IMAGE ANALYSIS BY INTEGRATION
OF DISPARATE INFORMATION

Jacqueline Le Moigne
Information Science and Technology Office and
Center of Excellence in Space Data and Information Sciences
Mail Codes 930.4 and 930.5
NASA Goddard Space Flight Center
Greenbelt, MD 20771
Tel: (301) 286-8723
Fax: (301) 286-5152
lemoigne@nibbles.gsfc.nasa.gov

ABSTRACT

Image analysis often starts with some preliminary segmentation which provides a representation of the scene needed for further interpretation. Segmentation can be performed in several ways, which are categorized as pixel-based, edge-based, and region-based. Each of these approaches are affected differently by various factors, and the final result may be improved by integrating several or all of these methods, thus taking advantage of their complementary nature.

In this paper, we propose an approach that integrates pixel-based and edge-based results by utilizing an iterative relaxation technique. This approach has been implemented on a massively parallel computer and tested on some remotely sensed imagery from the Landsat-Thematic Mapper (TM) sensor.

1. INTRODUCTION

After pre-processing of some original data, image segmentation is the process which generates a spatial description of an image as a set of specific parts, regions or objects. The "segmented" output is then utilized by a higher-level image interpretation process. There is no single standard approach to segmentation which would be "successful" for any type of data, but some general methods have been developed based on the two main characteristics of regions or objects in an image:

(1) each region or object exhibits an internal uniformity with respect to some image property (e.g., gray level, color, texture),

(2) each region or object presents some contrast with its surroundings.

These two properties lead to three different types of segmentations, pixel-based, region-based and edge-based segmentations.

Each of these approaches is affected differently by various factors. Pixel-based methods form their decision only based on the information given at each pixel, while the two other types of segmentation take into account the information contained in the surrounding pixels. Usually in a pixel-based approach, all of the original information is utilized, thus avoiding a selection process. Such methods are also easier to integrate in a learning process, but their main

drawback is that they do not take into account spatial information. Conversely, edge- and region-based approaches base their decision on spatial information. Edge-based methods measure the variation of intensity between pixels belonging to different objects; they may produce excellent results for unevenly illuminated images, but they also can be very sensitive to noise. Region-based approaches, which measure the internal uniformity of some intensity or texture function, often produce spurious segmentation under non-uniform lighting conditions, but are usually less sensitive to noise.

In general, these three types of segmentation may be improved by integrating them and by taking advantage of their complementary nature. We previously proposed an approach that integrates region segmentation and edge detection results by interpreting a binary tree representation (Le Moigne, 1992), thus producing a refined region segmentation. This algorithm has been tested on Landsat-TM data. The integration of edge and region data may also be performed by a relaxation method and has been proposed in (Le Moigne, 1989); in the work described in this paper,we refine this relaxation method for the purpose of integrating edge and classification information and we implement it on a massively parallel computer, the MasPar MP-1. Then, we test this approach on remotely sensed imagery, such as Landsat-TM data.

2. RELAXATION TECHNIQUES

a. Overview

A large number of iterative relaxation schemes have been proposed to improve the results given by such basic processes as edge detection, region segmentation or pixel classification (Davis, 1981; Hummel, 1987; Faugeras, 1981; Peleg, 1980; Zucker, 1977). The principle of these algorithms is to utilize contextual information for iteratively changing the initial labeling of the objects in a scene toward optimal labeling.We will concentrate on relaxation methods for which the decisions at each point are taken in a probabilistic fashion. This general class of relaxation techniques is described in (Davis, 1981). Let us assume that we have a set of N objects $\{O_1, O_2, \ldots, O_N\}$ (e.g., the pixels) to be labeled into one of L classes $\{C_1, C_2, \ldots, C_L\}$ and that $p_i^n[C_k]$ is the probability that the object O_i is assigned to the class C_k at the iteration n. The principle of the relaxation algorithms, then, is to build a series of probability sets $\{p_i^n[C_k] ; 1 \le i \le N; 1 \le k \le L\}$, where each new iteration step, n, adjusts the probabilities according to the contextual information, and these probabilities satisfy the conditions (1) and (2):

for every i, for every k,
$$0 \le p_i^n[C_k] \le 1, \qquad (1)$$
and
$$\text{for every } i, \ \sum_{k=1}^{L} p_i^n[C_k] = 1. \quad (2)$$

Thus, the problem is to define the updating formula for $p_i^n[C_k]$. The simplest updating (which we use in this work) is described below.

The relaxation scheme assumes that the class assignments of each object depend on the class assignment of the "other" objects; the "other" objects can be defined, for example, by neighboring objects. Therefore, we define

c[i,k;j,l] as the compatibility coefficient between the object O_i with the label C_k and the object O_j with the label C_l. We assume that the c[i,k;j,l]'s belong to the range interval [0,+1] and are positive if there is compatibility and equal to 0 if there is high incompatibility. The coefficient c[i,k;j,l] can be defined as a conditional probability, $p(i \in C_k / j \in C_l)$, which provides a probabilistic framework. Let $q_i^n[C_k]$ be the global compatibility for the object O_i with the label C_k; it is defined by

$$q_i^n[C_k] = \frac{1}{V[i]} \sum_{j=1}^{j=V[i]} \sum_{l=1}^{l=L} c[i,k;j,l] \, p_j^n[C_l] \quad (3)$$

where $V[i]$ is the number of objects in the neighborhood of object O_i. Then $q_i^n[C_k]$ is the "increment" which is applied to update $p_i^n[C_k]$ and compute the new probability set $\{p_i^{(n+1)}[C_k]\}$ (see (Davis, 1981) for details):

$$p_i^{n+1}[C_k] = \frac{p_i^n[C_k] \times q_i^n[C_k]}{\sum_{l=1}^{l=L} p_i^n[C_l] \times q_i^n[C_l]} \quad (4)$$

This multiplication still ensures that the conditions (1) and (2) are satisfied. Besides, if the global compatibility of the object O_i with the label C_k is higher than all the compatibilities of O_i with the other labels C_l, then the $p_i^n[C_k]$'s increase relative to the other $p_i^n[C_l]$'s. That means that this scheme provides an overall improvement of the labeling but it does not guarantee the convergence toward stable labeling.

Other relaxation schemes (Faugeras, 1981; Peleg, 1980) utilize different formulas or different frameworks. For example in (Faugeras, 1981), another way of updating the probabilities $p_i^n[C_k]$ is given; the principle of this other method is to minimize a criterion by the "projected gradient" optimization method. The criterion measures the ambiguity and the consistency of the current labeling at each step. This algorithm provides us with a converging sequence of probabilities P^n. Starting from an initial point P^0, the method converges toward a local minimum in the vicinity of P^0.

b. Utilizing Relaxation to Integrate Disparate Information

We now describe how a relaxation technique, such as the one described above, can be utilized to integrate knowledge from edge detection and pixel classification.

Previously, Zucker and Hummel (Zucker, 1977) simultaneously used edge and region data in a relaxation process for labeling dots. Their goal was to provide a low-level description of the roles of dots in cluster analysis. Unlike our approach, they used these two types of data "equally", i.e., both edges and regions defined the labels and the initial probabilities. There were ten different labels, with eight edge labels at various orientations, one "region" label called "interior point" label, and one "noise" label.

In our approach, the labels are all region labels, and edge data are utilized to update the labeling by way of the compatibility coefficients. This approach tends to be more general and this algorithm can be very easily applied to the fusion of various types of data. In our work, the

definition of the labels and the initial probabilities are provided by a neural network pixel classification (Chettri, 1992), then the relaxation process updates the initial labeling by using some edge detection results, e.g., from a Canny operator (Canny, 1986). These standard techniques, pixel classification and edge detection, can easily be changed without altering the definition of the overall relaxation algorithm.

Once the initial probabilities are given, the fusion is realized by computing the compatibilities between neighboring pixels. Coefficients $c[i,k;j,l]$ in formula (3) represent the compatibilities between object O_i with the label C_k and object O_j with the label C_l. If a relaxation is performed only on regions, these coefficients represent the compatibilities of neighboring pixels belonging to given regions. If both regions and edges are considered, both information can be integrated into the compatibility coefficients by utilizing the following equation:

$$c[i,k;j,l] = p_r(C_k/C_l) \times F_{k\&l}(i,j) \quad (5)$$

where $p_r(C_k/C_l)$ is the "region-probability" (independent of edges) of i belonging to the class C_k if the neighbor j belongs to the class C_l; this probability could be estimated from some previous ground truth data. The last term in Eq. (5), $F_{k\&l}(i,j)$, is a function that varies in [0,1] and is:

- closer to 1 if $C_k = C_l$ (or, more generally, if regions C_k and C_l are similar in tone) and i is not an edge point, or if regions C_k and C_l are different and i is an edge point,
- closer to 0 if $C_k = C_l$ and i is an edge point, or if regions C_k and C_l are different and i is not an edge point.

For example, F can be defined by:

$$F_{k\&l}(i,j) = K_{k\&l} \times Mag(i) + (1 - K_{k\&l}) \times (1 - Mag(i))$$

where $K_{k\&l}$ is equal to 0 if k=l and equal to 1 if k≠l, and Mag(i) is the magnitude of the gradient computed at the point i and normalized between 0 and 1.

The magnitude of the gradient at neighboring point j would also be taken into account and we could utilize the following formula:

$$F_{k\&l}(i,j) = K_{k\&l} \times Mag(i) \times Mag(j) + (1 - K_{k\&l}) \times (1 - Mag(i) \times Mag(j))$$

This updating has not been implemented yet, but will be considered in future work.

Therefore, both region and edge information participate simultaneously in the updating of the labeling. Results are presented in Section 4.

3. PARALLEL IMPLEMENTATION

Computation time is the main concern of relaxation techniques. However these techniques are characterized by parallel local processing which can be easily implemented on an architecture that favors computations between adjacent pixels (Fishler, 1987). The method discussed in this paper has been implemented on a MasPar MP-1. The MasPar Parallel Processor is a fine-grained, massively parallel SIMD architecture, with 16,384 parallel processing elements arranged in a 128x128 matrix and connected by an eight nearest neighbors interconnection network.

The parallel implementation of the relaxation algorithm is straightforward, with the quantities $p_i^n[C_k]$ and $q_i^n[C_k]$ computed in parallel at each pixel. Timings are given in Table 1 for various size images and various numbers of labels.

4. RESULTS

Figure 1 shows the results of this algorithm on a test image. In this example, we assume that the "ideal" image (or "ground truth") is composed of two distinct regions and we choose some compatibility coefficients which reflect this assumption; the initial probabilities present three classes, two classes corresponding to the two "ideal" regions and one "artefact" class. The edge image presents strong edges at the border between the two "ideal" regions and the edge magnitudes decrease with the distance to this border. For this example, we notice that if the relaxation algorithm is utilized without any edge information, one of the labels "takes over" the whole image after only three iterations. If the edge information is integrated in the relaxation process, the two regions are still separated after 10 iterations and the labeling seems to be stable.

Figures 2 to 5 present results of the algorithm applied on a Landsat-TM scene ("Washington D.C. region") shown at the top left corner of Figure 2. Initial probabilities for this scene have been obtained from a classification into seven labels, utilizing a probabilistic neural network (see (Chettri, 1993) for details). The 7 labels correspond to the classes "urban", "agriculture", "rangeland", "forest", "waterbodies", "wetland", and "bareland".

The algorithm described in section 2 was applied to this initial classification, using two different numbers of labels; first, we grouped these 7 labels into the 3 classes "urban", "agriculture", and "other". Figures 2 and 3 show the results without and with the edge information. When no edge information is utilized (Figure 2), one of the labels (label 0) has "disappeared" after 20 iterations. When the edge information is taken into account (Figure 3), the three labels are still present after 30 iterations, and the labeling seems to stay stable after 20 iterations.

Then, the 7 labels defined previously are considered, and we obtain similar results: see Figures 4 and 5. When no edge information is utilized, only three labels are left after 30 iterations and one of the three is covering most of the image; but when the edge information is integrated in the updating formula, all the initial labels are still represented after 30 iterations and the labeling seems to stabilize after 20 iterations. Also, the results obtained after 10 or 30 iterations can be compared to the ground truth data shown at the top left corner of Figure 5: qualitatively, we can observe an overall improvement of the segmentation (e.g., suppression of small isolated pixels or groups of pixels), but some features (e.g., a road) have been regrouped with a neighboring region. A quantitative evaluation of these results will be performed later.

The previous results show the importance of the edge information and how the integration of edge- and the pixel-based segmentations can improve the final result.

Other similar results have also been obtained on AVHRR data and will be presented at the conference.

137

5. CONCLUSION

The results presented in this paper show how the integration of complementary information, such as pixel- and edge-based techniques, can improve the final segmentation.

More work needs to be done, especially in the definition of the initial probabilities, in the choice of the compatibility coefficients, and in the quantitative evaluation of the results. Also, the results presented in section 4 seem to show that, when the edge information is utilized, the relaxation process becomes "stable" after a certain number of iterations: this issue of "convergence" will be studied, and different schemes, such as the one presented in (Faugeras, 1981) will be investigated.

Acknowledgments

REFERENCES

Canny, J., (1986), "Computational Approach to Edge Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. PAMI-8, 679-698.

Chettri, S.R., (1993), "The Probabilistic Neural Network Architecture for High Speed Classification of Remotely Sensed Imagery," submitted to 1993 Goddard Conference on Space Applications of Artificial Intelligence.

Davis, L.S., & Rosenfeld, A., (1981), "Cooperating Processes for Low-Level Vision: A Survey," Artificial Intelligence 17, 245-264.

Faugeras, O.D., & Berthod, M., (1981), "Improving Consistency and Reducing Ambiguity in Stocchastic Labeling: An Optimization Approach," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-3, No. 4, 412-424.

Fishler, M.A., & Firschein, O., (1987), "Parallel Computer Architectures for Computer Vision," in Readings in Computer Vision: Issues, Problems, Principles and Paradigms, edited by M.A. Fishler and O. Firschein, Morgan Kaufmann Publishers, Los Altos, California, 768-772.

Hummel, R.A., & Zucker, S.W., (1987), "On the Foundations of Relaxation Labeling Processes," in Readings in Computer Vision: Issues, Problems, Principles and Paradigms, edited by M.A. Fishler and O. Firschein, Morgan Kaufmann Publishers, Los Altos, California, 585-605.

Le Moigne, J., (1989), "Data Fusion by Relaxation," in Proc. of Intelligent Autonomous Systems Intern. Conference, Amsterdam, December 11-14, 368-377.

Le Moigne, J., & Tilton, J.C., (1992), "Refining Image Segmentation by Integration of Edge and Region Data," IGARSS'92, Houston, Texas, May 26-29.

Peleg, S., (1980), "A New Probabilistic Relaxation Scheme," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-2, No. 4, 362-369.

Zucker, S.W., & Hummel, R.A., (1977), "Computing the Shape of Dot Clusters, I: Labelling Edge, Interior, and Noise Points," University of Maryland Computer Science Center Technical Report No 543.

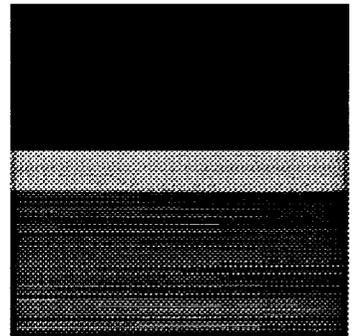| Image Size	# Labels	Edges ?	Time per Iteration
256*256	7	no	0.36
256*256	7	yes	0.41
256*256	3	no	0.10
256*256	3	yes	0.11
512*512	2	no	0.17
512*512	2	yes	0.17

Table 1

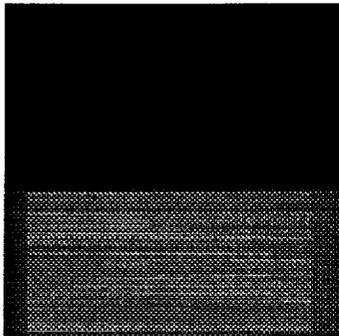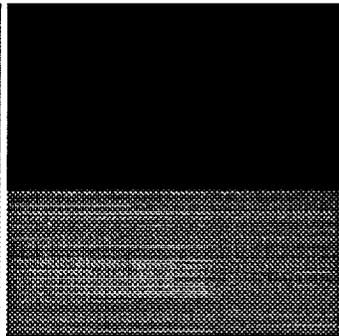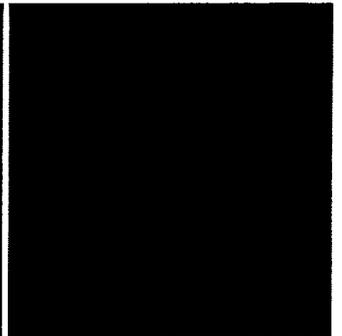*Timings Obtained for One Relaxation Iteration on a MasPar MP-1*
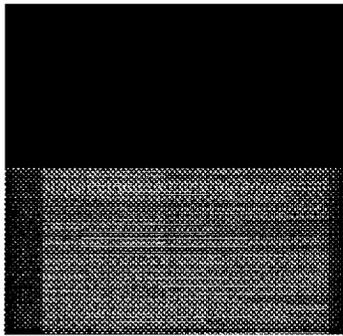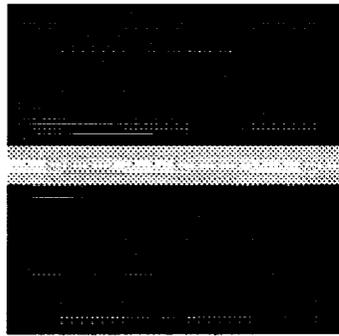
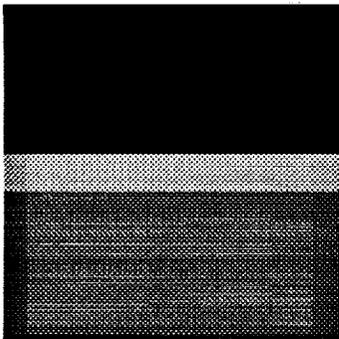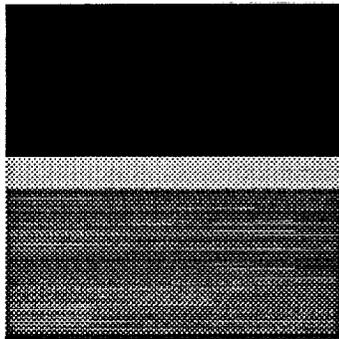"Ideal" Edges Initial Probabilities
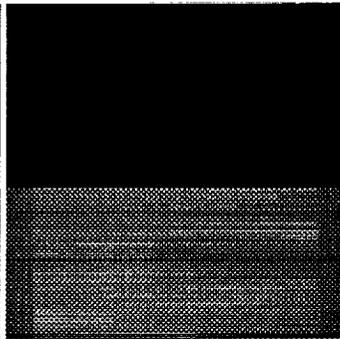
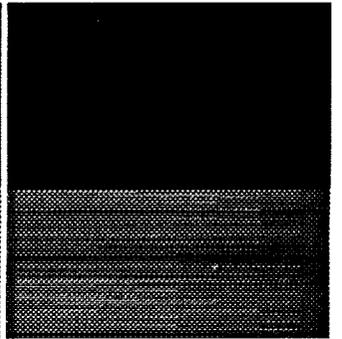Iter 1/ no edges Iter 2/ no edges Iter 3/ no edges Iter 10/ no edges

Iter 1/ edges Iter 2/ edges Iter 3/ edges Iter 10/ edges

<u>Figure 1</u>
Results of our Method on a Test Example
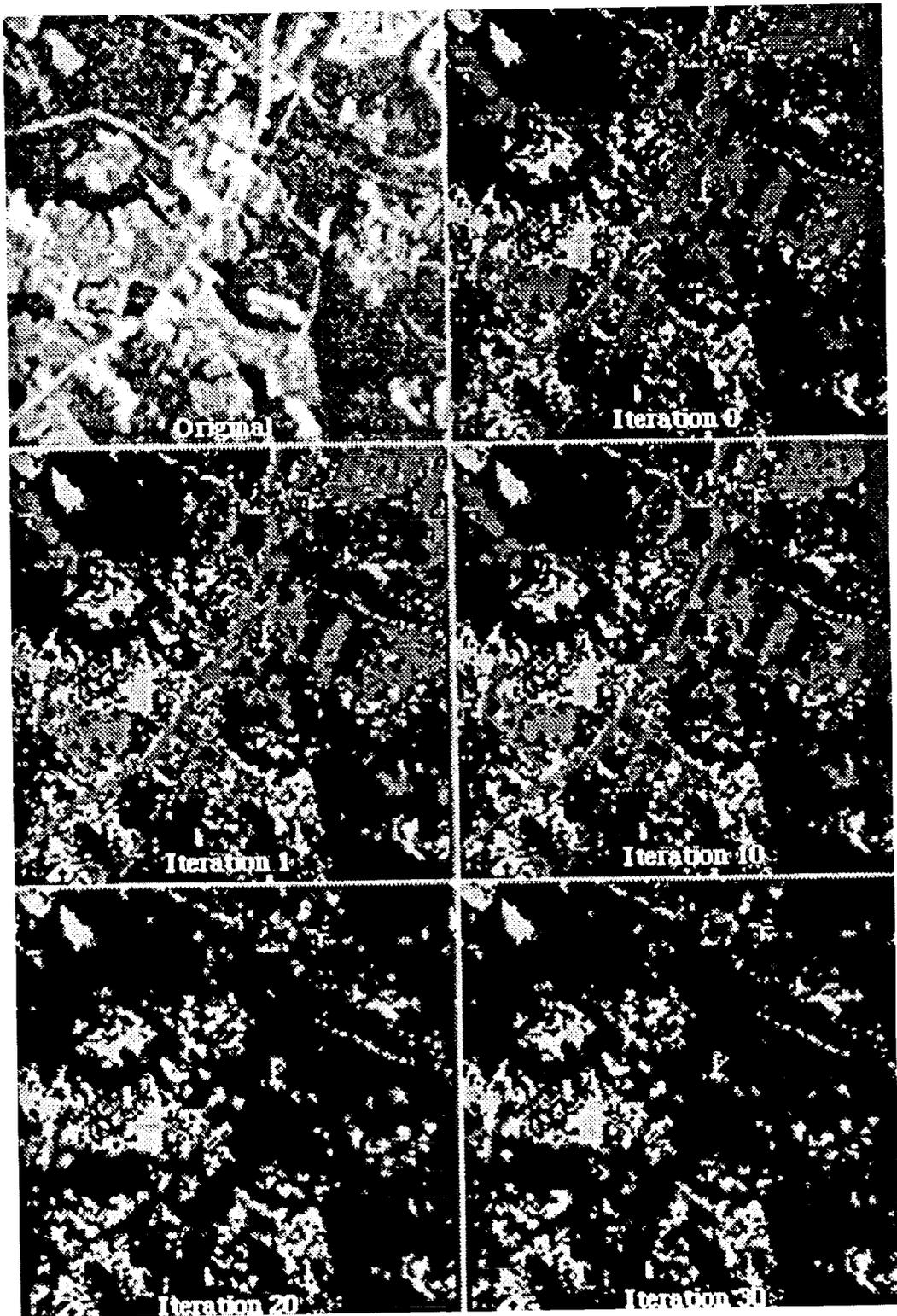
Figure 2
Relaxation Results Without Edge Information
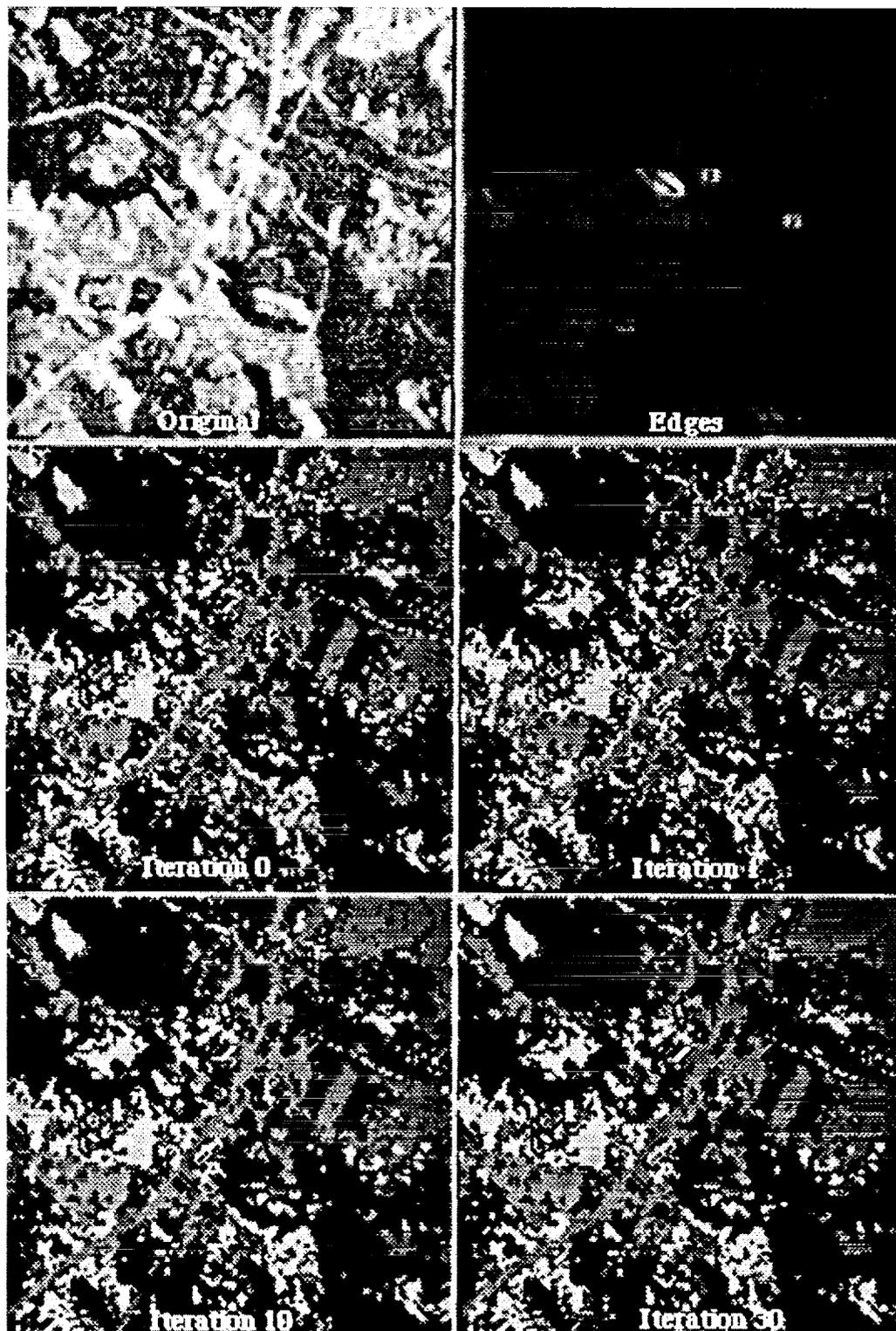*on a Landsat-TM scene (3 Labels)*

Figure 3
*Results of the "Relaxation With Edges" Method
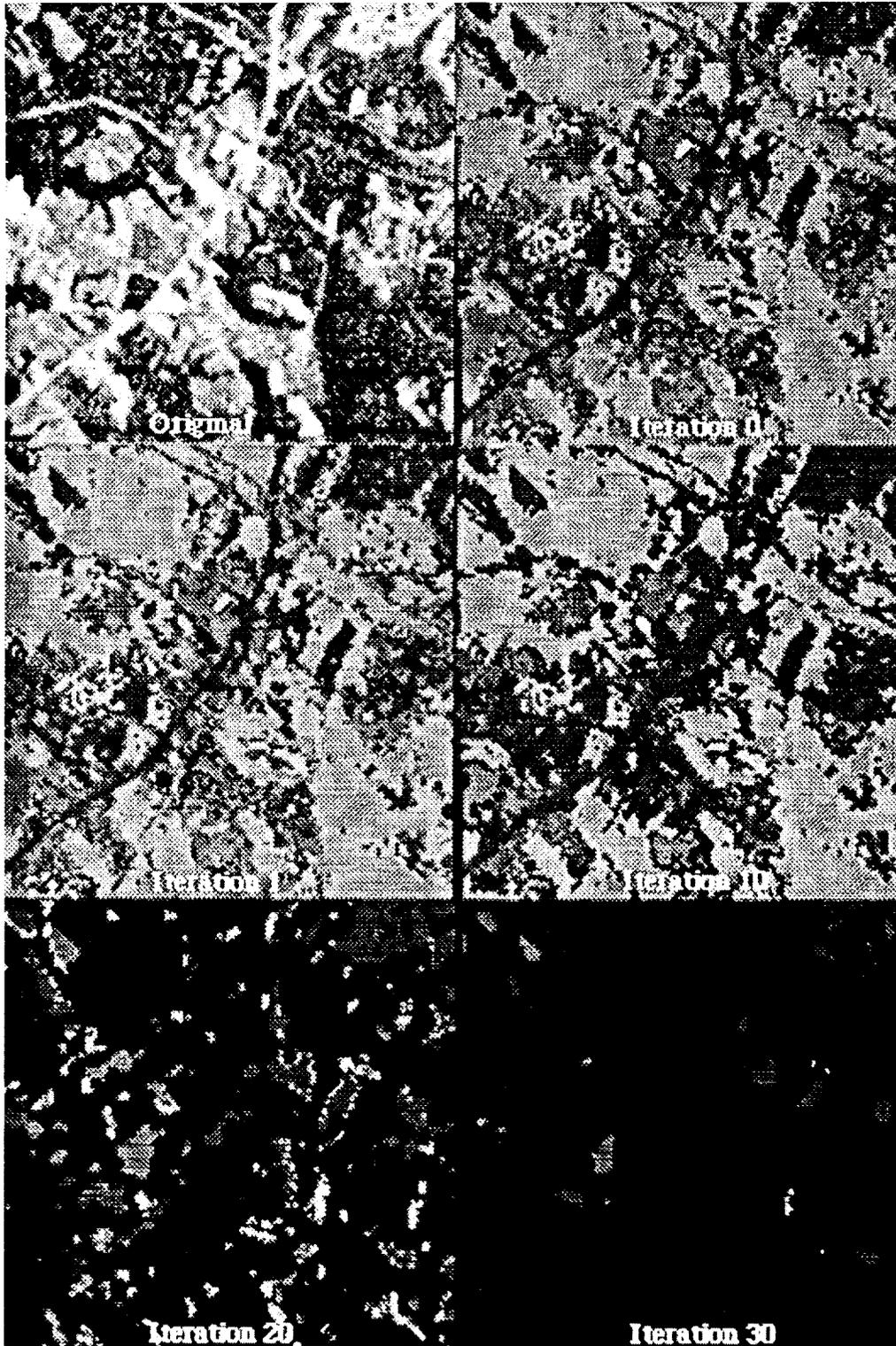on a Landsat-TM scene   (3 Labels)*

Figure 4
*Relaxation Results Without Edge Information
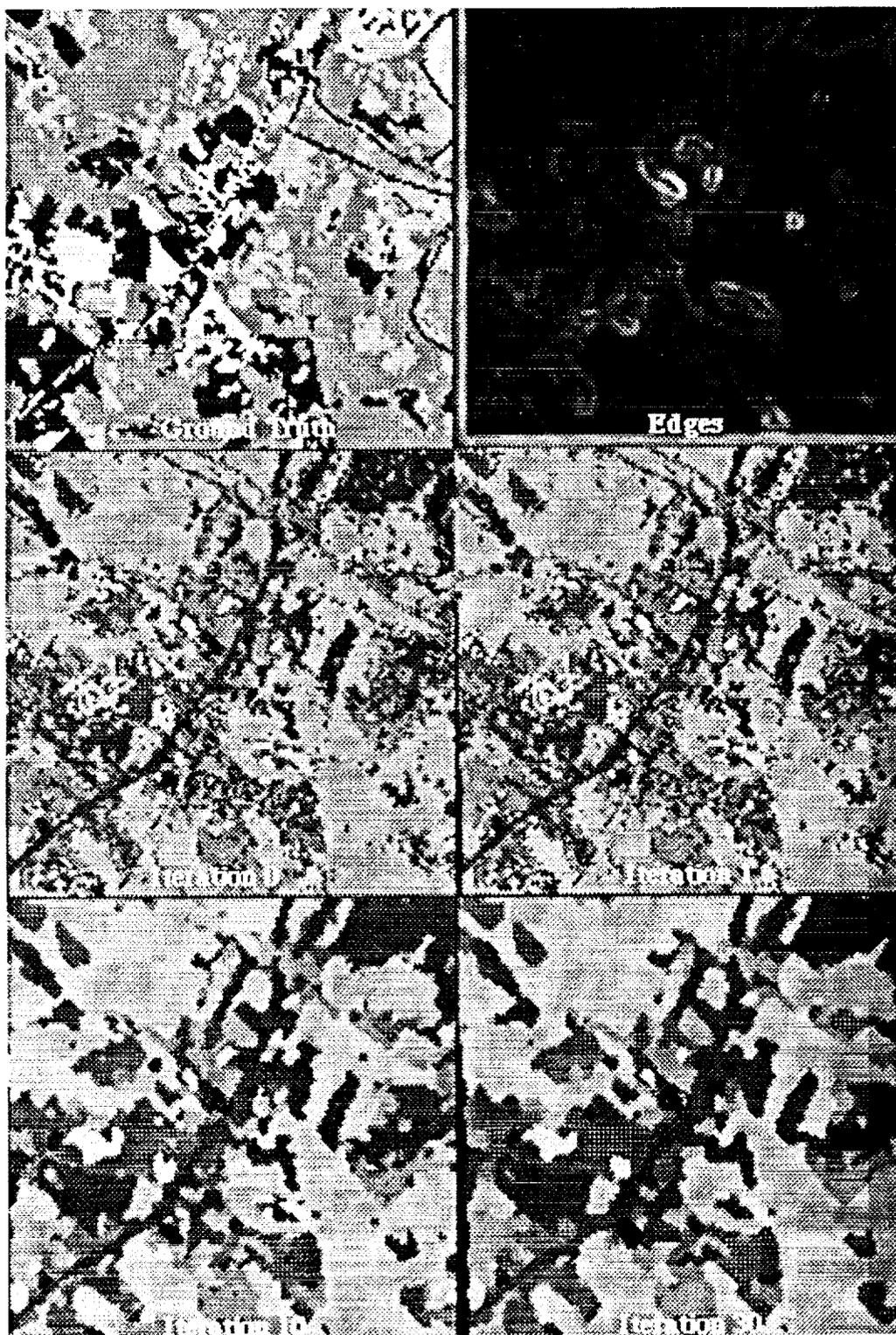on a Landsat-TM scene (7 Labels)*

Figure 5
*Results of the "Relaxation With Edges" Method
on a Landsat-TM scene (7 Labels)*