COMPUTER VISION, GRAPHICS, AND IMAGE PROCESSING 41, 233–260 (1988)

SURVEY

A Survey of Thresholding Techniques*

P. K. SAHOO, S. SOLTANI, AND A. K. C. WONG

Department of Systems Design Engineering, University of Waterloo, Waterloo, Canada N2L 3G1

AND

Y. C. CHEN

Department of Electrical Engineering, University of Waterloo, Waterloo, Canada N2L 3G1

Received July 7, 1984; accepted July 9, 1987

In digital image processing, thresholding is a well-known technique for image segmentation. Because of its wide applicability to other areas of the digital image processing, quite a number of thresholding methods have been proposed over the years. In this paper, we present a survey of thresholding techniques and update the earlier survey work by Weszka (*Comput. Vision Graphics & Image Process* **7**, 1978, 259–265) and Fu and Mu (*Pattern Recognit.* **13**, 1981, 3–16). We attempt to evaluate the performance of some automatic global thresholding methods using the criterion functions such as uniformity and shape measures. The evaluation is based on some real world images. © 1988 Academic Press, Inc.

1. INTRODUCTION

A popular tool used in image segmentation is thresholding. This paper represents a survey of a variety of thresholding (also known as binarization) techniques including both global and local thresholding. For the sake of discussion, global techniques are further classified as: point-dependent and region-dependent techniques. Several global thresholding methods are examined in detail to evaluate their performance based on a given set of test images. In this paper, we endeavor to update the earlier survey papers by Weszka [44] and Fu and Mu [13]. We also attempt to evaluate the performance of several global thresholding methods using a shape measure and a region uniformity measure as criterion functions. The evaluation is conducted on a set of images digitized from photographs of a portrait and natural scenes. Although our primary goal is to present a survey of bilevel thresholding methods, extensions of these methods to multithresholding problems are also treated in this paper.

Let N be the set of natural numbers, (x, y) be the spatial coordinate of a digitized image, and $G = \{0, 1, \ldots, l-1\}$ be a set of positive integers representing gray levels. Then, an image function can be defined as the mapping $f: N \times N \to G$. The brightness (i.e., gray level) of a pixel with coordinate (x, y) is denoted as $f(x, y)$.

Let $t \in G$ be a threshold and $B = \{b_0, b_1\}$ be a pair of binary gray levels and $b_0, b_1 \in G$. The result of thresholding an image function $f(\cdot, \cdot)$ at gray level t is a

*This work was supported by the Natural Sciences and Engineering Reearch Council, Canada, under Grants A4716 and G1367.

0734-189X/88 $3.00
Copyright © 1988 by Academic Press, Inc.
All rights of reproduction in any form reserved.

binary image function $f_t : N \times N \rightarrow B$, such that

$$f_t(x, y) = \begin{cases} b_0 & \text{if } f(x, y) < t. \\ b_1 & \text{if } f(x, y) \geq t. \end{cases}$$

In general, a thresholding method is one that determines the value t^* of t based on a certain criterion. If t^* is determined solely from the gray level of each pixel, then the thresholding method is point-dependent. If t^* is determined from the local property (e.g., the local gray level distribution) in the neighborhood of each pixel, then the thresholding method is region-dependent. A global thresholding technique is one that thresholds the entire image with a single threshold value, whereas a local thresholding technique is one that partitions a given image into subimages and determines a threshold for each of these subimages.

Let the number of pixels with gray level i be n_i. Then the total number of pixels in a given image is

$$n = \sum_{i=0}^{l-1} n_i.$$

The probability of occurrence of gray level i is defined as

$$p_i = \frac{n_i}{n}.$$

Also, by convention, the gray level 0 is the darkest and the gray level $l - 1$ is the lightest.

2. GLOBAL THRESHOLDING: POINT DEPENDENT TECHNIQUES

A. *p-tile Method*

One of the earliest thresholding methods is the p-tile method [10]. In this method, an image is assumed to consist of dark objects in a light background. By assuming that the percentage of the object area is known, the threshold is defined as the highest gray level which maps at least $(100 - p)\%$ of the pixels into the objects in the thresholded image. For example, suppose an object occupies 20% of an image, then the image should be thresholded at the highest gray level that allows at least 20% of the pixels to be mapped into the object. Obviously, this method is not applicable to images whose object area is not known.

B. *Mode Method*

For images with distinct objects and background, the histogram of the gray levels will be bimodal. In this case, a threshold can be chosen as the gray level that corresponds to the valley of the histogram. The technique is called the mode method [31]. Though this method is simple, it cannot be applied to images with extremely unequal peaks or to those with broad and flat valleys.

C. *Ostu Method*

This method, as proposed in [27], is based on discriminant analysis. In this method, the threshold operation is regarded as the partitioning of the pixels of an

image into two classes C_0 and C_1 (e.g., objects and background) at gray level t. That is, $C_0 = \{0, 1, \ldots, t\}$ and $C_1 = \{t+1, t+2, \ldots, l-1\}$. Let σ_W^2, σ_B^2, and σ_T^2 be the within-class variance, between-class variance, and the total variance, respectively. An optimal threshold can be determined by minimizing one of the following (equivalent) criterion functions with respect to t:

$$\lambda = \frac{\sigma_B^2}{\sigma_W^2}, \qquad \eta = \frac{\sigma_B^2}{\sigma_T^2}, \qquad \text{and} \qquad \kappa = \frac{\sigma_T^2}{\sigma_W^2}.$$

Of the above three criterion functions, η is the simplest. Thus, the optimal threshold t^* is

$$t^* = \operatorname{Arg} \operatorname*{Min}_{t \in G} \eta,$$

where

$$\sigma_T^2 = \sum_{i=0}^{l-1} (i - \mu_T)^2 p_i, \qquad \mu_T = \sum_{i=0}^{l-1} i p_i,$$

$$\sigma_B^2 = \omega_0 \omega_1 (\mu_1 \mu_0)^2, \qquad \omega_0 = \sum_{i=0}^{t} p_i, \qquad \omega_1 = 1 - \omega_0,$$

$$\mu_1 = \frac{\mu_T - \mu_t}{1 - \omega_0}, \qquad \mu_0 = \frac{\mu_t}{\omega_0}, \qquad \mu_t = \sum_{i=0}^{t} i p_i.$$

D. *Histogram Concavity Analysis Method*

For images with distinct objects and background it is possible to select the threshold from the gray level histogram using the mode method. For some images where valleys may not be found in their gray level histogram, it is often possible to define a good threshold at the "shoulder" of the histogram. Since both valleys and shoulders correspond to the concavities in the histogram, a threshold can be determined by analyzing the concavity of the histogram [35].

Let HS be a histogram defined over a set of gray levels $g_0, g_1, \ldots, g_{l-1}$. Denote the height of the histogram at these gray levels as $h(g_0), h(g_1), \ldots, h(g_{l-1})$, where $h(g_i) \neq 0$ for all i. Thus, HS can be regarded as a two-dimensional region.

To determine the concavities of HS, its convex hull is constructed. It is the smallest convex polygon $\overline{HS}$ which contains HS. The concavities of HS are determined from the set-theoretic difference $HS - \overline{HS}$. Let $\bar{h}(g_i)$ be the height of $\overline{HS}$ at gray level g_i. Possible threshold values are gray levels at which $\bar{h}(g_i) - h(g_i)$ has local maxima. However, not all these maxima are good candidates for thresholding because large concavities can also be introduced by noise spikes. Rosenfeld and De La Torre [35] call such concavities *spurious*. In order to eliminate maxima caused by spurious concavities, a balance measure

$$E_i = \left\{ \sum_{j=g_0}^{g_{i-1}} h(j) \right\} \left\{ \sum_{j=g_i}^{g_{l-1}} h(j) \right\}$$

(*i*) *Co-occurrence Matrix Method*

The co-occurrence matrix is introduced by Haralick for texture analysis [14]. In general, a co-occurrence matrix $M_{(d,\phi)}$ is one whose entries are the relative frequency of occurrence for two neighboring pixels with gray levels i and j, separated by distance d and with orientation ϕ. Obviously, the number of such matrices can be quite large, depending on the choice of d and ϕ. In Ahuja and Rosenfeld [2], the gray level co-occurrence matrix is defined as

$$M = M_{(1,0)} + M_{(1,\pi/2)} + M_{(1,\pi)} + M_{(1,3\pi/2)};$$

that is, the (i, j) element of M is the frequency that gray level i occurs as a 4-neighbor of gray level j.

Because of homogeneity, pixels interior to the objects or background should contribute mainly to the near-diagonal entries of M. Also, pixels near an edge should contribute mainly to the off-diagonal entries of M because of the gray level change near an edge. Therefore, the matrix M can be used to define two new histograms:

(a) A histogram based on the near-diagonal entries of M. This histogram should have a deep valley between the object and the background gray levels.

(b) A histogram based on the off-diagonal entries of M. This histogram should have a sharp peak between the object and the background gray levels.

A threshold for the image can then be chosen from the gray level range in which the valley in (a) overlaps with the peak in (b).

(ii) *(Gray Level, Local Average Gray Level) Scatter Plot Method*

This method [18] is quite similar to the co-occurrence matrix method discussed in the previous section. In a (gray level, local average gray level) scatter plot, the origin is defined as the upper left-hand corner, the gray level increases from left to right, and the local average gray level, taken over a window, increases from top to bottom. The intensity of a point on this plot is proportional to the occurrence frequency of the corresponding pair (gray level, local average gray level). In Kirby and Rosenfeld [18], the local averages are taken over a 3×3 square window.

The near-diagonal entries in the scatter plot represent pixels whose local average gray level is close to the gray level; such pixels are likely to be interior to the objects or background because of homogeneity. Thus, the histogram of the gray levels of these pixels should have a deep valley. On the other hand, the off-diagonal entries in the scatter plot represent pixels near an edge, so the gray level histogram obtained from these pixels should have a sharp peak. As in the co-occurrence matrix method, a threshold can be determined from the gray level range in which the valley and the peak of the above histogram overlap with each other.

C. *Deravi and Pal Method*

In this method [9], transition matrices similar to the co-occurrence matrix discussed in Section 3B(i) are used to define two "interaction measures" for threshold selection. The optimal threshold is determined by minimizing these interaction measures.

Using the notations of Section 3B(i), the transition matrices defined in [9] can be written as

$$T_h = M_{(1,0)},$$
$$T_v = M_{(1,3\pi/2)},$$
$$T_{vh} = T_v + T_h.$$

Any one of the above transition matrices can be used in threshold selection and we shall refer to the transition matrix as T in the following discussion. It was reported in [9] that the results obtained with different transition matrices do not differ significantly.

Let T_{ij} be the (i, j) entry of T and t be a threshold dividing the set of gray levels into two classes: $C_0 = \{0, 1, \ldots, t\}$ and $C_1 = \{t + 1, t + 2, \ldots, l - 1\}$. Then T can be partitioned into four regions defined by the following parameters:

$$a = \sum_{i=0}^{t} \sum_{j=0}^{t} T_{ij}, \qquad b = \sum_{i=t+1}^{l-1} \sum_{j=t+1}^{l-1} T_{ij},$$
$$c = \sum_{i=0}^{t} \sum_{j=t+1}^{l-1} T_{ij}, \qquad d = \sum_{i=t+1}^{l-1} \sum_{j=0}^{t} T_{ij}.$$

Thus a, b, c, and d represent the total number of transitions within C_0, within C_1, from C_0 to C_1, and from C_1 to C_0, respectively. The joint and conditional probabilities of the transition between C_0 and C_1 can be estimated by

$$P_j(t) = \frac{c + d}{a + b + c + d},$$

and

$$P_c(t) = \frac{1}{2}\left\{\frac{c}{a + c} + \frac{d}{b + d}\right\}.$$

Deravi and Pal [9] call $P_j(t)$ and $P_c(t)$ interaction measures. They also point out that $P_j(t)$ is similar to the "business measure" in [47] and that $P_c(t)$ does not depend directly on the histogram. The optimal threshold, t^*, is obtained by maximizing either one of the above interaction measures.

D. *Relaxation Methods*

The idea of relaxation was introduced by Southwell [38, 39] to improve the convergence of recursive solution for system of linear equations. In image segmentation, relaxation is applied as follows. The pixels of an image are first probabilistically classified into "light" and "dark" classes based on their gray levels. Then the probability of each pixel is adjusted according to the probabilities of the neighboring pixels. This adjustment process is iterated so that "light probabilities" (resp. "dark probabilities") become very high for pixels belonging to the light (resp. dark) regions. An attractive feature of these relaxation methods is that they are parallel-processing techniques, as opposed to the sequential techniques discussed so far.

(i) *Initial Classification*

For the initial classification of the pixels, Rosenfeld and Smith [36] have suggested the following method. Let d and l be the darkest and lightest gray levels, and g_i be the gray level of a pixel x_i. Then, for x_i, let

$$p^0_{i,\text{dark}} = \frac{l - g_i}{l - d},$$

and

$$p^0_{i,\text{light}} = \frac{g_i - d}{l - d}.$$

Though the above method is simple, it does not work in cases where the object and background gray levels do not lie on different halves of the gray level histogram. To avoid this problem, Rosenfeld and Smith [36] suggest another initialization scheme. Let m be the mean gray level. Then, if $g_i > m$, let

$$p^0_{i,\text{light}} = \frac{1}{2} + \frac{1}{2}\frac{g_i - m}{l - m},$$

and if $g_i \leq m$, let

$$p^0_{i,\text{dark}} = \frac{1}{2} + \frac{1}{2}\frac{m - g_i}{m - d}.$$

Fekete *et al.* [11] suggest an approach in which they assume the histogram can be divided into two Gaussian subpopulations so that the gray level distribution can be written as the sum of two Gaussian distributions. The parameters of these Gaussian distributions are determined by a method suggested in [7]. They find that faster convergence of the relaxation process can be obtained with this method.

(ii) *Iterative Updating of Probabilities*

As previously mentioned, the updating process consists of adjusting the probabilities of each pixel, based on neighboring probabilities. Let Λ be the set of class labels (e.g., the classes of dark and light pixels). Then, to direct the class updating, a compatibility coefficient, $r_{ij}(\lambda, \lambda')$, between a pixel x_i with label $\lambda \in \Lambda$ and another pixel x_j with label $\lambda' \in \Lambda$ is defined such that

$$r_{ij}(\lambda, \lambda') = \begin{cases} -1 & \text{if } \lambda \text{ and } \lambda' \text{ are } \textit{incompatible}. \\ 0 & \text{if } x_i \text{ and } x_j \text{ are } \textit{independent}. \\ 1 & \text{if } \lambda \text{ and } \lambda' \text{ are } \textit{compatible}. \end{cases}$$

Zucker *et al.* [52] propose the following equation for updating the probabilities:

$$p_i^{k+1}(\lambda) = \frac{p_i^k(\lambda)[1 + q_i^k(\lambda)]}{\sum_{\lambda \in \Lambda} p_i^k(\lambda)[1 + q_i^k(\lambda)]},$$

$$q_i^k(\lambda) = \frac{1}{8} \sum_{x_j \in N_i} \sum_{\lambda' \in \Lambda} r_{ij}(\lambda, \lambda') p_j^k(\lambda'),$$

where N_i is the 8-neighbor of x_i.

However, Pavlidis [29] has shown that $p_i^k(\lambda)$ always changes during the updating, therefore the above scheme violates the natural expectation that the labeling should not change if neighboring pixels are independent. Peleg [30] resolves this theoretical flaw by suggesting the following updating equation:

$$p_i^{k+1}(\lambda) = \frac{p_i^k(\lambda) \sum_{x_j \in N_i} \sum_{\lambda' \in \Lambda} r_{ij}(\lambda, \lambda') p_j^k(\lambda')}{\sum_{\lambda \in \Lambda} p_i^k(\lambda) \sum_{x_j \in N_j} \sum_{\lambda' \in \Lambda} r_{ij}(\lambda, \lambda') p_j^k(\lambda')}.$$

E. *Gradient Relaxation Methods*

In gradient relaxation, the optimal labeling scheme is determined by maximizing a criterion function with gradient optimization.

Let λ_1 and λ_2 be the labels of the classes of light and dark pixels, let $\{[p_i(\lambda_1), p_i(\lambda_2)]^T, i = 0, 1, \ldots, l-1\}$ be the set of probability vectors associated with the gray levels and let $\{[q_i(\lambda_1), q_i(\lambda_2)]^T, i = 0, 1, \ldots, l-1\}$ be the set of compatibility vectors, where $q_i(\lambda_k) = \frac{1}{8}\Sigma_{x_j \in N_i} p_j(\lambda_k)$, N_i is the 8-neighbor of the pixel x_i. Bhanu and Feugeras [4] suggest that the optimal labeling scheme for images with unimodal gray level distribution can be determined by

$$\text{maximizing} \qquad C(p) := \sum_{i=0}^{l-1} p_i^T(\lambda) q_i(\lambda) \qquad \text{subject to}$$

$$p_i(\lambda) \in K := \left\{ p(\lambda) = (p(\lambda_1), p(\lambda_2)) \mid p(\lambda_j) \geq 0, \sum_{j=1}^{2} p(\lambda_j) = 1 \right\}.$$

They point out that maximizing $C(p)$ is equivalent to reducing the inconsistency and ambiguity of the pixel labeling.

Here we suggest a similar method that uses a different criterion function derived from information theory. In this method the optimal labeling scheme is determined by

$$\text{maximizing} \qquad \Psi(p) := \sum_{i=0}^{l-1} I(p_i, q_i)$$

$$\text{subject to} \qquad p_i \in K.$$

The quantity $I(p_i, q_i)$ is defined as

$$I(p_i, q_i) = \sum_{j=1}^{2} p_i(\lambda_j)\log_e\left(\frac{p_i(\lambda_j)}{q_i(\lambda_j)}\right).$$

F. *Other Methods*

Thresholding techniques for optical character recognition systems have also received much attention. Because of the wide range of print quality distortions over a single document, a combination of threshold operators is often used, with each operator sensitive to a different type of distortion. Bartz [3], for example, combines four linear threshold operators to form a single threshold. An example of these operators is $T = kv + c$, where v is the average contrast over previously scanned characters, and k and c are optimizing parameters. Wolfe [50] suggests a two-step procedure to deal with the problem of shading in printed characters. During the first step, the mean gray level of each pixel is computed by averaging the gray levels over a 4×4 window. A pixel X is considered as belonging to a character if its mean gray level is darker than the mean gray levels of the two pixels whose orientations differ by 180° and are distance 8 apart. The second step is similar to the first step, but a larger window is used.

In Ullmann [41], the threshold of a pixel X is selected based on the gray levels in a 5×5 window, W, of X. Only the pixels which are labeled n contribute to the threshold decision (Fig. 1).

Two experimental rules, to be described later, are used to determine the threshold of X. Let n_X be the highest (brightest) gray level in W. Then rule (1) is applied if $n_X \leq 40$ and rule (2) is applied when $n_X > 40$. The two rules are:

(1) Label X as an object point if for some point n in W, we have $g(X) - g(n) < \tau$, where $g(\cdot)$ means the gray level of $(\cdot)$ and τ is some predefined threshold; otherwise label it as a background point.

(2) Label X as an object point if for at least one n in W, we have $g(X) < g(n)/\mu$, where μ is some predefined constant; otherwise label it as a background point.

Morrin [25] uses gray level versus gradient value plots to convert gray scale images with superior resolution and contrast for thresholding. Panda [28] suggests a method in which the threshold applied at a point depends on both the gray level and the edge value of that point. Some results and comparisons are given in [28].

o	n	n	n	o
n	n	o	n	n
n	o	X	o	n
n	n	o	n	n
o	n	n	n	o

FIG. 1. 5×5 window used in Ullmann [41].

4. LOCAL THRESHOLDING

In local thresholding, the original image is partitioned into smaller subimages and a threshold is determined for each of the subimages. This yields a thresholded image with gray level discontinuities at the boundaries of two different subimages. The threshold of a region can be determined by either the point-dependent method or the region-dependent method. A smoothing technique is then applied to eliminate the discontinuities.

Chow and Kaneko [7, 8] suggest the use of a 7×7 window for local thresholding. In their method, the original image is divided into 7×7 subimages and a threshold is computed for each subimage. However, a threshold is not computed for subimages with unimodal gray level histogram. Thresholds for such subimages are interpolated from neighboring subimages. For a bimodal subimage, the threshold is computed as follows. First the gray level histogram for a subimage is approximated by a sum of two Gaussian distributions, then the threshold is obtained by minimizing the classification error with respect to the threshold value. Some experiments on this method were done in [26].

For X-ray angiograms Fernando and Monro [12] suggest a local thresholding method. Global thresholding techniques have been found to be unsatisfactory for these images which are usually unimodal with a very narrow peak. According to this new method, the original image is partitioned with 16 nonoverlapping subimages and the entropic thresholding technique of Pun [32] is applied to determine the threshold value for each of these subimages. Finally, the entire thresholded image is processed by a low-pass filter to eliminate the gray level discontinuities at the boundaries of subimages.

5. MULTITHRESHOLDING METHODS

Many global thresholding methods, such as Ostu [27], Pun [32, 33], Kapur *et al.* [16], moment preserving [40], minimum error [19] can be extended to the case of multithresholding. In this section we discuss three multithresholding methods which are not included in the earlier sections of this paper.

A. *Amplitude Segmentation Method*

This method was proposed by Boukharouba *et al.* [5] and uses the intrinsic properties of cumulative distribution function of an image to be thresholded. In this method, the curvature of the distribution function is examined to obtain information regarding the threshold values. The distribution function $F(k)$ at point k is given by

$$F(k) = \frac{\sum_{g=0}^{k} h(g)}{\sum_{g=0}^{l-1} h(g)}.$$

The curvature of F is then defined by

$$C(x) = F''(x)\left[1 + (F'(x))^2\right]^{-3/2},$$

where F' and F'' are the first and second derivatives of F, respectively. It is pointed out in [5] that $C(x)$ is noisy and oscillatory and should be smoothed and approximated for use in thresholding. The zeros of the curvature determines the thresholds as well as the gray level to be assigned to each class.

B. *Wang and Haralick Method*

This is a recursive technique [42] for multiple threshold selection on digital images. In this method, pixels are first classified as edge pixels or nonedge pixels. Edge pixels are then classified, on the basis of their neighborhoods, as being relatively dark or relatively light. A histogram of the gray levels is obtained for those pixels which are edge pixels and relatively dark, and another histogram is obtained for those pixels which are edge pixels and relatively light. A threshold is selected based on the gray level intensity value corresponding to one of the highest peaks from the two histograms. To get multiple thresholds, the procedure is recursively applied using only those pixels whose intensities are smaller than threshold first and then using only those pixels whose intensities are larger than the threshold.

C. *Uniform Contrast Method*

This is a recursive threshold selection method proposed by Kohler [21]. The method is based on the following idea. The optimum threshold for segmentation of the image is that threshold which detects more high contrast edges and fewer low contrast edges than any other threshold [21]. In this method, a histogram of the average contrast $\mu(t)$ for each possible threshold t is created and the highest peak in the histogram corresponds to the optimal threshold. The average contrast $\mu(t)$ is calculated from the relation

$$\mu(t) = \frac{C(t)}{N(t)}$$

with $\mu(t) = 0$ if $N(t) = 0$ and $C(t)$ is the total contrast detected by threshold t, and $N(t)$ is the number of edges detected by t. For multithresholding any initial threshold is first selected and then a new histogram of $\mu(t)$ is computed by removing the contribution of the already detected edges by the initial threshold. This procedure is continued until the maximum average contrast for any threshold fell below some minimum average contrast criterion $\theta > 1$.

6. EXPERIMENTAL RESULTS

Not all the methods surveyed in this paper are automatic in nature (i.e., they require no human interactions). Some require user's feedback since unique optimal threshold is not possible. In local thresholding methods, the image is divided into smaller subimages and suitable global methods are adopted to compute the optimal thresholds for each subimage. It seems that global methods are frequently used for thresholding. For this reason, we intend to evaluate some of the global thresholding methods, namely the Ostu method [27], the Pun method [32], the Johannsen and Bille Method [15], the method of Kapur *et al.* [16], the co-occurrence matrix method [2], the histogram concavity analysis method [35], the method due to Deravi and Pal [9], the moments preserving method [40], and the minimum error method

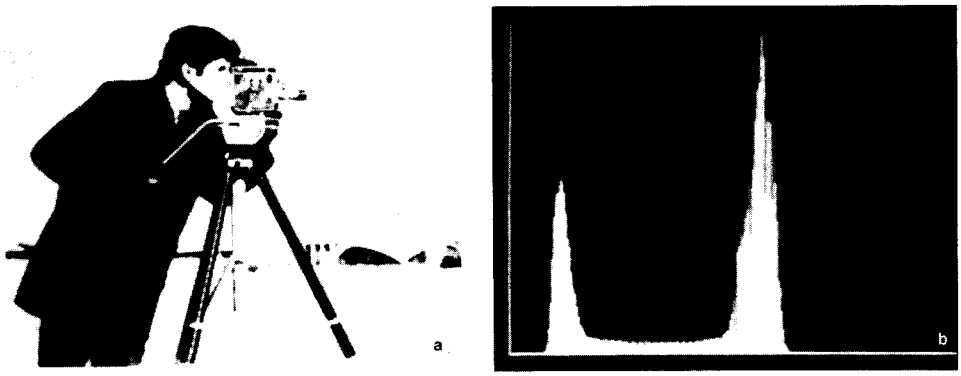

FIG. 2. (a) Digitized cameraman image; (b) histogram of cameraman image.

[19]. These methods are considered for a comparative study since they are automatic and in most cases do not require a subjective judgement for finding the best threshold from a set of optimal threshold values. The above mentioned methods are applied to three images, "cameraman", the "building", and the "model". The cameraman image is digitized from the printed version of the same image in [32]. This is done with a raster of 415 × 395 and the quantization is 256 gray levels. The digitized image and its gray level histogram are shown in Fig. 2. The building image is also digitized from the printed version of the same image in [32] with a raster of 411 × 403 and quantized to 256 gray levels. The image of the model is digitized with a raster of 321 × 314 and quantized to 256 levels. The building and the image of the model are shown in Figs. 3 and 4 along with their gray level histograms. The threshold values obtained by methods considered above are presented in Table 1.

For the co-occurrence matrix method the distance between cells d was chosen to be 1 and 5% of the elements of the M matrix that lie farthest from the diagonal was projected onto the diagonal. Incidentally a threshold value of 63 was found for all three images. This value did not change when 10% of the elements instead of 5% was used. However, when d changed from 1 to 2 (resp. 3) the threshold values for the

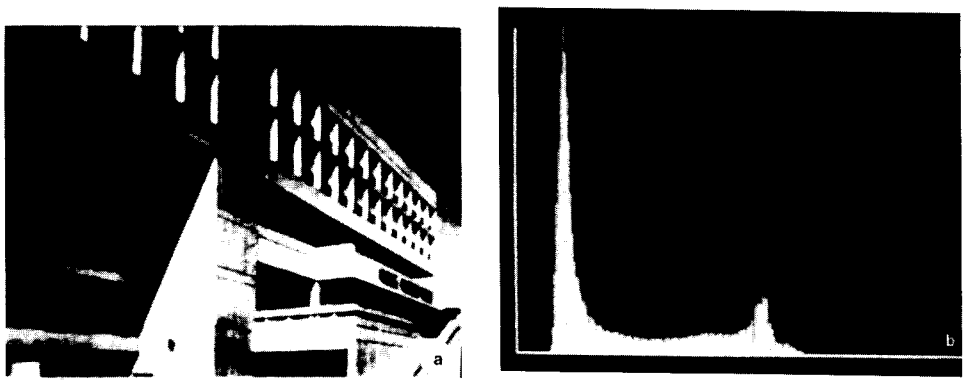

FIG. 3. (a) Digitized building image; (b) histogram of building image.

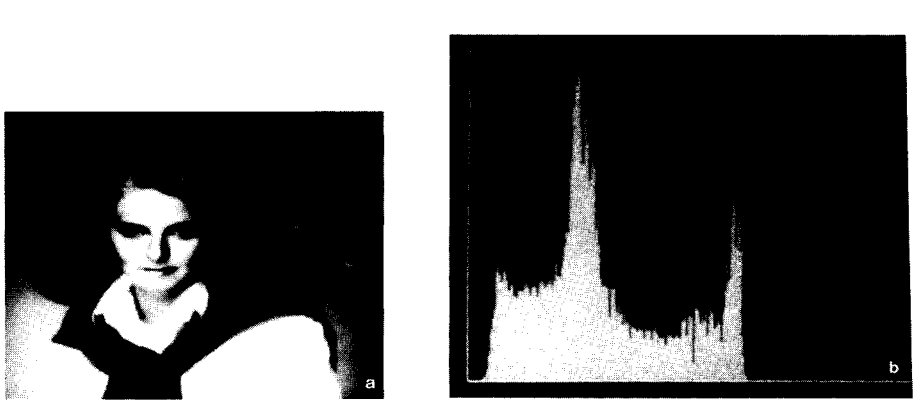

FIG. 4. (a) Digitized image of the model; (b) histogram of the model image.

cameraman and the building images did not change, but the threshold value of the model image changed to 30 (resp. 33). On changing the percentage of elements of M projected onto the diagonal to 1% and 2% while keeping d values to 2 and 3, respectively, we have obtained a threshold value 63 in both cases. Thus, whenever higher values of d are used, it is advisable to decrease the percentage of elements of M that lie farthest from the diagonal and projected onto the diagonal.

The Deravi and Pal method [9] was implemented for the cameraman picture only. With their method, two local minima were found, one at 45 and the other at 116. Since thresholding at 45 does not provide a better binary image we select 116 as the optimal threshold value. Figures 5–7 show the binary images obtained by the thresholding methods considered in this section.

7. MEASURES FOR EVALUATING THRESHOLDING METHODS

In digital images, uniformity and shape of the objects play great roles in separating objects from the background. The amount of agreement of these two aspects of every binary image with the real image has been evaluated for the three test images considered in the previous section. The uniformity measure used in this

TABLE 1
Optimal Threshold Value for the Set of Test Images

Method	Optimal threshold value		
	Cameraman	Building	Model
Co-occurrence matrix	63	63	63
Deravi & Pal	116		
Histogram concavity	127	47	127
Johannsen & Bille	75	79	114
Kapur *et al.*	123	69	81
Minimum error	111	37	113
Moment-preserving	91	76	76
Ostu	86	73	91
Pun [32]	131	28	63

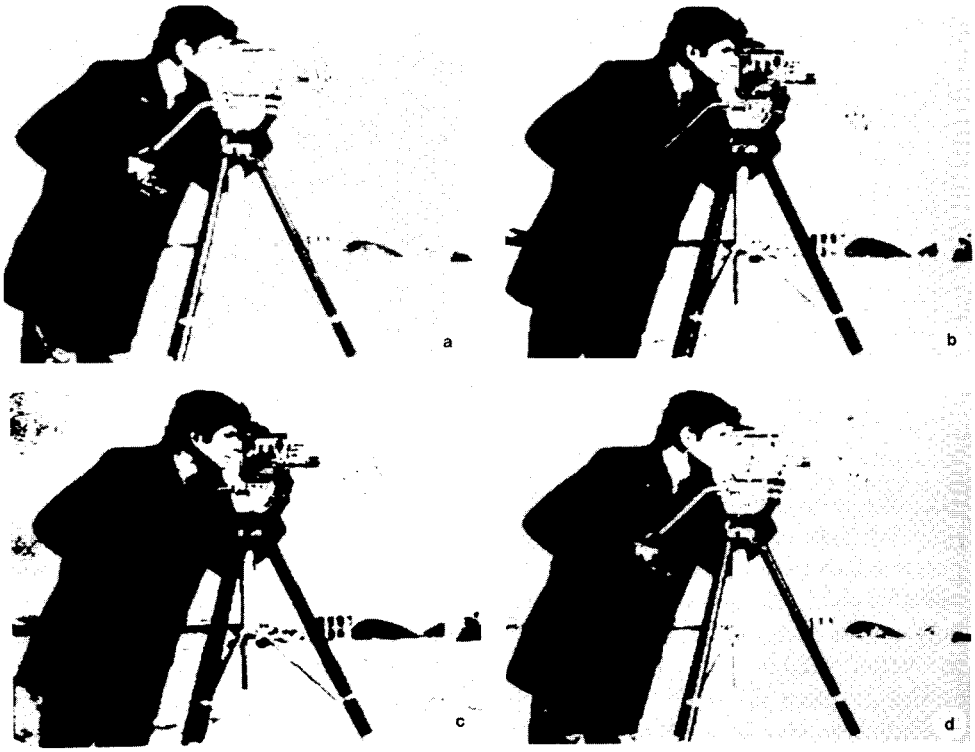

FIG. 5. Binary images of the cameraman: (a) co-occurrence matrix method ($t^* = 63$); (b) Deravi and Pal method ($t^* = 116$); (c) histogram concavity method ($t^* = 127$); (d) Johannsen and Bille method ($t^* = 75$); (e) Kapur, Sahoo, and Wong method ($t^* = 123$); (f) minimum error method ($t^* = 111$); (g) moment-preserving method ($t^* = 91$); (h) Ostu method ($t^* = 86$); (i) Pun method [32] ($t^* = 131$).

paper is adopted from Levine and Nazif [22]. For a given threshold value t, the *uniformity measure* $U(t)$ is given by

$$U(t) = 1 - \frac{\sigma_1^2 + \sigma_2^2}{C},$$

where

$$\sigma_i^2 = \sum_{(x,y)\in R_i} (f(x,y) - \mu_i)^2,$$

R_i = Segmented region i

$f(x,y)$ = The gray level of the pixel (x,y)

$$\mu_i = \frac{\sum_{(x,y)\in R_i} f(x,y)}{A_i},$$

A_i = Number of pixels in R_i, $\quad i = 1, 2,$

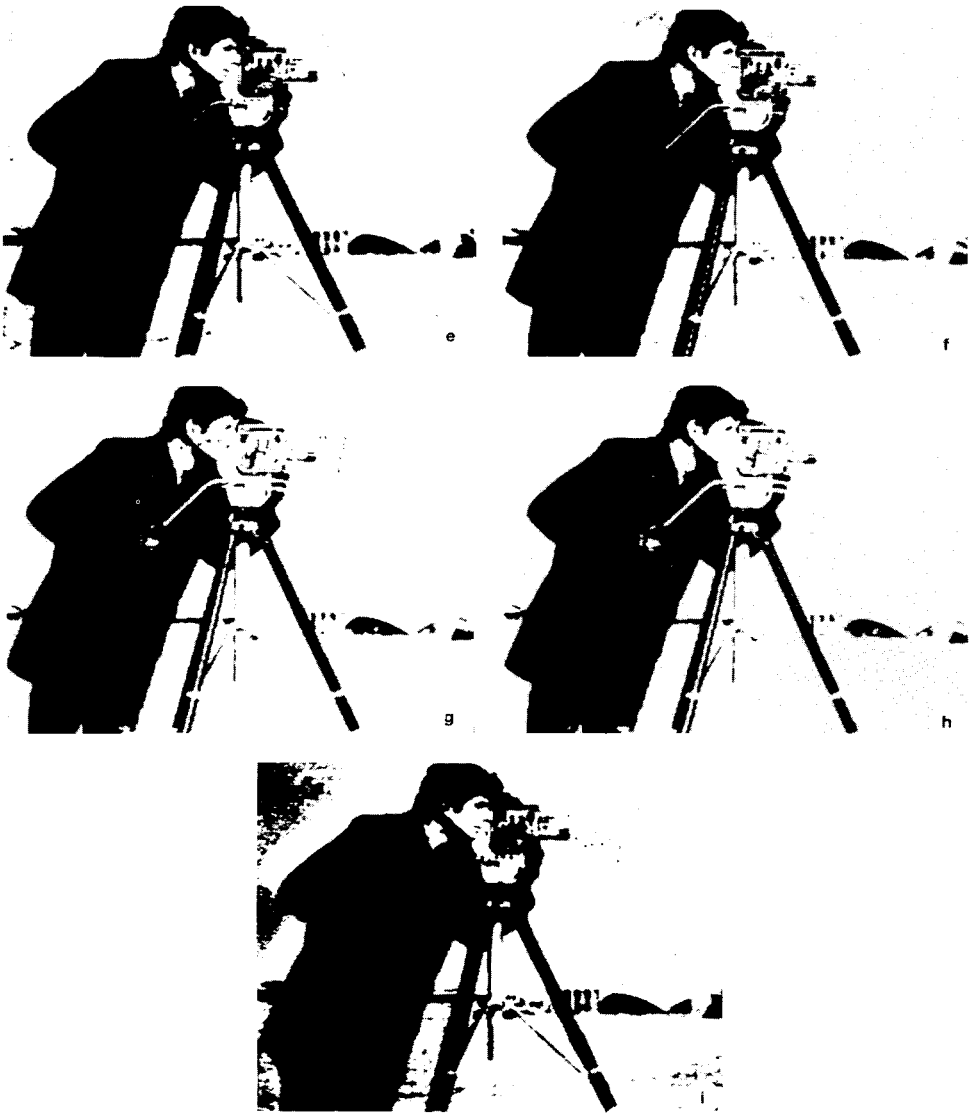

FIG. 5.—*Continued*

and

$$C = \text{A normalization factor.}$$

Levine and Nazif [22] have included a weighting factor in the calculation of uniformity measure U for images with more than two regions.

A measure, called *shape measure* S, is used for the measurement of the shape of the object in the test images. The measure S for a given image is calculated as follows: (a) assign a generalized gradient value $\Delta(x, y)$ to every pixel (x, y); (b) if the pixel (x, y) has a gray value higher than the average of its neighbors then assign the "+" sign to the generalized gradient value $\Delta(x, y)$, else assign the "−" value;

FIG. 6. Binary images of the building: (a) co-occurrence matrix method ($t^* = 63$); (b) histogram concavity method ($t^* = 47$); (c) Johannsen and Bille method ($t^* = 79$); (d) Kapur, Sahoo, and Wong method ($t^* = 69$); (e) minimum error method ($t^* = 37$); (f) moment-preserving method ($t^* = 76$); (g) Ostu method ($t^* = 73$); (h) Pun method [32] ($t^* = 28$).

(c) compute the shape measure S by using the formula

$$S = \frac{\sum_{(x,y)} \mathrm{Sgn}\left(f(x,y) - \overline{f_{N(x,y)}}\right)\Delta(x,y)\mathrm{Sgn}(f(x,y) - t)}{C},$$

where $\overline{f_{N(x,y)}}$ is the average gray value in the neighborhood $N(x, y)$, t is the threshold value of the image, C is a normalization factor, and

$$\mathrm{Sgn}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0. \end{cases}$$

The computation of the generalized gradient value $\Delta(x, y)$ of the pixel (x, y) is carried out using the formula

$$\Delta(x,y) = \left[\sum_{k=1}^{4} D_k^2 + \sqrt{2}\,D_1(D_3 + D_4) - \sqrt{2}\,D_2(D_3 - D_4)\right]^{1/2},$$

where

$$\begin{aligned} D_1 &= f(x+1, y) - f(x-1, y), \\ D_2 &= f(x, y-1) - f(x, y+1), \\ D_3 &= f(x+1, y+1) - f(x-1, y-1), \\ D_4 &= f(x+1, y-1) - f(x-1, y+1). \end{aligned}$$

Using both measures, the threshold value obtained according to each method mentioned earlier are evaluated. Tables 2a, 2b and 2c show the results of this

FIG. 6.—*Continued*

evaluation. Note that these values are renormalized according to the best possible threshold for each measure and the rank of each method is indicated within the square bracket.

8. DISCUSSION OF EVALUATION RESULTS

Analyzing Table 2a, we find that methods of Ostu, moment-preserving, and Johannsen and Bille are better threshold selection methods for the cameraman image with regards to region uniformity and shape measure. For bimodal images (see Table 3), the optimal values of uniformity and shape measure do not differ

FIG. 7. Binary images of the model: (a) co-occurrence matrix method ($t^* = 63$); (b) histogram concavity method ($t^* = 127$); (c) Johannsen and Bille method ($t^* = 114$); (d) Kapur, Sahoo, and Wong method ($t^* = 81$); (e) minimum error method ($t^* = 113$); (f) moment-preserving method ($t^* = 76$); (g) Ostu method ($t^* = 91$).

significantly. Thus for the cameraman image, if a particular method is found to be good it will also be found equally good with respect to the shape measure. Therefore, the above mentioned methods in Table 2a have the same rank with respect to uniformity as well as shape. On the whole, the evaluation results on this image for the above mentioned methods vary very little.

The building image does not have a distinct bimodal gray level histogram. Thus, unlike the cameraman image, the Ostu method (see Table 2b) ranks third according to shape measure and first with respect to region uniformity measure when applied to the building image. Moment-preserving method ranks second according to uniformity and ranks first according to shape measure. The performances of Johannsen and Bille and Kapur *et al.* methods are next to the Ostu and moment-preserving method.

Notice that the image of the model does not have a bimodal histogram. For this image, we find (refer to Table 2c) that Ostu method ranks first with respect to uniformity and ranks fourth with respect to the shape measure. The co-occurrence matrix method is found to have first rank according to the shape measure and seventh rank with respect to the uniformity measure. The method due to Kapur *et al.* ranks second with respect to uniformity and third according to shape. This result shows that the Kapur *et al.* method is fairly consistent with both shape and uniformity measure for this type of images. The co-occurrence matrix method is shape oriented and the Ostu method is uniformity oriented. Thus for the image of the model, which does not have a bimodal histogram, the co-occurrence matrix

FIG. 7.—*Continued*

method is found to be a good method if only the shape measure is considered, whereas the Ostu method is good on the basis of uniformity measure alone.

From this objective evaluation (from Tables 2a, 2b, and 2c), we find the Ostu method, which is based on discriminant analysis to be one of the better thresholding methods despite of its many short commings (see [20, 34]). Since our measures for evaluating a thresholding method are based on uniformity and the shape only, it is not altogether surprising to find Ostu method as one of the better methods. The moment-preserving method of Tsai is found to be comparable to the Ostu method. The performance of methods such as Johannsen and Bille as well as Kapur *et al.* is almost next to Ostu and the moment-preserving methods.

The thresholded images obtained by various methods (listed in Table 1) reveal valuable information regarding the thresholding techniques. We incorporate this

TABLE 2a
Evaluation Results for Cameraman Image

Method	Threshold	Uniformity (U)	Shape (S)
Co-occurrence matrix	63	0.7121 [5]	0.7471 [4]
Deravi & Pal	116	0.6190 [6]	0.5740 [6]
Histogram concavity	127	0.3556 [8]	0.3387 [8]
Johannsen & Bille	75	0.9089 [3]	0.9452 [3]
Kapur *et al.*	123	0.4775 [7]	0.4148 [7]
Minimum error	111	0.7133 [4]	0.6951 [5]
Moment-preserving	91	0.9911 [2]	0.9839 [2]
Ostu	86	0.9990 [1]	1.0000 [1]
Pun [32]	131	0.2205 [9]	0.2390 [9]

TABLE 2b
Evaluation Results for Building Image

Method	Threshold	Uniformity (U)	Shape (S)
Co-occurrence matrix	63	0.9098 [5]	0.8863 [5]
Histogram concavity	47	0.5889 [6]	0.5597 [6]
Johannsen & Bille	79	0.9833 [3]	0.9990 [2]
Kapur *et al.*	69	0.9792 [4]	0.9647 [4]
Minimum error	37	0.3820 [7]	0.2987 [7]
Moment-preserving	76	0.9980 [2]	0.9995 [1]
Ostu	73	0.9986 [1]	0.9931 [3]
Pun [32]	28	0.1988 [8]	0.0131 [8]

TABLE 2c
Evaluation Results for Image of Model

Method	Threshold	Uniformity (U)	Shape (S)
Co-occurrence matrix	63	0.5483 [7]	0.9597 [1]
Histogram concavity	127	0.6246 [6]	0.2775 [7]
Johannsen & Bille	114	0.8334 [5]	0.4543 [6]
Kapur *et al.*	81	0.9192 [2]	0.7636 [3]
Minimum error	113	0.8465 [3]	0.4682 [5]
Moment-preserving	76	0.8445 [4]	0.7872 [2]
Ostu	91	0.9986 [1]	0.7418 [4]
Pun [32]	63	0.5483 [7]	0.9597 [1]

visual information to supplement our findings. For visual analysis of the binary images, it seems reasonable to consider some important features such as facial details and camera of the cameraman picture, edges of the building on the right-hand side of the building image, and facial and hair features in the model image. Along with these observations, one should also consider the amount of distortion and loss of information in the thresholded image. A close look at the binary images of the cameraman in Fig. 5, we observe that methods such as Pun (Fig. 5i), histogram concavity (Fig. 5c) and co-occurrence matrix (Fig. 5a) do not provide a good threshold for this image. This can also be observed from Table 2a. We also observe that the co-occurrence matrix method yields a binary image in which the details of the face and the camera of the cameraman are lost, whereas Pun's and histogram concavity methods yield binary images with distortions. Also, Ostu method and moment-preserving method do not retain the details of the face in the cameraman image. The methods which retain the facial details and other valuable information of the image are the Kapur *et al.* method and the minimum error method. Notice that the Kapur *et al.* method ranks seven (in Table 2a) according to uniformity as well as shape measure, but it gives a good binary image, which retains the facial details of the cameraman.

A similar examination of the images in Fig. 6 reveals that the moment-preserving (Fig. 6f), co-occurrence matrix (Fig. 6a), Ostu (Fig. 6h), Kapur *et al.* (Fig. 6d), Johannsen and Bille (Fig. 6c) methods provide reasonably good threshold values for

TABLE 3
Optimal Value for the Uniformity and Shape Measures

Measure	Cameraman	Building	Model
Uniformity	87	74	93
Shape	86	78	57

FIG. 8. Binary images of the cameraman: (a) optimal uniformity ($t^* = 87$); (b) optimal shape ($t^* = 86$).

the building picture. Among these methods, the Johannsen and Bille method and the moment-preserving method yield good binary images. Regarding the binary images in Fig. 7, we observe that the Kapur *et al.* method and the moment-preserving method are better threshold selection techniques. For the same image, the Ostu method also provides a good threshold value. However, it is not as good as the two methods mentioned.

Thus, from the visual aspect of the binary images as well as the objective evaluation based on the uniformity and shape measures, we find that Johannsen and Bille, Kapur *et al.*, moment-preserving, and Ostu methods are good thresholding methods. In most cases, Pun [32], histogram concavity, and minimum error methods do not render good performance in comparison to other methods mentioned in this section. In an earlier study [12], unsatisfactory performance of the Pun method [32] was observed by Fernando and Monro.

9. COMMENTS AND CONCLUSION

The indicators used in this paper for evaluating an automatic thresholding method are *shape measure* and *uniformity measure*. Table 3 presents the gray levels which optimize the uniformity and shape measure for the test images. The set of digital test images, when binarized at the gray levels where S and U attain maximum, are shown in Figs. 8, 9, and 10. From these binary images it is evident that one may use these measures for image thresholding.

In summary, we have presented a survey of various thresholding methods and evaluated the performance of some global methods which are automatic in nature. All the methods described here optimize some criterion functions and provide

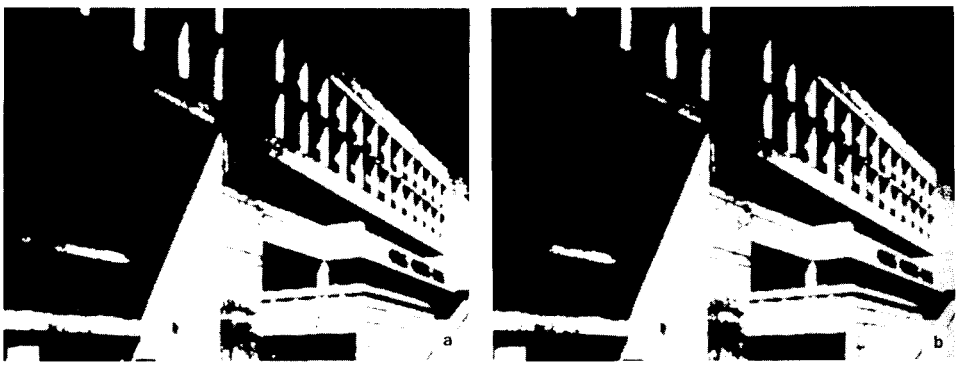

FIG. 9. Binary images of the building: (a) optimal uniformity ($t^* = 74$); (b) optimal shape ($t^* = 78$).

FIG. 10. Binary images of the model: (a) optimal uniformity ($t^* = 93$); (b) optimal shape ($t^* = 57$).

justifications for such optimization. We only investigate how the optimal values behave corresponding to the uniformity and shape measures to a given set of test images. From this study, we conclude that for the above set of test images Johannsen and Bille, Kapur *et al.*, Tsai's moment-preserving, and Ostu methods are reasonably good thresholding methods if one demands more uniformity and better shape of the object in the binary image. However, whether the aforementioned methods, in general, will perform well or not remains open.

ACKNOWLEDGMENTS

We are thankful to the referee for helping to improve the presentation of this paper.

REFERENCES

1. N. Abramson, *Information Theory and Coding*, McGraw–Hill, New York, 1963.
2. N. Ahuja and A. Rosenfeld, A note on the use of second-order gray-level statistics for threshold selection, *IEEE Trans. Systems Man Cybernet.* **SMC-8**, 1978, 895–899.
3. M. R. Bartz, Optimizing a video processor for OCR, in *Proceedings, International Joint Conference on AI*, 1969, pp. 79–90.
4. B. Bhanu and O. Faugeras, Segmentation of images having unimodal distributions, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-4**, 1982, 408–419.

5. S. Boukharouba, J. M. Rebordao, and P. L. Wendel, An amplitude segmentation method based on the distribution function of an image, *Comput. Vision Graphics Image Process.* **29**, 1985, 47–59.
6. H. Bunke, H. Feistl, H. Neimann, G. Sagerer, F. Wolf, and G. Zhou, Smoothing, thresholding, and contour extraction in images from gated blood pool studies, in *First IEEE Symp. on Medical Imaging and Image Interpretation, Berlin 1982.*
7. C. K. Chow and T. Kaneko, Boundary detection of radiographic images by a threshold method, in *Proceedings, IFIP Congress 71*, pp. 130–134.
8. C. K. Chow and T. Kaneko, Automatic boundary detection of left ventricle from cineangiograms, *Comput. Biomed. Res.* **5**, 1972, 338–410.
9. F. Deravi and S. K. Pal, Gray level thresholding using second-order statistics, *Pattern Recognit. Lett.* **1**, 1983, 417–422.
10. W. Doyle, Operation useful for similarity-invariant pattern recognition, *J. Assoc. Comput. Mach.* **9**, 1962, 259–267.
11. G. Fekete, J. O. Eklundh, and A. Rosenfeld, Relaxation: Evaluation and applications, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-3**, 1981, 460–469.
12. S. M. X. Fernando and D. M. Monro, Variable thresholding applied to angiography, in *Proceedings, 6th International Conference on Pattern Recognition, 1982.*
13. S. K. Fu and J. K. Mu, A survey on image segmentation, *Pattern Recognit.* **13**, 1981, 3–16.
14. R. M. Haralick, K. Shanmugam, and I. Dinstein, Texture features for image classification, *IEEE Trans. Systems Man Cybernet.* **SMC-3**, 1973, 610–621.
15. G. Johannsen and J. Bille, A threshold selection method using information measures, in *Proceedings, 6th Int. Conf. Pattern Recognition, Munich, Germany, 1982*, pp. 140–143.
16. J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, A new method for gray-level picture thresholding using the entropy of the histogram, *Comput. Vision Graphics Image Process.* **29**, 1985, 273–285.
17. Y. H. Katz, Pattern recognition of meteorological satellite cloud photography, *Proceedings, Third Symp. on Remote Sensing of Environment, 1965*, pp 173–214.
18. R. L. Kirby and A. Rosenfeld, A note on the use of (gray level, local average gray level) space as an aid in thresholding selection, *IEEE Trans. Systems Man Cybernet.* **SMC-9**, 1979, 860–864.
19. J. Kittler and J. Illingworth, Minimum error thresholding, *Pattern Recognit.* **19**, 1986, 41–47.
20. J. Kittler and J. Illingworth, On threshold selection using clustering criterion, *IEEE Trans. Systems Man Cybernet.* **SMC-15**, 1985, 652–655.
21. R. Kohler, A segmentation system based on thresholding, *Comput. Graphics Image Process.* **15**, 1981, 319–338.
22. M. D. Levine and A. M. Nazif, Dynamic measurement of computer generated image segmentations, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-7**, 1985, 155–164.
23. D. Mason, I. J. Lauder, D. Rutoritz, and G. Spowart, Measurement of C-Bands in human chromosomes, *Comput. Biol. Med.* **5**, 1975, 179–201.
24. D. L. Milgram, Region extraction using convergent evidence, *Comput. Graphics Image Process.* **11**, 1979, 1–12.
25. T. H. Morrin, A black–white representation of a gray-scale picture, *IEEE Trans. Comput.* **23**, 1974, 184–186.
26. Y. Nakagawa and A. Rosenfeld, Some experiments on variable thresholding, *Pattern Recognit.* **11**, 1979, 191–204.
27. N. Ostu, A threshold selection method from gray-level histogram, *IEEE Trans. Systems Man Cybernet.* **SMC-8**, 1978, 62–66.
28. D. P. Panda, *Segmentation of FLIR Images by Pixel Classification,* TR-508, University of Maryland Computer Science Center, 1977.
29. T. Pavlidis, *Structural Pattern Recognition*, Springer-Verlag, New York, 1977.
30. S. Peleg, A new probabilistic relaxation scheme, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-2**, 1980, 362–369.
31. J. M. S. Prewitt and M. L. Mendelsohn, The analysis of cell images, in *Ann. New York Acad. Sci.* Vol 128, pp 1035–1053, New York Acad. Sci., New York, 1966.
32. T. Pun, A new method for gray-level picture thresholding using the entropy of the histogram, *Signal Process.* **2**, 1980, 223–237.
33. T. Pun, Entropic thresholding: A new approach, *Comput. Vision Graphics Image Process.* **16**, 1981, 210–239.
34. S. S. Reddi, S. F. Rudin, and H. R. Keshavan, An optimal multiple threshold scheme for image segmentation, *IEEE Trans. Systems Man Cybernet.* **SMC-14**, 1984, 661–665.

35. A. Rosenfeld and P. De La Torre, Histogram concavity analysis as an aid in threshold selection, *IEEE Trans. Systems Man Cybernet.* **SMC-13**, 1983, 231–235.
36. A. Rosenfeld and R. C. Smith, Thresholding using relaxation, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-3**, 1981, 598–606.
37. P. K. Sahoo, *Theory and Applications of Some Measure of Uncertainties*, Ph.D. thesis, Department of Applied Mathematics, University of Waterloo, Waterloo, 1986.
38. R. Southwell, *Relaxation Methods in Engineering Science, A Treatise on Approximate Computation*, Oxford Univ. Press, London, 1940.
39. R. Southwell, *Relaxation Methods in Theoretical Physics*, Oxford Univ. Press (Clarendon), London, 1946.
40. W. Tsai, Moment-preserving thresholding: A new approach, *Comput. Vision Graphics Image Process.* **29**, 1985, 377–393.
41. J. R. Ullman, Binarization using associative addressing, *Pattern Recognit.* **6**, 1974, 127–135.
42. S. Wang and R. M. Haralick, Automatic multithreshold selection, *Comput. Vision Graphics Image Process.* **25**, 1984, 46–67.
43. S. Watanabe, and the CYBEST group, An automated apparatus for cancer processing CYBEST, *Comput. Vision Graphics Image Process.* **3**, 1974, 350–358.
44. J. S. Weszka, A survey of threshold selection techniques, *Comput. Vision Graphics Image Process.* **7**, 1978, 259–265.
45. J. S. Weszka, R. N. Nagel, and A. Rosenfeld, A threshold selection technique, *IEEE Trans. Comput.* **C-23**, 1974, 1322–1326.
46. J. S. Weszka and A. Rosenfeld, *Threshold Selection 4*, TR-336, University of Maryland Computer Science Center, 1974.
47. J. S. Weszka and A. Rosenfeld, Threshold evaluation techniques, *IEEE Trans. Systems Man Cybernet.* **SMC-8**, 1978, 622–629.
48. J. S. Weszka and A. Rosenfeld, Histogram modification for threshold selection, *IEEE Trans. Systems Man Cybernet.* **SMC-9**, 1979, 38–51.
49. J. S. Weszka, J. A. Veretuon, and A. Rosenfeld, *A Technique for Facilitating Threshold Selection for Objects Extraction from Digital Pictures*, TR-243, University of Maryland Computer Science Center, 1973.
50. R. N. Wolfe, A dynamic thresholding scheme for quantization of scanned image, in *Proceedings, Automatic Pattern Recognition*, 1969, pp. 143–162.
51. A. Y. Wu and A. Rosenfeld, Threshold selection using quadtree, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-4**, 1982, 90–94.
52. S. Zucker, R. Hummel, and A. Rosenfeld, An application of relaxation labelling to line and curve enhancement, *IEEE Trans. Comput.* **C-26**, 1977, 394–403.