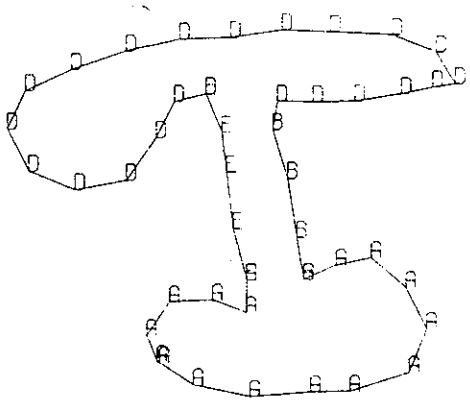


Fig. 10. Colon 4. Number of points=45; cluster time=1.50 s.

By taking into account the presence of a certain number of small triangular submatrices which do not merge with each other, $(T_i, \cdots, T_j)$, an investigation is carried out to test whether this region could be a concavity set in the shape boundary. Let P_a, P_b be the endpoints of the region defined by $T_i, \cdots, T_j$ on the shape. Let, furthermore, Z be the triangular submatrix of ML_I obtained by merging together $T_i, \cdots, T_j$ (i.e., Z is the submatrix of $ML_I(s, t)$ where $a \leqslant s \leqslant t, a \leqslant t \leqslant b$). The test is based on the evaluation of all the elements belonging to Z which do not belong to any $T_i, \cdots, T_j$. This region, called test region, is indicated in Fig. 2 by dotted lines. If all the elements in the test region are zeros, then the algorithm looks for a mirror of S_i. The detection of such a region is made by looking for a square submatrix of ML_I made of only ones, PRISM (s, t), in the band $a + 1 \leqslant s \leqslant b - 1, 1 \leqslant t \leqslant a - 1$, or in the band $b + 1 \leqslant s \leqslant k, a + 1 \leqslant t \leqslant b - 1$, where k is the number of rows of the ML_I matrix.

If such a region does exist, we say that the points belonging to the "intrusion" [e.g., $P_{22}, P_{23}, P_{24}, P_{25}, P_{26}, P_{27}$ in Fig. 1(a)] can be considered as a whole set. By using this technique combined with the procedure described in the previous section, the algorithm divides the shape in the following sets:

$$(P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8),$$

$$(P_8, P_9, P_{10}, P_{11}, P_{12}, P_{13}, P_{14}),$$

$$(P_{14}, P_{15}, P_{16}, P_{17}, P_{18}, P_{19}, P_{20}, P_{21}, P_{22}),$$

$$(P_{22}, P_{23}, P_{24}, P_{25}, P_{26}, P_{27}, P_1).$$

The method described in [1] produces as a single element of the decomposition that one formed by the points $P_8, \cdots, P_{14}, P_{22}, \cdots, P_{27}, P_1$. The fact that our method groups $P_8, \cdots, P_{14}$ and $P_{22}, \cdots, P_{27}, P_1$ separately suggests that our technique may be producing finer divisions than the previous one. The algorithm was implemented in Fortran and run on an IBM 370-145. The execution time is, for Fig. 1, 1.05 s. The value of the correlation factor is $CORR_1 = 0.7$ for the first iteration, $CORR_2 = CORR_1 - 0.12$ for the second iteration, and $CORR_3 = CORR_2 - 0.12$ for the third iteration. The same approach is used to carry out the following examples.

V. Some Example and Time Evaluation

We report here the original boundary, the obtained decomposition, and the CPU execution time in the cases listed in Figs. 5–10.

Acknowledgment

The authors wish to thank Prof. Haralick and Dr. Shapiro for the interesting discussions about the problem described in this correspondence and the essential help in furnishing data for the experiments.

References

[1] L. Shapiro and R. Haralick, "Decomposition of two-dimensional shapes by graph-theoretic clustering," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-1, Jan. 1979.

Segmentation of Images Having Unimodal Distributions

BIR BHANU AND OLIVIER D. FAUGERAS

Abstract—A gradient relaxation method based on maximizing a criterion function is studied and compared to the nonlinear probabilistic relaxation method for the purpose of segmentation of images having unimodal distributions. Although both methods provide comparable segmentation results, the gradient method has the additional advantage of providing control over the relaxation process by choosing three parameters which can be tuned to obtain the desired segmentation results at a faster rate. Examples are given on two different types of scenes.

Index Terms—Gradient relaxation, image segmentation, nonlinear relaxation, optimization, unimodal distribution.

I. Introduction

Various approaches based on thresholding have been used by many researchers for the segmentation of both monochrome and color pictures [1], [2]. Normally, in the application of these techniques, the histogram shows two or more peaks in at least one of the spectral features corresponding to various homogeneous regions of an image. Very often preprocessing is done to improve the histograms, and local properties are used to compute the global, local, or dynamic thresholds. However, if the intensity (or color) histogram of the image is unimodal, then the application of such methods gives a poor segmentation and there are no criteria for automatic threshold selection. Unimodal distributions are typically obtained when the image consists mostly of a large background area with other small but significant regions. For example, in the biomedical area, the extraction of the boundaries of various types of cells is complicated by the fact that the cells are very close together, their boundaries are poorly defined, and the gray level histogram is unimodal. Similarly, in scenes with many different objects as in the case of aerial photographs, the histogram may have only one peak because the range of intensities for each object will probably overlap with the ranges of other subjects. Jain *et al.* [3] consider the segmentation of muscle cell pictures using ten low level operators which are very time-consuming and require the selection of five thresholds. Rosenfeld and Davis [4] and Peleg [5] use iterative methods to modify the histogram. Their methods do not take into ac-

Manuscript received August 4, 1980; revised December 4, 1981. This work was supported in part by DARPA Contract F-33615-80-C-1080 and NIH Grant GM23732.

B. Bhanu was with the Image Processing Institute and Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90007. He is now with the Aeronutronic Division, Ford Aerospace and Communications Corporation, Newport Beach, CA 92660.

O. D. Faugeras was with the Image Processing Institute, Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90007. He is now with INRIA, Rocquencourt, France, and the University of Paris XI, Paris, France.

0162-8828/82/0700-0408$00.75

count the possibility that small regions in an image may have significance even though they may not show significant peaks in the histogram of the image. Rosenfeld [6] considers thresholding by relaxation. In this correspondence, we study a gradient relaxation method based on maximizing a criterion function and compare it to the nonlinear probabilistic relaxation method [6], [7] for the purpose of segmenting images having unimodal gray level distributions. Results are illustrated with the aid of two different types of scenes.

II. Segmentation Schemes

Fig. 1 shows two 128 × 128 pixel 8-bit images. The background of the image in Fig. 1(a) consists of a confluent monolayer of human skin cancer cells, and the small circular-shaped objects are human lymphocytes and red blood cells. The objective is to get the boundaries of all the cells. The image in Fig. 1(b) is part of an aerial photograph. Here the objective is to detect significant features such as roads, etc. Gray level histograms of these images are shown in Fig. 2. Note that the histograms are almost unimodal, and as a consequence, there is no reliable way of automatically choosing a threshold for segmenting these pictures. Although one could argue that the gray level intensity 16 in Fig. 2(b) might be a valley, thresholding the image at this value produces a very poor segmentation.

Commonly used difference operators such as gradient, Laplacian, and Sobel were applied to the images shown in Fig. 1. For example, Fig. 3(a) shows the Laplacian of the cell image obtained by convolution of the cell image by a 3 × 3 mask. The gray level histogram of the image in Fig. 3(a) is shown in Fig. 3(b). We also considered the methods based on thresholding the histogram of the picture where the gradient, Laplacian, and edge values are high [1]. However, the pictures so obtained have a unimodal histogram and lack the criterion for segmenting them at the valley of two peaks. Thresholding at the gray level corresponding to the mode or mean of the filtered histogram gave very poor results. Edge detection has also been done by convolving the images in Fig. 1 with 5 × 5 masks corresponding to the ideal step edges in six directions [8]. Thresholding of the magnitude image does not show good segmentation. A number of bar masks of various sizes and orientations have also been used, but the results were poor (for the image shown in Fig. 1(a) the width of an edge is about 5–6 pixels).

III. Segmentation Using Relaxation Methods

A. Gradient Relaxation Algorithm

In [9], Faugeras and Berthod proposed a relaxation algorithm which is based upon the explicit use of consistency and ambiguity to define a global criterion upon the set of units. This criterion has the inherent problem that the consistency and ambiguity tend to go in opposite directions. Therefore, in the present study, we consider a simpler criterion [10] based on the inner product of probability vector $\boldsymbol{p}_i$ and compatibility vector $\boldsymbol{q}_i$. N is the number of pixels in the image, $\boldsymbol{q}_i$ is a function of $\boldsymbol{p}_i$'s as discussed below, and we defined the criterion as

$$C(\boldsymbol{p}_1, \boldsymbol{p}_2, \cdots, \boldsymbol{p}_N) = \sum_{i=1}^{N} \boldsymbol{p}_i \cdot \boldsymbol{q}_i \tag{1}$$

and carry out its maximization using the gradient projection approach.

Suppose we have a set of N pixels $i = 1, 2, \cdots, N$ which fall into two classes λ_1 and λ_2 corresponding to the white (gray value = 255) and black (gray value = 0). The relaxation process is specified by choosing a model of interaction between pixels. We attach to every pixel i the set V_i of its eight nearest neighbors. Assuming that objects of interest in the picture are continuous, we will make like reinforce like and define a compatibility function c such that

$$\begin{aligned} c(i, \lambda_k, j, \lambda_l) &= 0, \quad k \neq l, \quad j \text{ in } V_i \text{ for all } i \\ c(i, \lambda_k, j, \lambda_k) &= 1, \quad k = 1, 2 \quad j \text{ in } V_i \text{ for all } i. \end{aligned} \tag{2}$$

The compatibility vector $\boldsymbol{q}_i$ is then defined as

$$q_i(\lambda_k) = \tfrac{1}{8} \sum_{j \in V_i} \sum_{l=1}^{2} c(i, \lambda_k, j, \lambda_l)\, p_j(\lambda_l), \quad k = 1, 2, \quad i = 1, \cdots, N. \tag{3}$$

In effect, $q_i(\lambda_k)$ is the mean neighborhood probability of the ith pixel for the class under consideration, i.e., $q_i(\lambda_k) = (\frac{1}{8}) \sum_{j \in V_i} p_j(\lambda_k)$. The choice of compatibility function in (2) will give the desired result in the interior of the region, but on the border of a region, the pixel label may be ambiguous because of two different classes of neighbors and may cause a little distortion of the boundary.

The maximization of the global criterion (1) means that we are seeking a local maximum close to the initial labeling $\boldsymbol{p}_i^{(o)}$ $(i = 1, \cdots, N)$ subject to the constraints that the $\boldsymbol{p}_i$'s are the probability vectors. The maximization of (1) results in a reduced inconsistency and ambiguity. Inconsistency is defined as the error between $\boldsymbol{p}_i$ and $\boldsymbol{q}_i$. Intuitively, this means the discrepancy between what every pixel "thinks" about its own labeling $(\boldsymbol{p}_i)$ and what its neighbors "think" about it $(\boldsymbol{q}_i)$. Ambiguity is measured by the quadratic entropy and results from the fact that initial labeling $\boldsymbol{p}_i^{(o)}$ is ambiguous ($\boldsymbol{p}_i^{(o)}$ are not unit vectors). We are therefore trying to align the vectors $\boldsymbol{p}_i$ and $\boldsymbol{q}_i$ while turning them into unit vectors. Indeed, it can be easily seen that each term $\boldsymbol{p}_i \cdot \boldsymbol{q}_i$ is maximum for $\boldsymbol{p}_i = \boldsymbol{q}_i$ (maximum consistency) and $\boldsymbol{p}_i = \boldsymbol{q}_i$ = unit vector (maximum unambiguity).

The initial assignment of probabilities to every pixel is very important. It affects the convergence rate and the results of relaxation schemes. The simplest way to compute the initial probabilities [6], [11] is to define

$$p_i(\lambda_1) = \frac{I(i)}{G - 1} \tag{4}$$

where $I(i)$ is the intensity at the ith pixel and G is the number of possible gray levels $(0 \leqslant I(i) \leqslant G - 1)$. This has the problem of completely ignoring any *a priori* information that we may have about the contents of the image. It is possible to include a rough knowledge of the relative number of white pixels versus black pixels. If $\bar{I}$ denotes the mean of the image, we may have an approximate idea of the value of the number of white versus black pixels (the ratio $r_0 = n_{\text{white}}/n_{\text{black}}$). If we estimate this ratio for the image, we obtain

$$r = \frac{n_{\text{white}}}{n_{\text{black}}} = \frac{E(\lambda_1)}{E(\lambda_2)} = \frac{\frac{1}{N}\sum_i p_i(\lambda_1)}{\frac{1}{N}\sum_i p_i(\lambda_2)} = \frac{\bar{I}}{G - 1 - \bar{I}}. \tag{5}$$

If we know *a priori* that there are more white than black pixels in the image, we may want to modify the distribution of gray levels so as to make the ratio r closer to the known value r_0. One very simple way to do so is to define

$$I' = \text{FACT} * (I - \bar{I}) + \bar{I}_0 \tag{6}$$

where $\bar{I}_0$ is a desired mean and FACT is a function of the intensity which is taken to be equal to 1 if $I > \bar{I}$ and less than 1 if $I < \bar{I}$. In our experiments, the initial assignment of probabilities has been obtained by

$$p_i(\lambda_1) = \text{FACT} * \left(\frac{I(i) - \bar{I}}{255} \right) + 0.5. \tag{7}$$

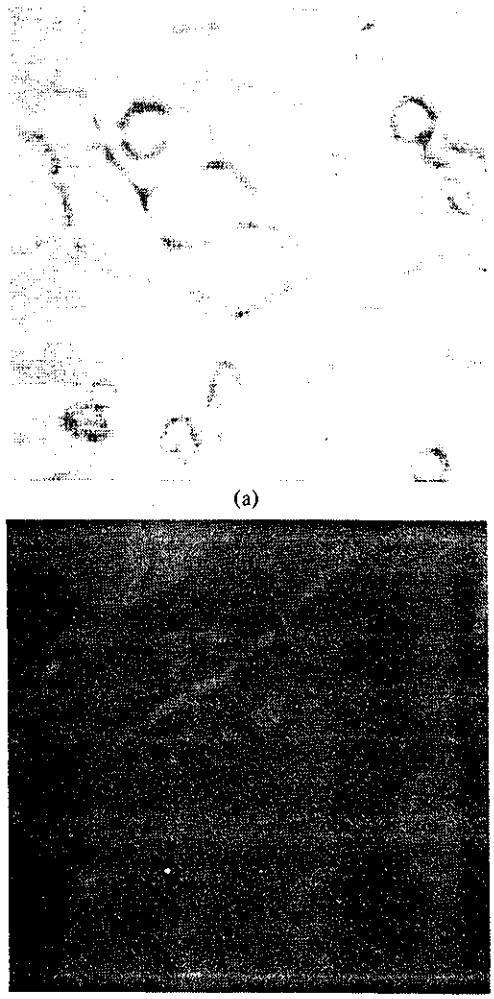


Fig. 1. Two typical 128 × 128, 8-bit images. (a) Cell image. (b) An aerial image.

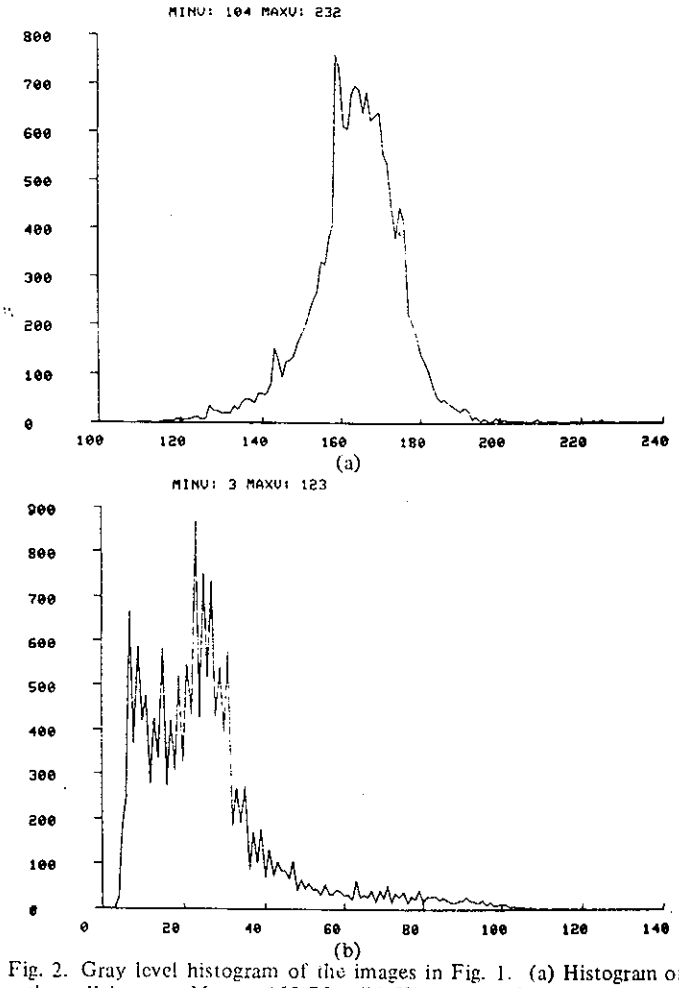


Fig. 2. Gray level histogram of the images in Fig. 1. (a) Histogram of the cell image. Mean = 163.76. (b) Histogram of the aerial image. Mean = 26.47.

When $I(i) < \bar{I}$, FACT has usually been taken between 0.7 and 1. Of course, if the first term of (7) happens to be greater than 0.5 or less than −0.5, then a probability of one or zero, respectively, is assigned to that pixel.

The gradient of the criterion C in (1) is obtained as

$$\frac{\partial C}{\partial p_i(\lambda_1)} = 2q_i(\lambda_1) \tag{8}$$

$$\frac{\partial C}{\partial p_i(\lambda_2)} = 2q_i(\lambda_2) \tag{9}$$

and the iteration of the $\boldsymbol{p}_i$'s is given by

$$p_i^{(n+1)}(\lambda_1) = p_i^{(n)}(\lambda_1) + \rho^{(n)} P_i^{(n)}\left[\frac{\partial C}{\partial p_i(\lambda_1)}\right] \tag{10}$$

$$p_i^{(n+1)}(\lambda_2) = p_i^{(n)}(\lambda_2) + \rho^{(n)} P_i^{(n)}\left[\frac{\partial C}{\partial p_i(\lambda_2)}\right] \tag{11}$$

where $\rho^{(n)}$ is a positive step size.

In order to have $p_i^{(n+1)}(\lambda_1) + p_i^{(n+1)}(\lambda_2) = 1$, the projection of the gradient should be such that

$$P_i^{(n)}\left[\frac{\partial C}{\partial p_i(\lambda_1)}\right] = 2q_i(\lambda_1) - \frac{1}{2}\left[\frac{\partial C}{\partial p_i(\lambda_1)} + \frac{\partial C}{\partial p_i(\lambda_2)}\right] \tag{12}$$

and

$$P_i^{(n)}\left[\frac{\partial C}{\partial p_i(\lambda_2)}\right] = 2q_i(\lambda_2) - \frac{1}{2}\left[\frac{\partial C}{\partial p_i(\lambda_1)} + \frac{\partial C}{\partial p_i(\lambda_2)}\right], \tag{13}$$

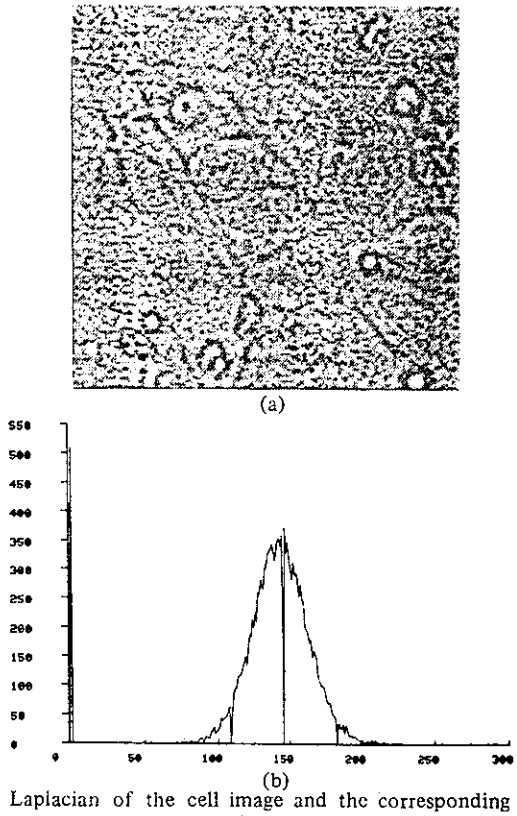


Fig. 3. Laplacian of the cell image and the corresponding gray level histogram.

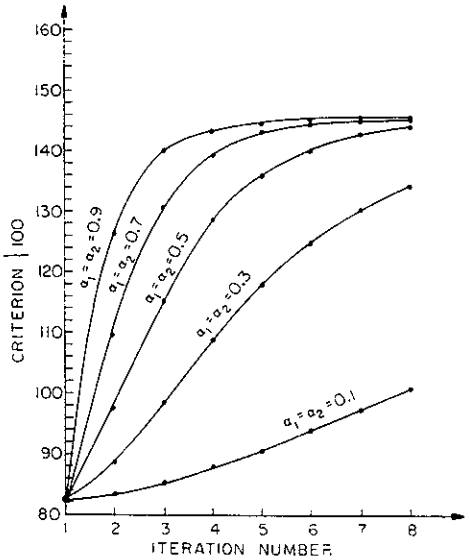
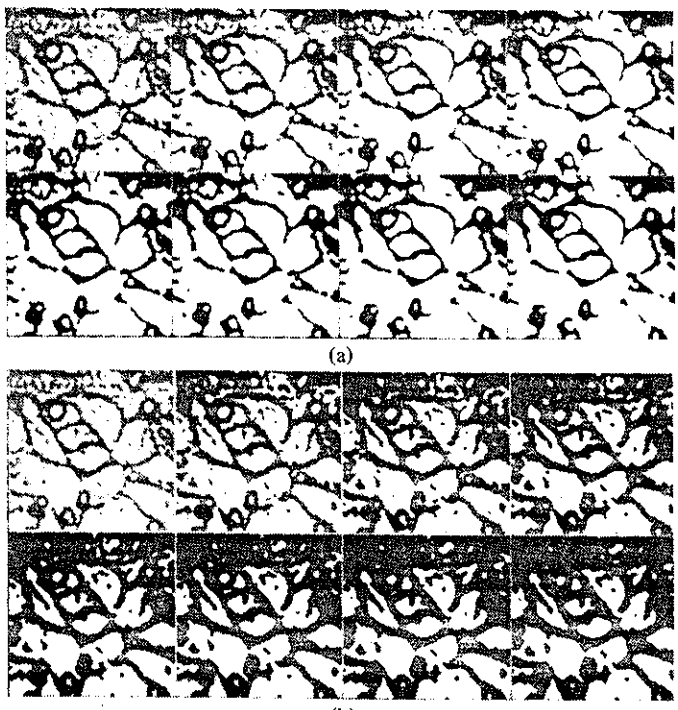


Fig. 4. Variation of criterion (1) with the iteration number for various values of α_1 and α_2 for the cell image. α_1/α_2 = FACT = 1 in all cases.



(a) (b)

Fig. 5. Effect of different values of α_1 and α_2 on the cell image. FACT = 1 in both cases. In each figure, the first row contains images for the first four iterations, and the second for the next four iterations. This figure illustrates the biasing of a class. (a) $\alpha_1 = 0.4, \alpha_2 = 0.2$. (b) $\alpha_1 = 0.2, \alpha_2 = 0.4$.

but

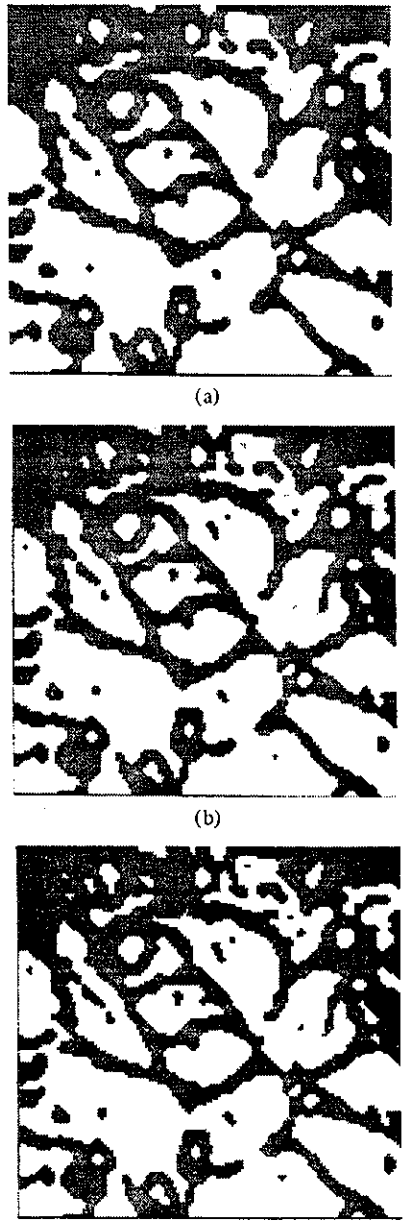
$$\frac{\partial C}{\partial p_i(\lambda_1)} + \frac{\partial C}{\partial p_i(\lambda_2)} = 2.$$

So the iterations in (10) and (11) reduce to

$$p_i^{(n+1)}(\lambda_1) = p_i^{(n)}(\lambda_1) + \rho^{(n)}\,[2q_i(\lambda_1) - 1] \tag{14}$$

$$p_i^{(n+1)}(\lambda_2) = p_i^{(n)}(\lambda_2) + \rho^{(n)}\,[1 - 2q_i(\lambda_1)]. \tag{15}$$

Normally, $\rho^{(n)}$ is kept constant for all pixels during each iteration, and it is determined to have the largest possible value such that the p_i's at the $(n+1)$st iteration still lie in the bounded convex region of the $2N$-dimensional Euclidean space defined



(a) (b) (c)

Fig. 6. Results showing that for a fixed ratio of α_1 and α_2, increasing both of them by a constant factor increases the speed of convergence. In each figure, FACT = 1 is taken. (a) $\alpha_1 = \alpha_2 = 0.2$, iteration 8. (b) $\alpha_1 = \alpha_2 = 0.5$, iteration 4. (c) $\alpha_1 = \alpha_2 = 0.9$, iteration 3.

by $p_i(\lambda_1) + p_i(\lambda_2) = 1$ and $p_i(\lambda_k) \geq 0$, $k = 1, 2$ and $i = 1, \ldots, N$. However, in the two-class case considered, it is easier to compute a $\rho_i^{(n)}$ for each pixel. This leads to a faster convergence rate. The maximum possible value for $\rho_i^{(n)}$ is obtained from (14) by setting $p_i^{(n+1)}(\lambda_1) = 1$ when $2q_i(\lambda_1) - 1 > 0$ and $p_i^{(n+1)}(\lambda_1) = 0$ when $2q_i(\lambda_1) - 1 < 0$. Thus,

$$\rho_{i\max}^{(n)} = \begin{cases} \left(\dfrac{1 - p_i^{(n)}(\lambda_1)}{2q_i(\lambda_1) - 1}\right), & \text{if } 2q_i(\lambda_1) - 1 > 0 \qquad (16) \\ \left(\dfrac{p_i^{(n)}(\lambda_1)}{1 - 2q_i(\lambda_1)}\right), & \text{if } 2q_i(\lambda_1) - 1 < 0. \qquad (17) \end{cases}$$

Since we want to be able to control the rate of convergence and the number of pixels within each class, we actually took

$$p_i^{(n+1)}(\lambda_k) = p_i^{(n)}(\lambda_k) + \rho_i^{(n)}\,[2q_i(\lambda_k) - 1] \tag{18}$$

and

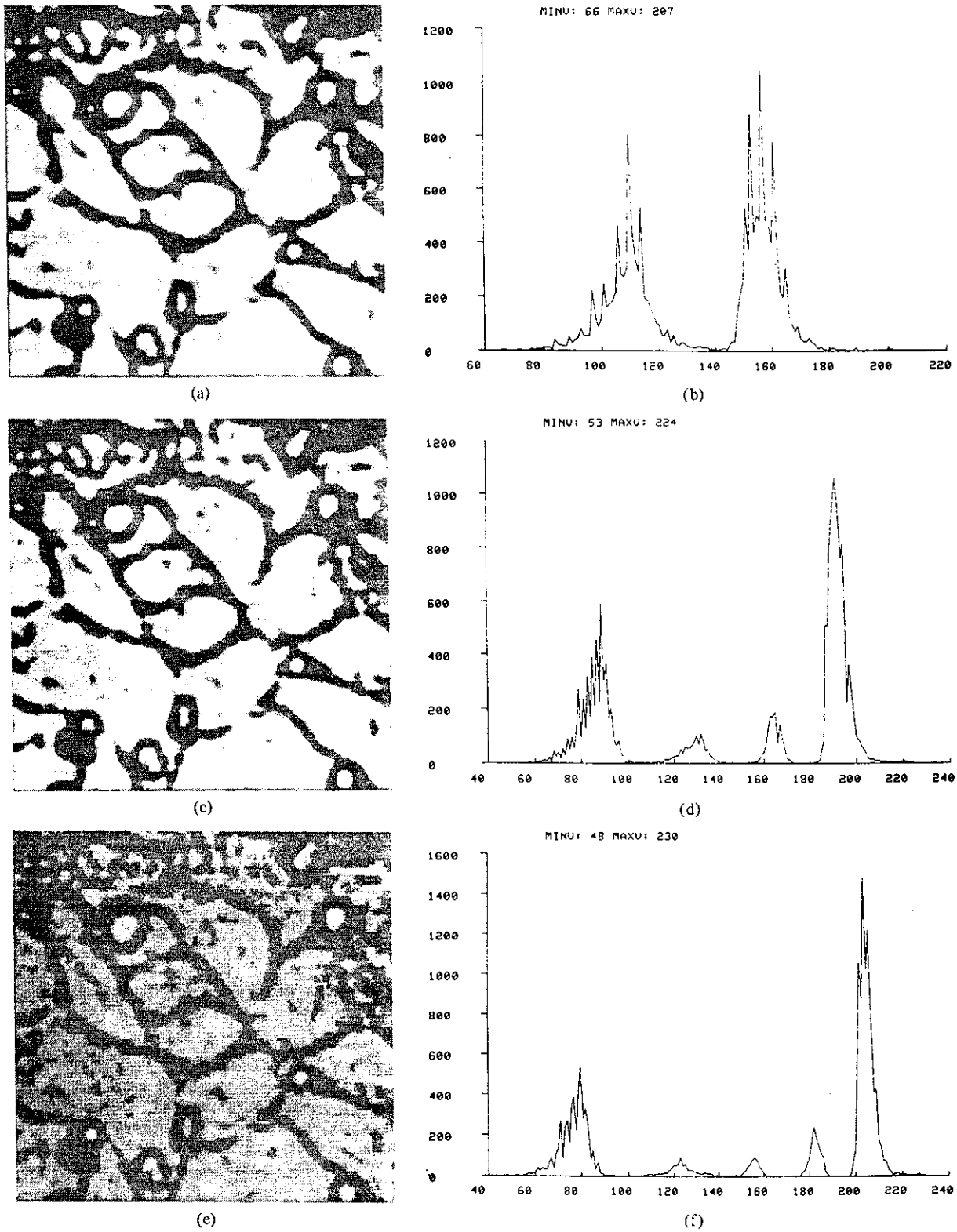
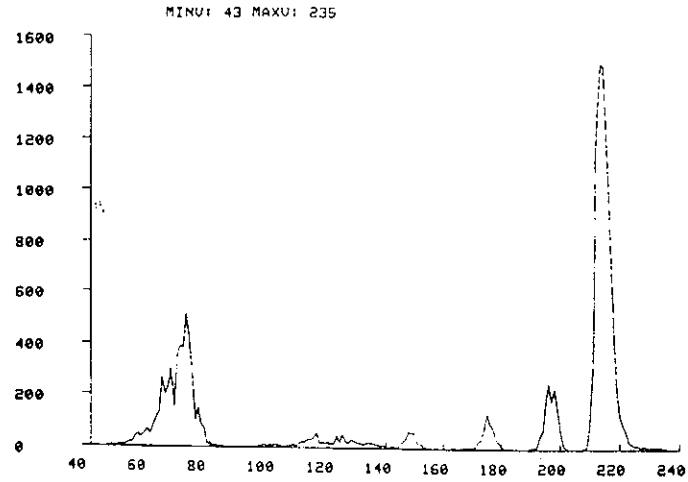
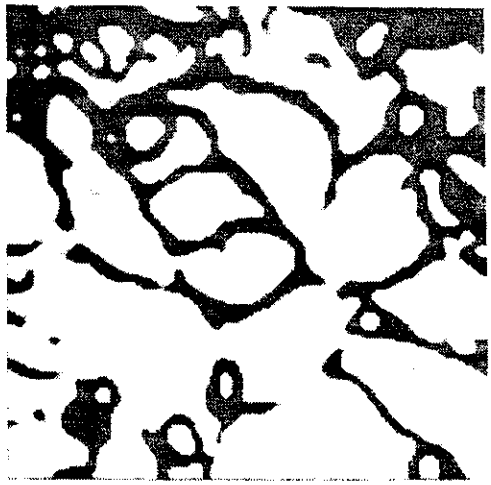


Fig. 7. Results of gradient relaxation method at various iterations and corresponding histograms for the cell image. FACT = 0.9, $\alpha_1 = 0.2$, $\alpha_2 = 0.1$. (a) Iteration 1. (b) Histogram of Fig. 7(a). (c) Iteration 3. (d) Histogram of Fig. 7(c). (e) Iteration 4. (f) Histogram of Fig. 7(e).

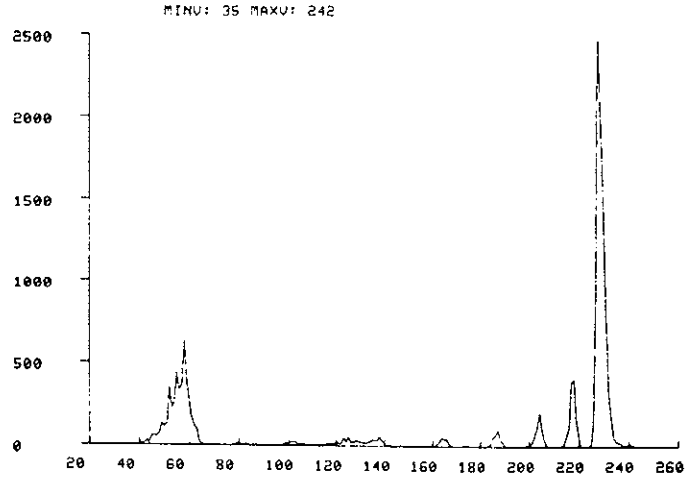
(g)



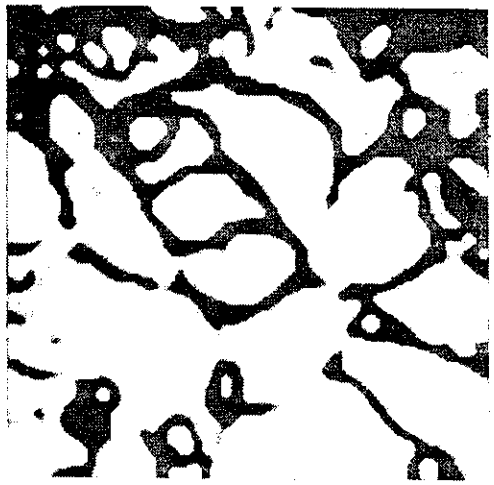
(h)



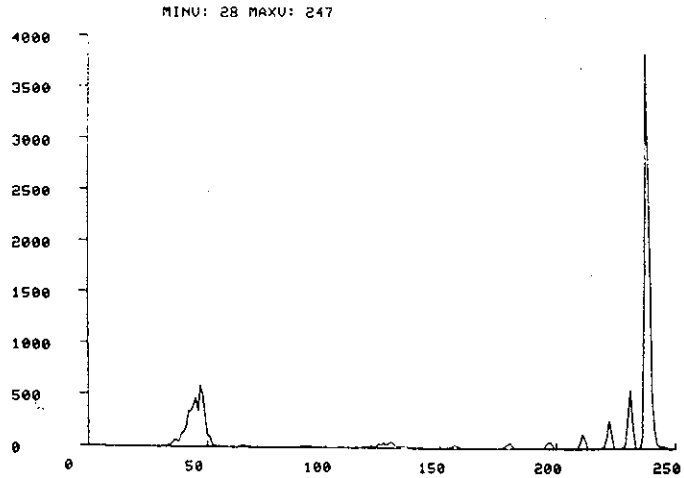
(i)



(j)



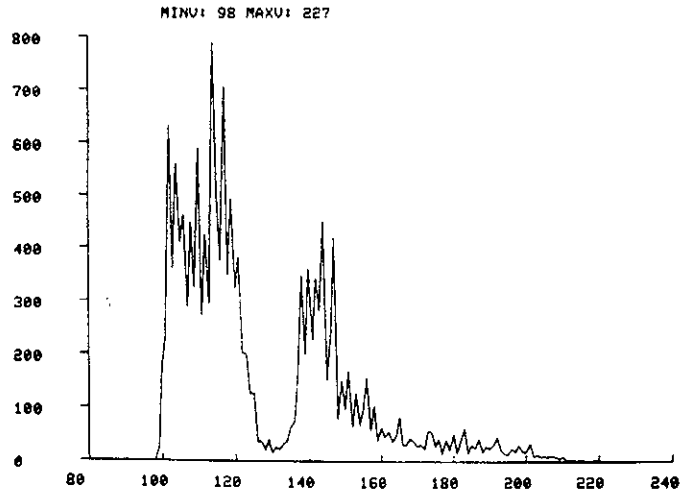
(k)



(l)

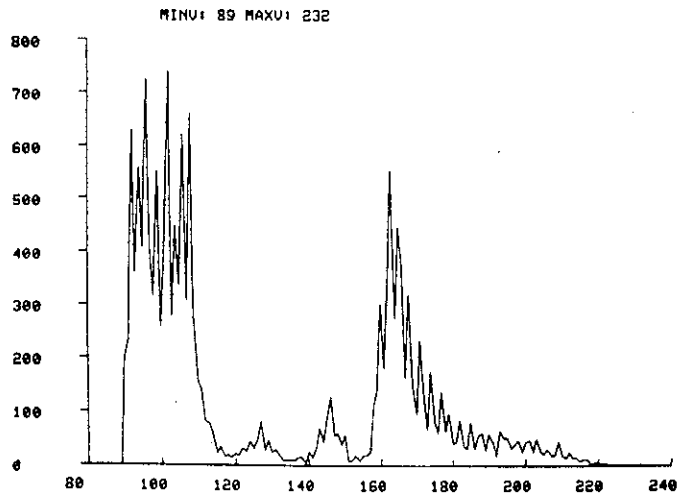
Fig. 7. (*Continued.*) (g) Iteration 5. (h) Histogram of Fig. 7(g). (i) Iteration 7. (j) Histogram of Fig. 7(i). (k) Iteration 9. (l) Histogram of Fig. 7(k).

(a)



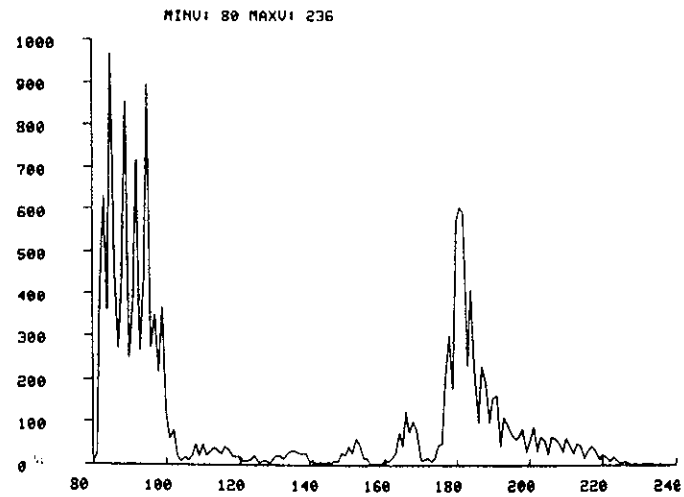
(b)

(c)



(d)

(e)



(f)

Fig. 8. Results of gradient relaxation method at various iterations and corresponding histograms for the aerial image. FACT = 1, $\alpha_1 = 0.1$, $\alpha_2 = 0.05$. (a) Iteration 1. (b) Histogram of Fig. 8(a). (c) Iteration 3. (d) Histogram of Fig. 8(c). (e) Iteration 5. (f) Histogram of Fig. 8(e).

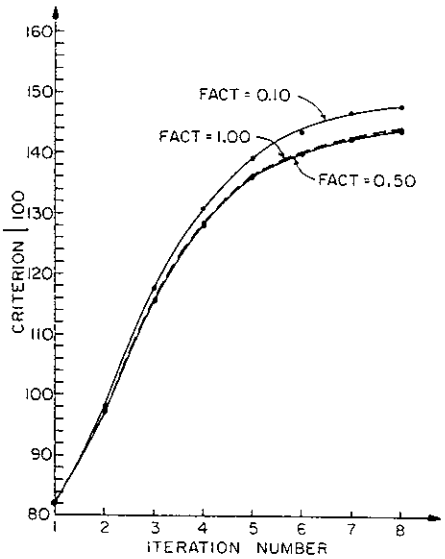


Fig. 9. Variation of criterion with the iteration number for various values of FACT for the cell image. $\alpha_1 = \alpha_2 = 0.50$ in all cases.

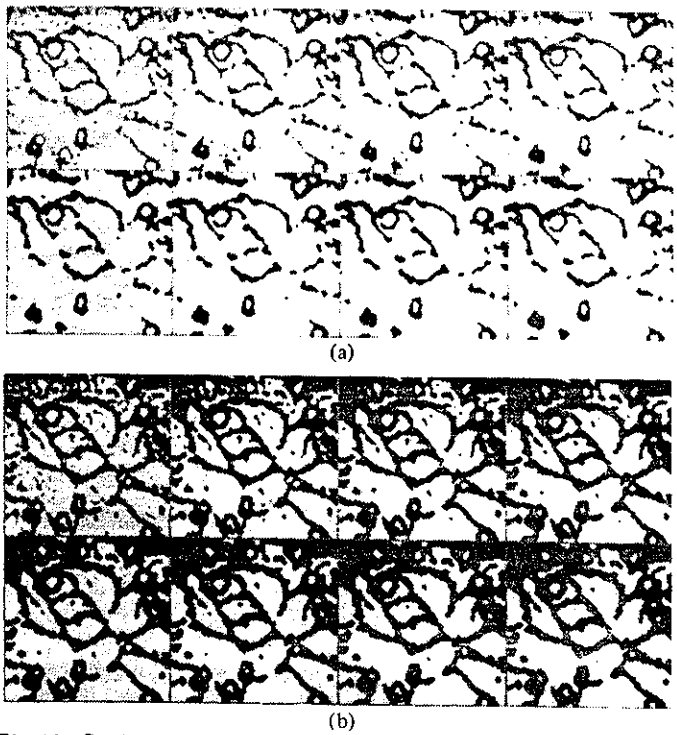


Fig. 10. Gradient relaxation results at various iterations for different values of FACT for the cell image. $\alpha_1 = \alpha_2 = 0.5$ in both cases. (a) FACT = 0.1. (b) FACT = 1.0.

$$\rho_i^{(n)} = \begin{cases} \alpha_1 \rho_{i\max}^{(n)}, & \text{if } 2q_i(\lambda_1) - 1 > 0 \\ \alpha_2 \rho_{i\max}^{(n)}, & \text{if } 2q_i(\lambda_1) - 1 < 0 \end{cases} \tag{19}$$

where $k = 1, 2$ and α_1 and α_2 are constants less than unity. In the actual implementation of the algorithm, we only need to evaluate (18) for $k = 1$ or 2 only since $p_i^{(n+1)}(\lambda_1) + p_i^{(n+1)}(\lambda_2) = 1$. A side effect of computing $\rho_i^{(n)}$ for every pixel is that we may not be following the gradient exactly. However, it can be expected that we are approximately in the direction of the gradient, and the criterion (1) is still maximized. It is our conjecture that for the two-class case, the criterion will always increase. Fig. 4 shows the behavior of the criterion as the iteration number increases for the cell image when $\alpha_1 = \alpha_2$. The same behavior is obtained when α_1 and α_2 are not equal. The values of α_1 and α_2 can be used to bias a class and control the speed of convergence, hence the control over the relaxation process. Figs. 5 and 6 illustrate these effects for different values of α_1 and α_2 on the cell image.

Fig. 5 shows how changing the ratio of α_1 and α_2 allows the control of where we converge by biasing one class. At every step, the probabilities are displayed as gray levels, i.e., $I^{(n)}(i) = p_i^{(n)}(\lambda_1) * (G - 1)$. From Fig. 4, it can be seen that for a fixed ratio of α_1 and α_2, increasing both of them by a constant factor increases the speed of convergence, and Fig. 6 verifies the fact that indeed we converge towards a similar result. What we seem to be losing when increasing α_1 and α_2 is some amount of smoothing: there are a few more isolated small black blobs in Fig. 6(b) and (c) than in Fig. 6(a). The magnitude of α_1 and α_2 controls the degree of smoothing at each iteration and their ratio controls the bias. Therefore, changing the values of α_1 and α_2 not only allows one to control where one wants to converge, but also how fast one converges. Thus, control over the relaxation process is achieved.

Figs. 7 and 8 show the results of gradient relaxation method at various iterations and corresponding histograms for the cell and aerial image, respectively. Observe that at the first iteration itself, we get two peaks separated by a valley which can be used to automatically select the threshold to obtain segmentation. As the number of iterations increases, the two peaks move apart, average brightness increases, and the convergence of probabilities takes place as expected. When the peaks are far apart, thresholding can be done at the mean value. Also, it has been found that the actual value of FACT does not affect the value of criterion (see Fig. 9) very much, hence the speed of convergence, but it affects the segmentation results as expected. Fig. 10 shows the gradient relaxation results at various iterations for different values of FACT when $\alpha_1 = \alpha_2 = 0.5$ for the cell image.

B. Nonlinear Probabilistic Relaxation Algorithm

Rosenfeld [6], [7] has used the nonlinear probabilistic relaxation algorithm for thresholding. In this algorithm, the probability of labeling the ith pixel with label λ_k at the $(n+1)$st iteration is given by

$$p_i^{(n+1)}(\lambda_k) = \frac{p_i^{(n)}(\lambda_k)\, q_i^{(n)}(\lambda_k)}{\sum_{l=1}^{2} p_i^{(n)}(\lambda_l)\, q_i^{(n)}(\lambda_l)},$$

$$k = 1, 2, \quad i = 1, \cdots, N. \tag{20}$$

Assuming the same compatibility function and the initial assignment of probabilities as in the discussion of the gradient technique, the results on the images shown in Fig. 1 are illustrated in Figs. 11 and 12. It can be seen from the histogram plots that in the first few iterations, histogram equalization takes place, resulting in the enhanced image, and then black and white classes get separated. However, in this method, unlike the gradient method, we do not have any control over where we converge and how fast we converge. As an aside, Helland [12] has shown that the compatibility function defined in (2) results in the maximum convergence speed of the Peleg relaxation scheme [13] for the two-class problem. In this case, Peleg's updating rule becomes the same as (20). Thus, it is possible to control the convergence rate of the nonlinear relaxation method by adjusting other parameters such

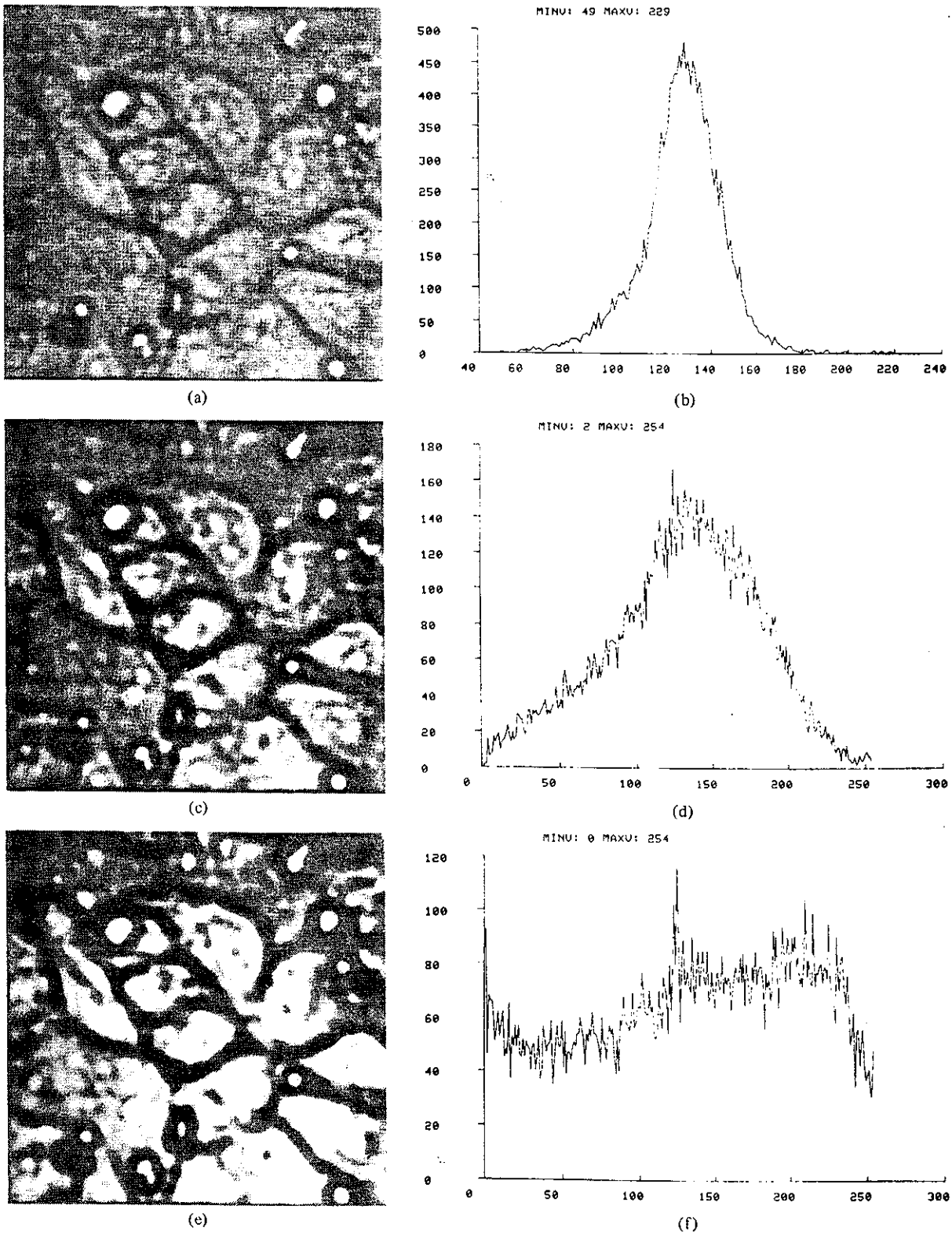
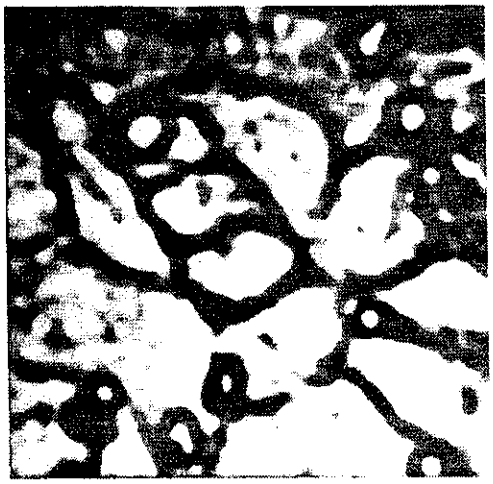
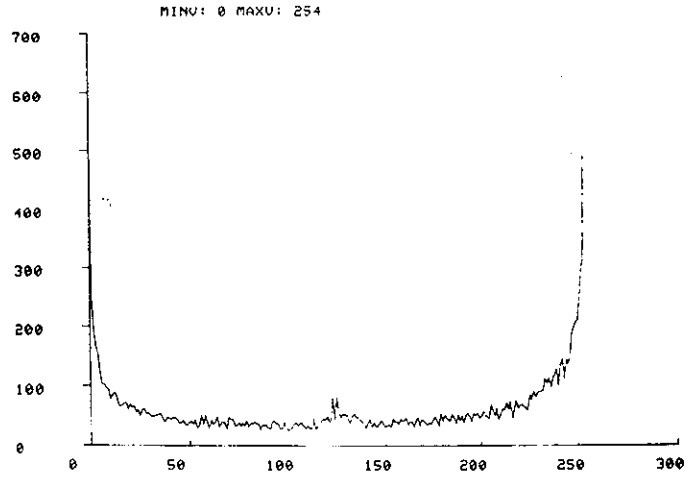


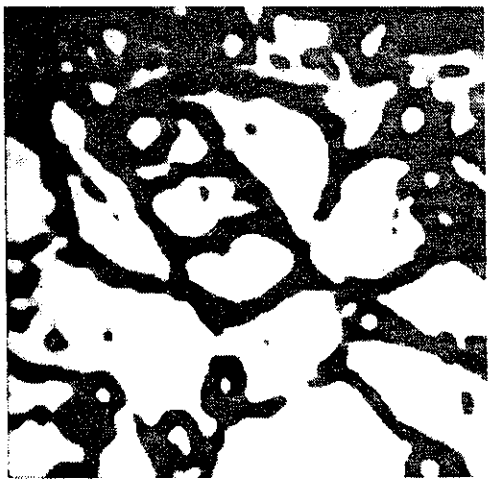
Fig. 11. Results of nonlinear relaxation method at various iterations and corresponding histograms for the cell image. FACT = 0.9. (a) Iteration 1. (b) Histogram of Fig. 11(a). (c) Iteration 3. (d) Histogram of Fig. 11(c). (e) Iteration 4. (f) Histogram of Fig. 11(e).



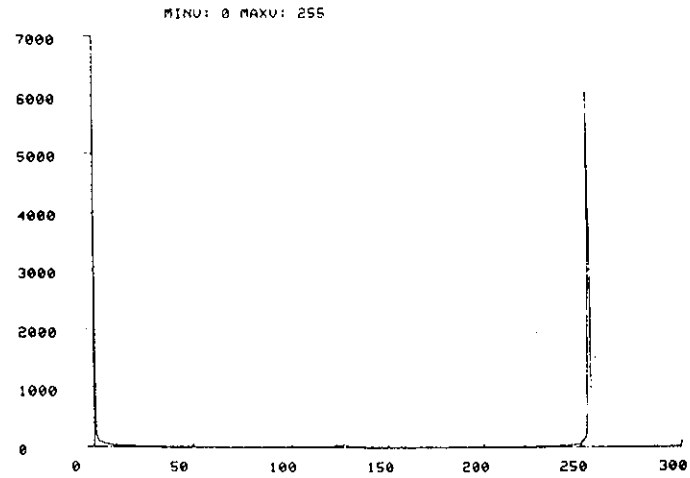
(g)



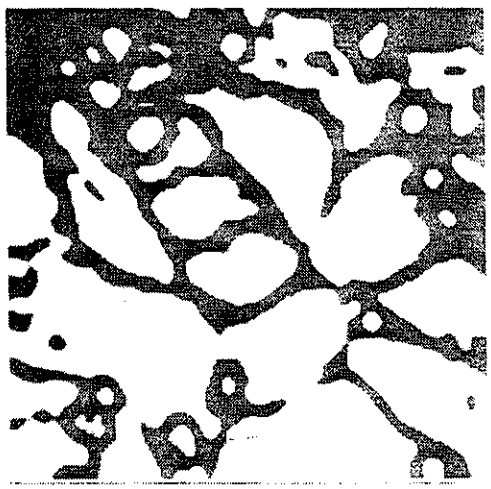
(h)



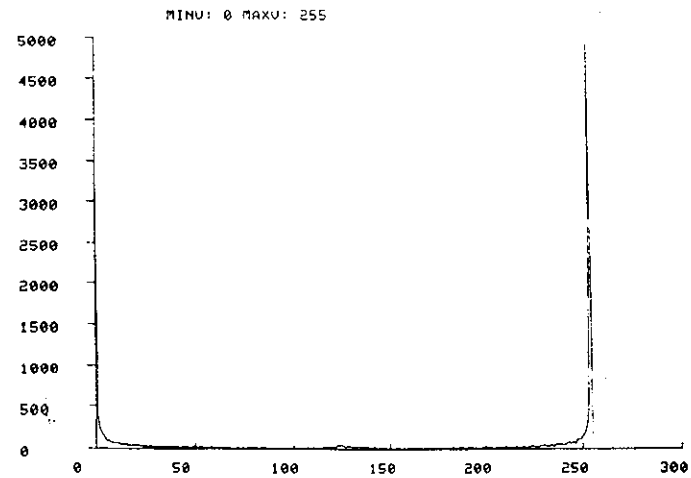
(i)



(j)



(k)



(l)

Fig. 11. (*Continued.*) (g) Iteration 5. (h) Histogram of Fig. 11(g). (i) Iteration 7. (j) Histogram of Fig. 11(i). (k) Iteration 9. (l) Histogram of Fig. 11(k).

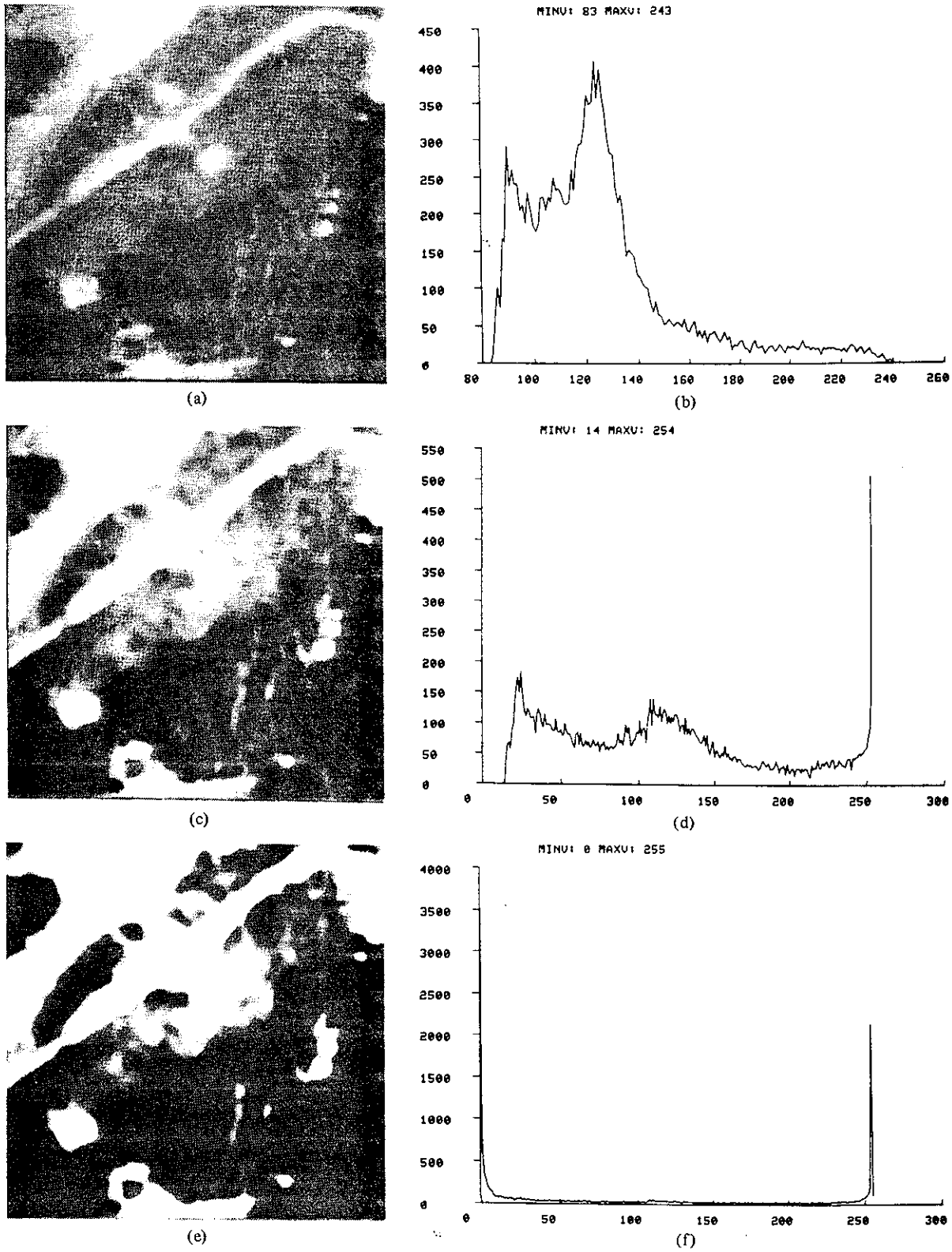


Fig. 12. Results of nonlinear relaxation method at various iterations and corresponding histograms for the aerial image. FACT = 1. (a) Iteration 1. (b) Histogram of Fig. 12(a). (c) Iteration 3. (d) Histogram of Fig. 12(c). (e) Iteration 5. (f) Histogram of Fig. 12(e).

as the compatibility coefficients. However, the results of relaxation can be sensitive to the choice of these coefficients.

IV. Conclusion

From the images and histograms shown in Figs. 7, 8, 11, and 12, two observations can be made. First, the gradient and nonlinear relaxation methods provide comparable segmentation results. Second, it is not possible to compare these two methods directly because they do not converge to exactly the same limit. However, as illustrated in Figs. 5, 6, and 10, it is only the gradient method that provides the control over the relaxation process by choosing the α_1, α_2, and FACT param-

eters which can be tuned to obtain the desired segmentation results at a faster rate. The magnitude of the α's controls the degree of smoothing at each iteration and their ratio controls the bias. The magnitude of FACT controls the initial assignment of probabilities. Although we have presented a two-class problem for the segmentation of images having unimodal distributions, the method can be easily generalized to include more classes [14].

REFERENCES

[1] J. S. Weszka, "A survey of threshold selection techniques," *Comput. Graphics Image Processing*, vol. 7, pp. 259–265, 1978.

[2] R. Ohlander, K. Price, and D. Reddy, "Picture segmentation using a recursive region splitting method," *Comput. Graphics Image Processing*, vol. 8, pp. 313–333, 1978.

[3] A. K. Jain, S. P. Smith, and E. Backer, "Segmentation of muscle cell pictures: A preliminary study," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-2, pp. 232–242, May 1980.

[4] A. Rosenfeld and L. S. Davis, "Iterative histogram modification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-8, pp. 300–302, 1978.

[5] S. Peleg, "Iterative histogram modification, 2," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-8, pp. 555–556, 1978.

[6] A. Rosenfeld, "Image understanding project status report," in *Proc. DARPA Image Understanding Workshop*, Apr. 1979, pp. 14–24.

[7] A. Rosenfeld, R. Hummel, and S. W. Zucker, "Scene labeling by relaxation operations," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, pp. 420–433, 1976.

[8] R. Nevatia and K. R. Babu, "Linear feature extraction and description," *Comput. Graphics Image Processing*, vol. 13, pp. 257–269, 1980.

[9] O. D. Faugeras and M. Berthod, "Improving consistency and reducing ambiguity in stochastic labeling: An optimization approach," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-3, pp. 412–424, July 1981.

[10] M. Berthod and O. D. Faugeras, "Using context in the global recognition of a set of objects: An optimization approach," in *Proc. 8th World Comput. Congr. (IFIP'80)*, Tokyo, Japan, pp. 695–698.

[11] G. Fekete, "Relaxation: Evaluation and applications," Dep. Comput. Sci., Univ. Maryland, College Park, Tech. Rep. 796, Aug. 1979.

[12] A. R. Helland, "Convergence properties of two label relaxation," in *Proc. DARPA Image Understanding Workshop*, Apr. 1980, pp. 176–181.

[13] S. Peleg, "A new probabilistic relaxation scheme," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-2, pp. 362–369, July 1980.

[14] B. Bhanu and O. D. Faugeras, "Segmentation of images having unimodal gray level distributions," Image Processing Inst., Univ. Southern California, Los Angeles, USCIPI Rep. 960, Mar. 1980.

A Medial Axis Transformation for Grayscale Pictures

SHYUAN WANG, AZRIEL ROSENFELD, AND ANGELA Y. WU

Abstract—Blum's medial axis transformation (MAT) for binary pictures yields medial axis points that lie midway between opposite borders of a region or along angle bisectors. This note discusses a generalization of the MAT in which a score is computed for each point P of a grayscale picture based on the gradient magnitudes at pairs of points that have P as their midpoint. These scores are high at points that lie midway between pairs of antiparallel edges or along angle bisectors, so that they define a MAT-like "skeleton," which we may call the GRADMAT. However, this skeleton is rather sensitive to the presence of noise edges or to irregularities in the region edges, and it also is subject to artifacts created by pairs of edges belonging to different objects.

Index Terms—Image approximation, image processing, medial axis transformation, region representation.

Manuscript received January 9, 1980; revised December 28, 1981. This work was supported by the U.S. Air Force Office of Scientific Research under Grant AFOSR-77-3271.

S. Wang and A. Rosenfeld are with the Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD 20742.

A. Y. Wu is with the Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD 20742 and the Department of Mathematics, Statistics, and Computer Science, American University, Washington, DC 20016.

I. INTRODUCTION

In the early 1960's, Blum [1] introduced the medial axis transformation (MAT) of a set S; this is basically the set of centers and radii of the maximal disks that are contained in S, or equivalently, the set of points of S whose distances to the complement $\bar{S}$ are local maxima, together with these distances. It is not hard to see that medial axis points tend to lie midway between opposite borders of S or along the bisectors of angles formed by the borders. Thus, these points constitute a kind of "skeleton" of S. For an introduction to the MAT, see [2, sect. 9.2.3].

Blum's MAT is defined for a picture only after the picture has been segmented into S and $\bar{S}$. Several generalizations of the MAT to grayscale pictures have been suggested. We can define the gray-weighted length of a path as proportional to the sum of the gray levels at the points of the path; the gray-weighted distance between two points can then be defined as the lowest gray-weighted length of a path between them, and the gray-weighted MAT (GMAT) of a picture can be defined as the set of points whose gray-weighted distances to the set of 0's in the picture are local maxima, together with these distances [3]. Note that this definition still requires segmentation of the picture, since it treats 0's as "background" and regions of nonzero values as objects. Another generalization is based on finding maximal homogeneous disks in the given picture; the set of centers, radii, and average gray levels of these disks defines a generalized MAT, called the SPAN (spatial piecewise approximation by neighborhoods), since this information can be used to generate approximations to the picture [4].

This note discusses a generalization of the MAT in which a score is computed for each point P of the picture based on the gradient magnitudes at pairs of points that have P as their midpoint. These scores are high at points that lie midway between pairs of antiparallel edges or along angle bisectors, so that they define a MAT-like "skeleton," which we may call the GRADMAT. However, this skeleton is rather sensitive to the presence of noise edges or to irregularities in the region edges, and it is also subject to artifacts created by pairs of edges belonging to different objects. Note that the GRADMAT scores are not thresholded; the GRADMAT is a "gray" skeleton, not a binary one.

II. THE GRADMAT

The basic idea of the GRADMAT is to compute a score for every P based on the gradient magnitudes at all pairs of points that have P as their midpoint. Evidently, this score will be very high at the center of a circle (or region that has a high degree of central symmetry); and it will also be high along the midline of a parallel-sided strip. There will be weaker responses at points that lie on local axes of symmetry, e.g., on angle bisectors, since such points are midway between at least one pair of edges.

These examples show that the GRADMAT is in many ways analogous to the MAT. However, it should be realized that the analogy is only partial. To see this, consider the GRADMAT

0162-8828/82/0700-0419$00.75

BB

IEEE TRANSACTIONS ON

PATTERN ANALYSIS AND MACHINE INTELLIGENCE

JULY 1982 VOLUME PAMI-4 NUMBER 4 (ISSN 0162-8828)

PUBLISHED BY THE IEEE COMPUTER SOCIETY

In Cooperation With

Aerospace and Electronic Systems Society
Control Systems Society
Engineering in Medicine and Biology Society
Information Theory Group
Sonics and Ultrasonics Group
Systems, Man, and Cybernetics Society

PAPERS

CORRESPONDENCE

tern Recognition, Kyoto, Japan, Session-B8, Nov. 1978, pp. 1047-1049.
[21] D. B. Fry, A. B. Abramson, P. D. Eimas, and A. M. Liberman, "The identification and discrimination of synthetic vowels," *Language and Speech*, vol. 5, pp. 171-189, 1962.
[22] K. N. Stevens and A. S. House, "Perturbation of vowel articulation by consonantal context: An acoustical study," *J. Speech Hearing Res.*, vol. 6, pp. 111-128, 1963.
[23] B. E. F. Lindblom and M. Studdert-Kenedy, "On the role of formant transitions in vowel recognition," *J. Acoust. Soc. Amer.*, vol. 42, p. 830, 1967.
[24] J. B. Miller and W. A. Ainsworth, "Identification of synthetic isolated vowels and vowels in h-d context," *Acustica*, vol. 27, pp. 278-282, 1972.
[25] P. C. Delattre, A. M. Liberman, F. S. Cooper, and L. J. Gerstman, "An experimental study of the acoustic determination of vowel color: Observations on one-and-two formant vowels synthesized from spectrographic patterns," *World*, vol. 8, pp. 195-210, 1952.
[26] L. J. Gerstman, "Classification of self normalized vowels," *IEEE Trans. Audio Electroacoust.*, vol. AU-16, 1968.
[27] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. New York, Wiley, 1958, p. 147.
[28] P. Ladgefoged, *Three Areas of Experimental Phonetics*. Oxford: Oxford Univ. Press, 1967.

Correction to "Segmentation of Images Having Unimodal Distributions"

BIR BHANU AND OLIVIER D. FAUGERAS

In the above paper,[1] three figures have been inadvertently switched. The correct figures are as follows.

Fig. 7(a) should be Fig. 7(e).
Fig. 7(e) should be Fig. 7(a).
Fig. 8(c) should be Fig. 8(e).
Fig. 8(e) should be Fig. 8(c).
Fig. 11(j) should be Fig. 11(l).
Fig. 11(l) should be Fig. 11(j).

Manuscript received July 12, 1982.

B. Bhanu is with the Aeronutronic Division, Ford Aerospace and Communications Corporation, Newport Beach, CA 92660.
O. D. Faugeras is with INRIA, Rocquencourt, France, and the University of Paris XI, Paris, France.

[1] B. Bhanu and O. D. Faugeras, *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-4, pp. 408-419, July 1982.

0162-8828/82/1100-0689$00.75

IEEE TRANSACTIONS ON

PATTERN ANALYSIS AND MACHINE INTELLIGENCE

NOVEMBER 1982 VOLUME PAMI-4 NUMBER 6 (ISSN 0162-8828)

PUBLISHED BY THE IEEE COMPUTER SOCIETY

In Cooperation With

Aerospace and Electronic Systems Society
Control Systems Society
Engineering in Medicine and Biology Society
Information Theory Group
Sonics and Ultrasonics Group
Systems, Man, and Cybernetics Society

PAPERS

CORRESPONDENCE