in a minimum. The homogeneity criterion is obtained from the models of spectral and textures distances. 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 constitute therefore optimum training sets for each class. The statistical behavior of these training sets is used to classify the pixels of the image in one member of a set of classes. Comparing the statistical behavior of a sliding window with that of each class does the classification. The size of this window is the same as the one employed in the region-growing algorithm. The centered pixel of the sliding window is labeled as belonging to a class if its spectral or texture distance is a minimum to the class. Such distance is evaluated using the statistical content of the class and the sliding window as input to the model of spectral or textural separability. A series of examples, employing synthetic and natural images, are presented to show the value of this classifier. The goodness of the segmentation is evaluated by means of the Kappa coefficient and a matrix of distances derived from the mentioned model.

Key words: Region growing, contextual classifier, segmentation.

1 Introduction

A digital multispectral image, such as those gathered by satellite sensors, contains spectral, contextual and textural information related to the scene of interest. For optical images, the detail of information depends upon a series of factors, such as: number of spectral bands, size of the pixel, number of quantization levels, and signal to noise ratio. For radar images the factors are: pixel size, polarization, wavelength, geometric aspect, and signal to noise ratio. A given pixel in the image carries information of the related instantaneous field of view (IFOV). A pixel in the image is a numerical characterization of the average radiometric properties of the IFOV. Hence, a pixel is a statistical sample of the average response to the incoming radiation of the IFOV. In addition, a pixel is embedded in a certain spatial context. To derive the location and spatial organization of image objects a segmentation is required. A model of the scene is constructed by means of a segmentation of the image. By means of this model, some valuable aspects of scene behavior may be obtained.

Segmentation is a partition of the image in a number of regions, each region related to a spatial pattern 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 in a class using only its spectral properties and ignoring the context. This approach named per-pixel classification, proved to be limited in nature and applicable only to well spectral differentiated cases. In the last years, efforts (Gong and Howarth, 1992; Arai, 1993; Kontoes and Rokos, 1996) have been devoted to context classifiers. In this approach, a pixel is labeled to a class taking into account its spectral properties and the context of location. On the other hand, a texture is a spatial organization of pixel values; therefore, a texture classifier must be contextual by nature. A contextual classifier consistently produces higher classification accuracies than the per-pixel classifier.

A supervised classifier employs a-priori information of each determined class; this is usually done by means of training sets. These training sets are defined through closed polygons outlined by some interactive procedure on the image. In this definition, there is not a clear criterion to assume that the training sets are valid statistical samples of the classes. Therefore, 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 neighborhood.

In this work, a new contextual classifier is proposed that determine statistical samples of defined classes as a result of an automated region-growing algorithm. A pixel is then classified by comparing the shape of the density function associated to the pixel neighborhood and that of the classes. The comparison is done by means of a measure of similarity between density functions both: for spectral response and for texture content. As explained in the next sections, this scheme of classification is valid for both classifiers described in the present paper.

2 Contextual Classifier

2.1 Region growing scheme

The contextual classifier uses the training sets determined by an optimized region-growing algorithm (Lira and Frulla, 1998). The sets are statistical representations of defined classes, being these spectral or textural. 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 such as ground truth. Once the seeds are determined, one per class, the growing of the class regions starts. The growing is performed by pixel aggregation satisfying a homogeneity criterion. The criterion is evaluated in a window with optimum size. A pixel is aggregated into the region provided the difference between the homogeneity value of the seed centered window and the pixel-centered window does not exceed a certain threshold. The growing 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. The above may be established as follows

Let g(**r**) be the image and $\mathbf{p}_{ij}^{0} \in$ g, and let $R_0 = \{\mathbf{p}_{ij}^{0}\}$ be the initial sub-region. The pixel $\mathbf{p}_{ij}^{0}$ is known as the seed related to R_0. 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 $E(R_0)_v$ the value of the homogeneity criterion applied to the neighborhood 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

$$R_1 \equiv \left\{\mathbf{p}_{kl}^{1} : \left|E\left(\mathbf{p}_{kl}^{1}\right)_v - E\left(\mathbf{p}_{ij}^{0}\right)_v\right| \leq e\right\} \tag{1}$$

The real number ε is known as the parameter of uniformity. Once R_1 is been determined the previous step is repeated, so in general the region R_m is given by

$$R_m \equiv \left\{\mathbf{p}_{kl}^{m} : \left|E\left(\mathbf{p}_{kl}^{m}\right)_v - E\left(\mathbf{p}_{ij}^{0}\right)_v\right| \leq e\right\} \tag{2}$$

The homogeneity criterion E is always tested against the original sub-region R_0. The growing 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. Thus, the labeling of a tested pixel is carried out as

$$\mathbf{p}_{kl} \to \text{class } t\text{: } \left|E(\mathbf{p}_{kl})_v^t - E(\mathbf{p}_{ij}^0)_v^t\right| \leq e \tag{3}$$

Where $E(\mathbf{p}_{kl})_v^t$ is the homogeneity criterion applied to the tested pixel.

2.2 Estimation of optimum parameters for region growing

The optimum values for the window v and the threshold ε are obtained as follows. An odd sized window neighborhood is assumed for each seeded pixel. Begining from v = 3 pixels, the windows are systematically incremented in size. Only squared windows are considered. For multispectral images, the density function is estimated, for radar images the joint density function is derived by means of the co-occurrence matrix. Let $\mathsf{S}_v^a(\mathrm{i})$ be the density function in a window v for class a. Let $\mathsf{T}_v^s(\mathrm{i,j})$ be the joint density function in a window v for class s. These functions are normalized

$$\sum_{\mathrm{i}=1}^{256} \mathsf{S}_v^a(\mathrm{i}) = 1 \text{ and } \sum_{\mathrm{j}=1}^{256}\sum_{\mathrm{i}=1}^{256} \mathsf{T}_v^s(\mathrm{i,j}) = 1 \tag{4}$$

The above is assuming 256 quantization levels in the image, and $\mathsf{S}_v^a(\mathrm{i})$ is a k –dimensional vector where k is the number of bands. The joint density function $\mathsf{T}_v^s(\mathrm{i,j})$ is obtained by means of the co-occurrence matrix evaluated in the window v for directions 0°, 45°, 90° and 135° of the Freeman code. The optimum window size v is estimated when

$$\sum_{i=1}^{256} \left| \mathbf{S}_v^a(i) - \mathbf{S}_{v+1}^a(i) \right| \leq 0.03 \text{ or } \sum_{j=1}^{256}\sum_{i=1}^{256} \left| \mathrm{T}_v^s(i,j) - \mathrm{T}_{v+1}^s(i,j) \right| \leq 0.03 \tag{5}$$

In this sense, the window size is adjusted according to class heterogeneity. A smooth class requires a small window size; a heterogeneous class requires a greater window size. Equation (5) means the existence of an optimum window size for each class. The 3% indicated in equation 5 is derived from heuristic tests.

The third parameter handled in the region-growing algorithm is a threshold ε named the uniformity parameter. To estimate the value of this parameter ε_S for a multispectral image, let $d_{ab} = \sum_i \left| \mathbf{S}_{v_a}^a(i) - \mathbf{S}_{v_b}^b(i) \right|$, $\forall\, a \neq b$, be the minimum distance between spectral class-*a* and any other spectral class-*b*. Then, a pixel $\mathbf{p}_{kl}$ is incorporated into the region class-*a* if

$$\left| \mathbf{S}_{v_a}^a - \mathbf{S}_{v_a}^p \right| < d_{ab} \tag{6a}$$

and if

$$\left| \in_a - \in_p \right| \leq 0.03 \tag{6b}$$

Where $\in_a = \sigma_a/\mu_a$, is the heterogeneity of the initial optimum window for class *a*, and $\in_p = \sigma_p/\mu_p$, is the heterogeneity of the pixel window. The quantities μ and σ are the mean and the standard deviation respectively. The threshold ε_S is determined by expression (6a) with the restriction provided by (6b).

The estimate of the threshold ε_T for a textured radar image is as follows: a pixel p_{kl} is incorporated into the region class-*s* if

$$\left| \mathrm{T}_{v_s}^s - \mathrm{T}_{v_s}^p \right| < d_{st} \tag{7}$$

Where $d_{st} = \sum_i \left| \mathrm{T}_{v_s}^s(i,j) - \mathrm{T}_{v_s}^t(i,j) \right|$ is the minimum distance between texture class-*s* and any other texture class-*t*. The threshold ε_T is determined by expression (7).

2.3 Rationale of spectral classifier

The basic steps of the spectral classifier are the following:

i. - The bands selected for classification are loaded into RAM memory. Decorrelated bands are usually employed in this step.

ii. - Pixels are seeded in selected places of each spectral class defined for segmentation.

iii. - The optimum window and uniformity spectral parameter values are derived for each defined class according to equation (5) and (6).

iv. - The optimized region-growing algorithm is applied to the selected pixels, employing optimum parameters according to the previous step.

v. - For each region grown, the normalized density function is obtained. Each region represents a spectral class and is a valid statistical sample of the spectral behavior of the class.

vi. - A set of pixel centered neighborhoods formed by the optimum window sizes of the classes is considered to classify the pixels of the image. For each pixel neighborhood, the density function is obtained.
vii. - A pixel $\mathbf{p}_{kl}$ of the multispectral image is classified according to the following

$$\mathbf{p}_{kl} \rightarrow \text{Class } k\text{: } \sum_{i} \left| \mathbf{S}^{k}_{v_k}(i) - \mathbf{S}^{p}_{v_k}(i) \right| \text{, minimum, } \forall\ v_k \in \mu \qquad (8)$$

Where $\mathbf{p}_{kl}$ is a vector pixel with coordinates (k,l) and $\mu \equiv \{v_a, v_b, \ldots v_m\}$ is the set of optimum windows for m spectral classes.

2.4 Rationale of texture classifier

The rationale for the texture classifier is similar as the spectral classifier:
i. - The textured image is loaded into RAM memory. This image is usually a speckle filtered radar image.
ii. - Pixels are seeded in selected places of each texture class defined for segmentation.
iii. - The optimum window and uniformity texture parameter values are derived for each defined class according to equation (5) and (7).
iv. - The optimized region-growing algorithm is applied to the selected pixels, employing optimum parameters according to the previous step.
v. - For each region grown, a normalized co-occurrence matrix is obtained. This is the joint density function of the class. Each region is considered a statistically valid sample for the defined texture classes.
vi. - A set of pixel centered neighborhoods formed by the optimum window sizes of the classes is considered to classify the pixels of the image. For each pixel neighborhood the joint density function is obtained
vii. - A pixel p_{kl} of the texture image is classified according to the following

$$p_{kl} \rightarrow \text{Class } u\text{: } \sum_{i} \left| \mathrm{T}^{u}_{v_u}(i,j) - \mathrm{T}^{p}_{v_u}(i,j) \right| \text{, minimum, } \forall\ v_u \in \eta \qquad (9)$$

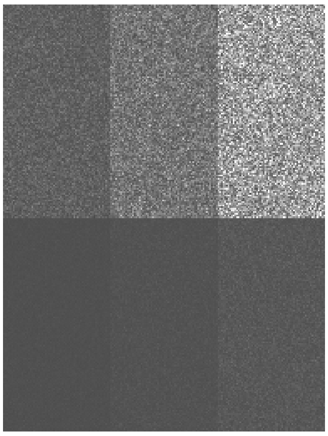
Where p_{kl} is a pixel with coordinates (k,l) and $\eta \equiv \{v_q, v_r, \ldots v_w\}$ is the set of optimum windows for *w* texture classes.

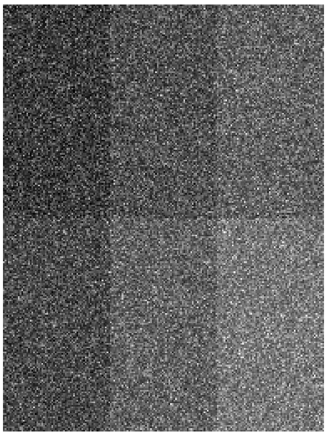
3 Results and discussion

Three examples are presented in this work. These examples are worked out on the grounds of: a). - A set of 36 synthetic images with well known statistical parameters for each class. The images in this set are singled band and contain six classes each. The dimension of these images is 192 x 256 pixels. The density function of such classes is Rayleigh-like, with varying mean and standard deviation. The set of synthetic images is generated as follows: for σ = 1, 2, 4 . . . 32, the distance between the class means $\Delta\mu$ is set as 1, 2, 4 . . . 32. b). - A multispectral SPOT image for which principal components were applied. The dimension of this image is 512 x 512 pixels, with a pixel size of 20 x 20 m^2, and covering a portion of central México. c). - A speckle filtered SEASAT radar image, gathered in the L band with four looks. The dimension of this image is 998 x 998 pixels, with a pixel size of 25 x 25 m^2, and covering a portion of northern México. This SEASAT image is of amplitude type.

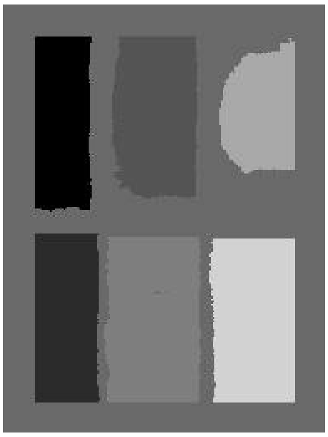
The series of images shown below resumes these results.

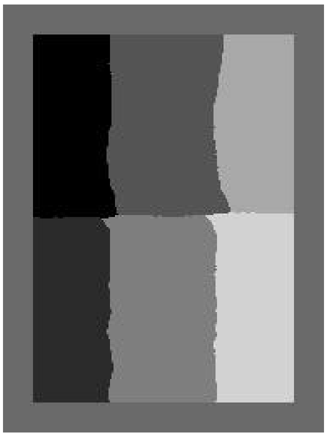
From the set of synthetic images, a multiband image of decorrelated bands was generated





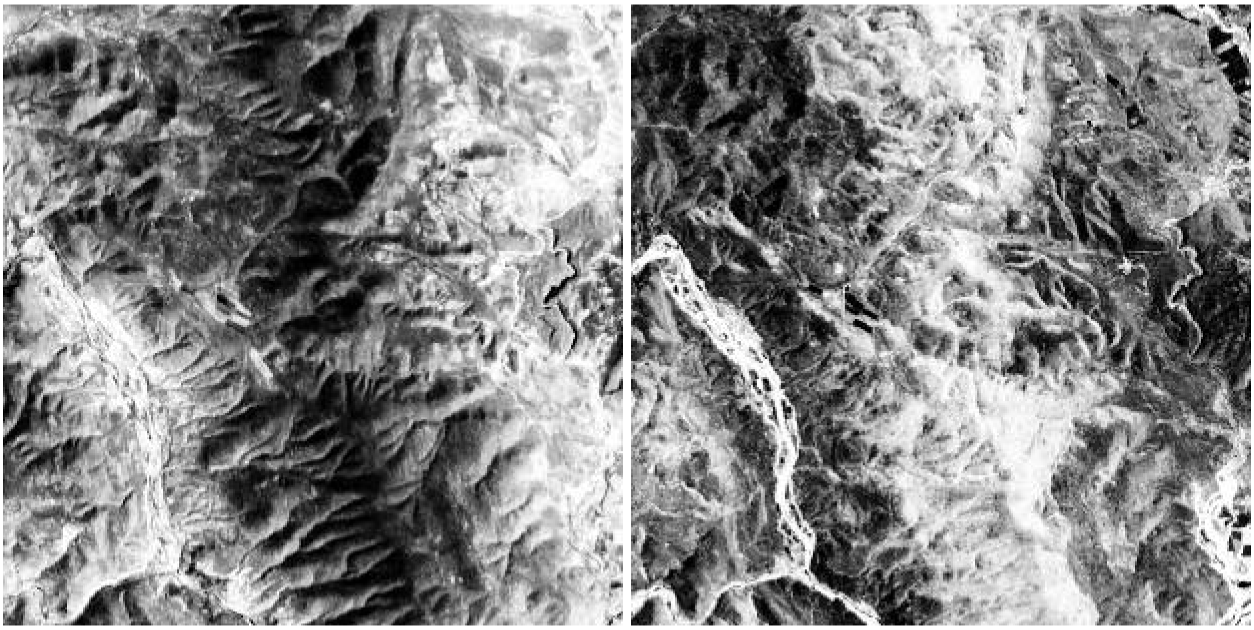
The region growing and the classification are shown in the following images



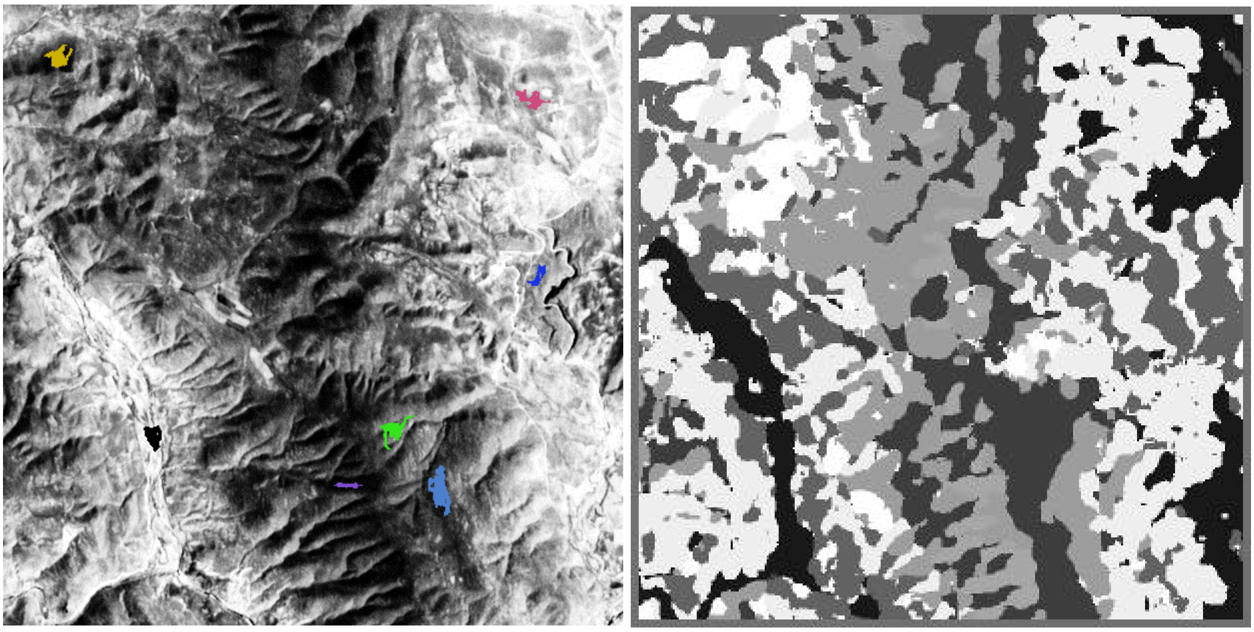


In the vertex of the four classes the density function is a class mixture generating a miss classification of some pixels. Both, for region growing and for classification, the window should be entirely contained in the image, hence the frame in the above two images.

The first two principal components of the SPOT image are the following:



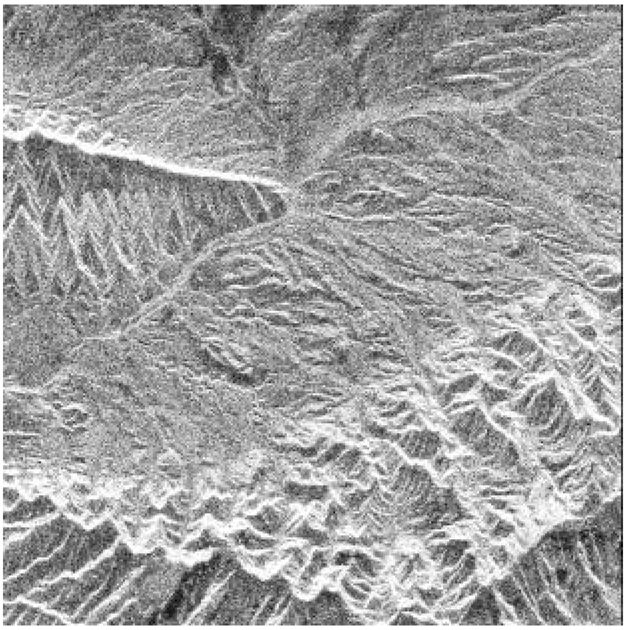
This two-band image depict a mountainous range (center) covered by heavy vegetation, a stream of a river (lower left), and soil mixed with spare vegetation (right). On this image, six pixels were seeded signaling six spectral classes. The following images show the result of the region growing and the spectral classification.



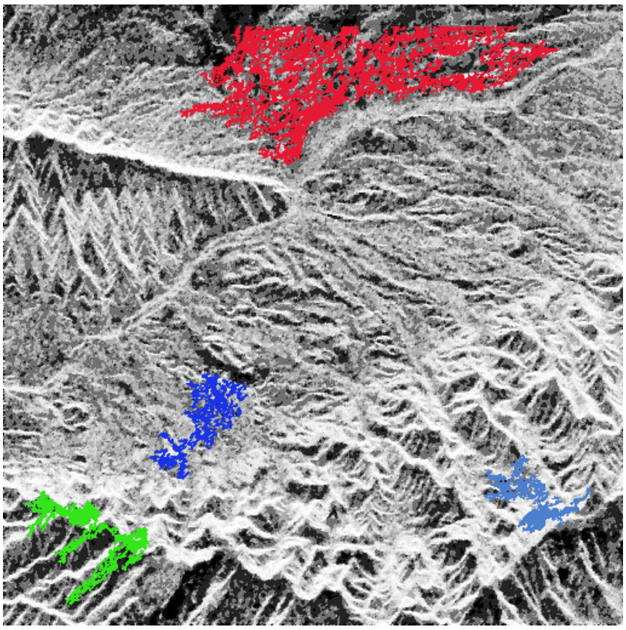
Those pixels whose associated density function have no intersection with the density function of any class are set as unclassified.

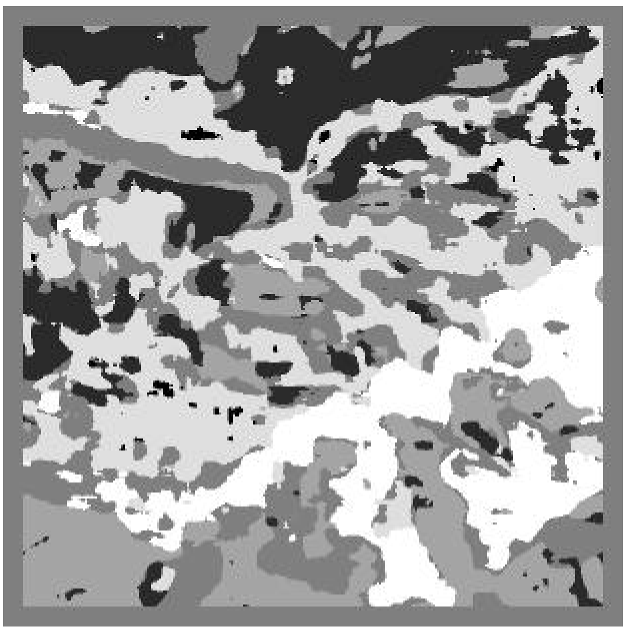
Anais IX Simpósio Brasileiro de Sensoriamento Remoto, Santos, Brasil, 11-18 setembro 1998, INPE, p. 1087-1095.

The SEASAT radar image, speckled and filtered are the following:



The speckle was filtered by means of a geometric filter (Lira and Frulla, 1998). On the speckle free image, six pixels were seeded signaling six texture classes. The following images show the result of the region growing and the texture classification





Four regions are shown as initial textures classes. The segmentation depicts four classes plus a non-classified pixels class (medium gray). The non-classified class includes the border of the image and parts of the image, this might be a fifht texture class.

4 Conclusions

A new contextual classifier based upon an automated region-growing algorithm has been developed and tested. This algorithm provides valid statistical samples of defined classes as input into a contextual classifier, both spectral and textural. The classification and growing of the regions are performed employing optimum 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, some classes show a gaussian behavior and some a Rayleigh like. The results are encouraging, although more research is needed, in particular the model for spectral and texture distance might be revised. However the rationale of classification is of general nature and might be adapted to new models of texture and spectral separabilities. In the present step of the research, no attempts have been made to identify the segmented classes with natural objects in the scene.

5 References

Arai, K. A Classification Method with a Spatial-Spectral Variability, *International Journal of Remote Sensing*, **14**, pp. 699-709, 1993.

Gong, P. And Howarth, P.J. Frequency-Based Contextual Classification and Gray-Level Vector Reduction for Land-Use Identification, *Photogrammetric Engineering & Remote Sensing*, **58**, pp. 423-437, 1992.

Kontoes, C.C. and Rokos, D. The Integration of Spatial Context Information in an Experimental Knowledge-Based System and the Supervised Relaxation Algorithm - Two Succesful Approaches to Improving SPOT-XS Classification, *International Journal of Remote Sensing*, **17**, pp. 3093-3106, 1996.

Lira, J and Frulla, L. An Automated Region Growing Algorithm for Segmentation of Texture Regions in SAR Images, *International Journal of Remote Sensing*, Accepted for Publication, 1988.