Journal of Electronic Imaging 13(1), 146–165 (January 2004).

Survey over image thresholding techniques and quantitative performance evaluation

Mehmet Sezgin
Tübitak Marmara Research Center
Information Technologies Research Institute
Gebze, Kocaeli
Turkey
E-mail: sezgin@btae.mam.gov.tr

Bülent Sankur
Boğaziçi University
Electric-Electronic Engineering Department
Bebek, Istanbul
Turkey

Abstract. *We conduct an exhaustive survey of image thresholding methods, categorize them, express their formulas under a uniform notation, and finally carry their performance comparison. The thresholding methods are categorized according to the information they are exploiting, such as histogram shape, measurement space clustering, entropy, object attributes, spatial correlation, and local gray-level surface. 40 selected thresholding methods from various categories are compared in the context of nondestructive testing applications as well as for document images. The comparison is based on the combined performance measures. We identify the thresholding algorithms that perform uniformly better over nondestructive testing and document image applications.* © *2004 SPIE and IS&T.* [DOI: 10.1117/1.1631316]

1 Introduction

In many applications of image processing, the gray levels of pixels belonging to the object are substantially different from the gray levels of the pixels belonging to the background. Thresholding then becomes a simple but effective tool to separate objects from the background. Examples of thresholding applications are document image analysis, where the goal is to extract printed characters,[1,2] logos, graphical content, or musical scores: map processing, where lines, legends, and characters are to be found:[3] scene processing, where a target is to be detected:[4] and quality inspection of materials,[5,6] where defective parts must be delineated. Other applications can be listed as follows: cell images[7,8] and knowledge representation,[9] segmentation of various image modalities for nondestructive testing (NDT) applications, such as ultrasonic images in Ref. 10, eddy current images,[11] thermal images,[12] x-ray computed tomography (CAT),[13] endoscopic images,[14] laser scanning confocal microscopy,[13] extraction of edge field,[15] image segmentation in general,[16,17] spatio-temporal segmentation of video images,[18] etc.

The output of the thresholding operation is a binary image whose one state will indicate the foreground objects, that is, printed text, a legend, a target, defective part of a material, etc., while the complementary state will correspond to the background. Depending on the application, the foreground can be represented by gray-level 0, that is, black as for text, and the background by the highest luminance for document paper, that is 255 in 8-bit images, or conversely the foreground by white and the background by black. Various factors, such as nonstationary and correlated noise, ambient illumination, busyness of gray levels within the object and its background, inadequate contrast, and object size not commensurate with the scene, complicate the thresholding operation. Finally, the lack of objective measures to assess the performance of various thresholding algorithms, and the difficulty of extensive testing in a task-oriented environment, have been other major handicaps.

In this study we develop taxonomy of thresholding algorithms based on the type of information used, and we assess their performance comparatively using a set of objective segmentation quality metrics. We distinguish six categories, namely, thresholding algorithms based on the exploitation of: 1. histogram shape information, 2. measurement space clustering, 3. histogram entropy information, 4. image attribute information, 5. spatial information, and 6. local characteristics. In this assessment study we envisage two major application areas of thresholding, namely document binarization and segmentation of nondestructive testing (NDT) images.

A document image analysis system includes several image-processing tasks, beginning with digitization of the document and ending with character recognition and natural language processing. The thresholding step can affect quite critically the performance of successive steps such as

Paper 02016 received Feb. 7, 2002; revised manuscript received Jan. 23, 2003; accepted for publication May 27, 2003.
1017-9909/2004/$15.00 © 2004 SPIE and IS&T.

classification of the document into text objects, and the correctness of the optical character recognition (OCR). Improper thresholding causes blotches, streaks, erasures on the document confounding segmentation, and recognition tasks. The merges, fractures, and other deformations in the character shapes as a consequence of incorrect thresholding are the main reasons of OCR performance deterioration.

In NDT applications, the thresholding is again often the first critical step in a series of processing operations such as morphological filtering, measurement, and statistical assessment. In contrast to document images, NDT images can derive from various modalities, with differing application goals. Furthermore, it is conjectured that the thresholding algorithms that apply well for document images are not necessarily the good ones for the NDT images, and vice versa, given the different nature of the document and NDT images.

There have been a number of survey papers on thresholding. Lee, Chung, and Park[19] conducted a comparative analysis of five global thresholding methods and advanced useful criteria for thresholding performance evaluation. In an earlier work, Weszka and Rosenfeld[20] also defined several evaluation criteria. Palumbo, Swaminathan and Srihari[21] addressed the issue of document binarization comparing three methods, while Trier and Jain[3] had the most extensive comparison basis (19 methods) in the context of character segmentation from complex backgrounds. Sahoo *et al.*[22] surveyed nine thresholding algorithms and illustrated comparatively their performance. Glasbey[23] pointed out the relationships and performance differences between 11 histogram-based algorithms based on an extensive statistical study.

This survey and evaluation, on the one hand, represents a timely effort, in that about 60% of the methods discussed and referenced are dating after the last surveys in this area.[19,23] We describe 40 thresholding algorithms with the idea underlying them, categorize them according to the information content used, and describe their thresholding functions in a streamlined fashion. We also measure and rank their performance comparatively in two different contexts, namely, document images and NDT images. The image repertoire consists of printed circuit board (PCB) images, eddy current images, thermal images, microscope cell images, ultrasonic images, textile images, and reflective surfaces as in ceramics, microscope material images, as well as several document images. For an objective performance comparison, we employ a combination of five criteria of shape segmentation goodness.

The outcome of this study is envisaged to be the formulation of the large variety of algorithms under a unified notation, the identification of the most appropriate types of binarization algorithms, and deduction of guidelines for novel algorithms. The structure of the work is as follows: Notation and general formulations are given in Sec. 2. In Secs. 3 to 8, respectively, histogram shape-based, clustering-based, entropy-based, object attribute-based, spatial information-based, and finally locally adaptive thresholding methods are described. In Sec. 9 we present the comparison methodology and performance criteria. The evaluation results of image thresholding methods, separately for nondestructive inspection and document processing applications, are given in Sec. 10. Finally, Sec. 11 draws the main conclusions.

2 Categories and Preliminaries

We categorize the thresholding methods in six groups according to the information they are exploiting. These categories are:

1. histogram shape-based methods, where, for example, the peaks, valleys and curvatures of the smoothed histogram are analyzed
2. clustering-based methods, where the gray-level samples are clustered in two parts as background and foreground (object), or alternately are modeled as a mixture of two Gaussians
3. entropy-based methods result in algorithms that use the entropy of the foreground and background regions, the cross-entropy between the original and binarized image, etc.
4. object attribute-based methods search a measure of similarity between the gray-level and the binarized images, such as fuzzy shape similarity, edge coincidence, etc.
5. the spatial methods use higher-order probability distribution and/or correlation between pixels
6. local methods adapt the threshold value on each pixel to the local image characteristics.

In the sequel, we use the following notation. The histogram and the probability mass function (PMF) of the image are indicated, respectively, by $h(g)$ and by $p(g)$, $g=0...G$, where G is the maximum luminance value in the image, typically 255 if 8-bit quantization is assumed. If the gray-value range is not explicitly indicated as $[g_{\min}, g_{\max}]$, it will be assumed to extend from 0 to G. The cumulative probability function is defined as

$$P(g)=\sum_{i=0}^{g} p(i).$$

It is assumed that the PMF is estimated from the histogram of the image by normalizing it to the total number of samples. In the context of document processing, the foreground (object) becomes the set of pixels with luminance values less than T, while the background pixels have luminance value above this threshold. In NDT images, the foreground area may consists of darker (more absorbent, denser, etc.) regions or conversely of shinier regions, for example, hotter, more reflective, less dense, etc., regions. In the latter contexts, where the object appears brighter than the background, obviously the set of pixels with luminance greater than T will be defined as the foreground.

The foreground (object) and background PMFs are expressed as $p_f(g)$, $0 \leqslant g \leqslant T$, and $p_b(g)$, $T+1 \leqslant g \leqslant G$, respectively, where T is the threshold value. The foreground and background area probabilities are calculated as:

$$P_f(T)=P_f=\sum_{g=0}^{T} p(g), \quad P_b(T)=P_b=\sum_{g=T+1}^{G} p(g). \qquad (1)$$

Table 1 Thresholding functions for the shape-based algorithms.

Shape_Rosenfeld[24]	$T_{\text{opt}}=\arg\max\{[p(g)-\text{Hull}(g)]\}$ by considering object attributes, such as busyness.
Shape_Sezan[32]	$T_{\text{opt}}=\gamma\{\text{first terminating zero of } \tilde{p}(g)\}+(1-\gamma)\times\{\text{first initiating zero of } \tilde{p}(g)\}$ $0\leqslant\gamma\leqslant 1$, $\tilde{p}(g)=d/dg[p(g)*\text{smoothing\_kernel}(g)]$, where $\gamma=1$ and kernel size is 55
Shape_Olivio	$T_{\text{opt}}=\langle T^0,$ valley at the highest resolution$\vert T^1,\ldots T^k,$ given valleys at k lower resolutions$\rangle$, $k=3$
Shape_Ramesh[30]	$T_{\text{opt}}=\min\left[\sum_{g=0}^{T}[b_1(T)-g]^2+\sum_{g=T+1}^{G}[b_2(T)-g]^2\right]$ where $b_1(T)=m_f(T)/P(T)$, $b_2(T)=m_b(T)[1-P(T)]$
Shape_Guo[28]	$T_{\text{opt}}=\min_g \frac{1}{\vert 1-\sum_{i=1}^{p}a_i\exp(-j2\pi gi/256)\vert^2}$, where $\{a_i\}_{i=1}^{p}$ the n'th order AR coefficients

The Shannon entropy, parametrically dependent on the threshold value T for the foreground and background, is formulated as:

$$H_f(T)=-\sum_{g=0}^{T} p_f(g)\log p_f(g),$$
$$H_b(T)=-\sum_{g=T+1}^{G} p_b(g)\log p_b(g). \tag{2}$$

The sum of these two is expressed as $H(T)=H_f(T)+H_b(T)$. When the entropy is calculated over the input image distribution $p(g)$ (and not over the class distributions), then obviously it does not depend on the threshold T, and hence is expressed simply as H. For various other definitions of the entropy in the context of thresholding, with some abuse of notation, we use the same symbols of $H_f(T)$ and $H_b(T)$.

The fuzzy measures attributed to the background and foreground events, that is, the degree to which the gray level, g, belongs to the background and object, respectively, and are symbolized by $\mu_f(g)$ and $\mu_b(g)$. The mean and variance of the foreground and background as functions of the thresholding level T can be similarly denoted as:

$$m_f(T)=\sum_{g=0}^{T} gp(g) \quad \sigma_f^2(T)=\sum_{g=0}^{T}[g-m_f(T)]^2p(g), \tag{3}$$

$$m_b(T)=\sum_{g=T+1}^{G} gp(g)$$
$$\sigma_b^2(T)=\sum_{g=T+1}^{G}[g-m_b(T)]^2p(g). \tag{4}$$

We refer to a specific thresholding method, which was programmed in the simulation analysis and whose formula appears in the table, with the descriptor, ''category_author.'' For example, Shape_Sezan and Cluster_Otsu, refer, respectively, to the shape-based thresholding method introduced in a paper by Sezan and to the clustering-based thresholding method first proposed by Otsu.

3 Histogram Shape-Based Thresholding Methods

This category of methods achieves thresholding based on the shape properties of the histogram (see Table 1). The shape properties come into play in different forms. The distance from the convex hull of the histogram is investigated in Refs. 20, and 24–27, while the histogram is forced into a smoothed two-peaked representation via autoregressive modeling in Refs. 28 and 29. A more crude rectangular approximation to the lobes of the histogram is given in Refs. 30 and 31. Other algorithms search explicitly for peaks and valleys, or implicitly for overlapping peaks via curvature analysis.[32–34]

Convex hull thresholding. Rosenfeld's method[24] (Shape_Rosenfeld) is based on analyzing the concavities of the histogram $h(g)$ vis-à-vis its convex hull, Hull(g), that is the set theoretic difference $\vert\text{Hull}(g)-p(g)\vert$. When the convex hull of the histogram is calculated, the deepest concavity points become candidates for a threshold. In case of competing concavities, some object attribute feedback, such as low busyness of the threshold image edges, could be used to select one of them. Other variations on the theme can be found in Weszka and Rosenfeld,[20,25] and Halada and Osokov.[26] Sahasrabudhe and Gupta[27] have addressed the histogram valley-seeking problem. More recently, Whatmough[35] has improved on this method by considering the exponential hull of the histogram.

Peak-and-valley thresholding. Sezan[32] (Shape_Sezan) carries out the peak analysis by convolving the histogram function with a smoothing and differencing kernel. By adjusting the smoothing aperture of the kernel and resorting to peak merging, the histogram is reduced to a two-lobe function. The differencing operation in the kernel outputs the triplet of incipient, maximum, and terminating zero-crossings of the histogram lobe $S=[(e_i,m_i,s_i),i=1,...2]$. The threshold sought should be somewhere between the first terminating and second initiating zero crossing, that is:

$T_{\text{opt}}=\gamma e_1+(1-\gamma)s_2$, $0\leqslant\gamma\leqslant 1$. In our work, we have found that $\gamma=1$ yields good results. Variations on this theme are provided in Boukharouba, Rebordao, and Wendel,[36] where the cumulative distribution of the image is first expanded in terms of Tschebyshev functions, followed by the curvature analysis. Tsai[37] obtains a smoothed histogram via Gaussians, and the resulting histogram is investigated for the presence of both valleys and sharp curvature points. We point out that the curvature analysis becomes effective when the histogram has lost its bimodality due to the excessive overlapping of class histograms.

In a similar vein, Carlotto[33] and Olivo[34] (Shape_Olivio) *consider* the multiscale analysis of the PMF and interpret its fingerprints, that is, the course of its zero crossings and extrema over the scales. In Ref. 34 using a discrete dyadic wavelet transform, one obtains the multiresolution analysis of the histogram, $p^s(g)=p(g)*\psi_s(g)$, $s=1,2...$, where $p^0(g)=p(g)$ is the original normalized histogram. The threshold is defined as the valley (minimum) point following the first peak in the smoothed histogram. This threshold position is successively refined over the scales starting from the coarsest resolution. Thus starting with the valley point $T^{(k)}$ at the k'th coarse level, the position is backtracked to the corresponding extrema in the higher resolution histograms $p^{(k-1)}(g)\ldots p^{(0)}(g)$, that is, $T^{(0)}$ is estimated by refining the sequence $T^{(1)}\ldots T^{(k)}$ (in our work $k=3$ was used).

Shape-modeling thresholding. Ramesh, Yoo, and Sethi[30] (Shape_Ramesh) use a simple functional approximation to the PMF consisting of a two-step function. Thus the sum of squares between a bilevel function and the histogram is minimized, and the solution for T_{opt} is obtained by iterative search. Kampke and Kober[31] have generalized the shape approximation idea.

In Cai and Liu,[29] the authors have approximated the spectrum as the power spectrum of multi-complex exponential signals using Prony's spectral analysis method. A similar all-pole model was assumed in Guo[28] (Shape_Guo). We have used a modified approach, where an autoregressive (AR) model is used to smooth the histogram. Here one begins by interpreting the PMF and its mirror reflection around $g=0$, $p(-g)$, as a noisy power spectral density, given by $\tilde{p}(g)=p(g)$ for $g\geqslant 0$, and $p(-g)$ for $g\leqslant 0$. One then obtains the autocorrelation coefficients at lags $k=0...G$, by the inverse discrete fourier transform (IDFT) of the original histogram, that is, $r(k)=\text{IDFT}[\tilde{p}(g)]$. The autocorrelation coefficients $\{r(k)\}$ are then used to solve for the n'th order AR coefficients $\{a_i\}$. In effect, one smoothes the histogram and forces it to a bimodal or two-peaked representation via the n'th order AR model ($n=1,2,3,4,5,6$). The threshold is established as the minimum, resting between its two pole locations, of the resulting smoothed AR spectrum.

4 Clustering-Based Thresholding Methods

In this class of algorithms, the gray-level data undergoes a clustering analysis, with the number of clusters being set always to two. Since the two clusters correspond to the two lobes of a histogram (assumed distinct), some authors search for the midpoint of the peaks.[38–41] In Refs. 42–45, the algorithm is based on the fitting of the mixture of Gaussians. Mean-square clustering is used in Ref. 46, while fuzzy clustering ideas have been applied in Refs. 30 and 47. See Table 2 for these algorithms.

Iterative thresholding. Riddler[38] (Cluster_Riddler) advanced one of the first iterative schemes based on two-class Gaussian mixture models. At iteration n, a new threshold T_n is established using the average of the foreground and background class means. In practice, iterations terminate when the changes $|T_n-T_{n+1}|$ become sufficiently small. Leung and Fam [39] and Trussel[40] realized two similar methods. In his method, Yanni and Horne[41] (Cluster_Yanni) initializes the midpoint between the two assumed peaks of the histogram as $g_{\text{mid}}=(g_{\max}+g_{\min})/2$, where $g_{\max}$ is the highest nonzero gray level and $g_{\min}$ is the lowest one, so that $(g_{\max}-g_{\min})$ becomes the span of nonzero gray values in the histogram. This midpoint is updated using the mean of the two peaks on the right and left, that is, as $g^*_{\text{mid}}=(g_{\text{peak1}}+g_{\text{peak2}})/2$.

Clustering thresholding. Otsu[46] (Cluster_Otsu) suggested minimizing the weighted sum of within-class variances of the foreground and background pixels to establish an optimum threshold. Recall that minimization of within-class variances is tantamount to the maximization of between-class scatter. This method gives satisfactory results when the numbers of pixels in each class are close to each other. The Otsu method still remains one of the most referenced thresholding methods. In a similar study, thresholding based on isodata clustering is given in Velasco.[48] Some limitations of the Otsu method are discussed in Lee and Park.[49] Liu and Li[50] generalized it to a 2-D Otsu algorithm.

Minimum error thresholding. These methods assume that the image can be characterized by a mixture distribution of foreground and background pixels: $p(g)=P(T).p_f(g)+[1-P(T)].p_b(g)$. Lloyd[42] (Cluster_Lloyd) considers equal variance Gaussian density functions, and minimizes the total misclassification error via an iterative search. In contrast, Kittler and Illingworth[43,45] (Cluster_Kittler) removes the equal variance assumption and, in essence, addresses a minimum-error Gaussian density-fitting problem. Recently Cho, Haralick, and Yi[44] have suggested an improvement of this thresholding method by observing that the means and variances estimated from truncated distributions result in a bias. This bias becomes noticeable, however, only whenever the two histogram modes are not distinguishable.

Fuzzy clustering thresholding. Jawahar, Biswas, and Ray[47] (Cluster_Jawahar_1), and Ramesh Yoo, and Sethi[30] assign fuzzy clustering memberships to pixels depending on their differences from the two class means. The cluster means and membership functions are calculated as:

$$m_k=\frac{\sum_{g=0}^{G} g.p(g)\mu_k^{\tau}(g)}{\sum_{g=0}^{G} p(g)\mu_k^{\tau}(g)}, \quad k=f,b,$$

Table 2 Thresholding functions for clustering-based algorithms.

Cluster_Riddler[34]	$T_{\text{opt}}=\lim_{n\to\infty}\frac{m_f(T_n)+m_b(T_n)}{2}$ where $m_f(T_n)=\sum_{g=0}^{T_n} gp(g) \quad m_b(T_n)=\sum_{g=T_n+1}^{G} gp(g)$
Cluster_Yanni[41]	$T_{\text{opt}}=(g_{\max}-g_{\min})\sum_{g=g_{\min}}^{g^*_{\text{mid}}} p(g)$
Cluster_Otsu[46]	$T_{\text{opt}}=\arg\max\left\{\frac{P(T)[1-P(T)][m_f(T)-m_b(T)]^2}{P(T)\sigma_f^2(T)+[1-P(T)]\sigma_b^2(T)}\right\}$
Cluster_Lloyd[24]	$T_{\text{opt}}=\arg\min\left[\frac{m_f(T)+m_b(T)}{2}+\frac{\sigma^2}{m_f(T)-m_b(T)}\log\frac{1-P(T)}{P(T)}\right]$ σ^2 is the variance of the whole image.
Cluster_Kittler[32]	$T_{\text{opt}}=\arg\min\{P(T)\log\sigma_f(T)+[1-P(T)]\log\sigma_b(T)-P(T)\log P(T)-[1-P(T)]\log[1-P(T)]\}$ where $\{\sigma_f(T),\sigma_b(T)\}$ are foreground and background standard deviations.
Cluster_Jawahar_1[47]	$T_{\text{opt}}=\arg\operatorname{equal}_g[\mu_f^\tau(g)=\mu_b^\tau(g)]$ where $d(g,m_k)=\sum_{g=0}^{T}(g-m_k)^2,\ k=b,f$
Cluster_Jawahar_2[47]	$T_{\text{opt}}=\arg\operatorname{equal}_g[\mu_f^\tau(g)=\mu_b^\tau(g)]$ where $d(g,m_k)=\frac{1}{2}\left(\frac{g-m_k}{\sigma_k}\right)^2+\log\sigma_k-\log\beta_k,\ k=b,f$

$$\mu_f^\tau(g)=\frac{1}{1+\left(\frac{d(g,m_f)}{d(g,m_b)}\right)^{2/\tau-1}},\quad \mu_b^\tau(g)=1-\mu_f^\tau(g).$$

In these expressions, $d(\ldots)$ is the Euclidean distance function between the gray-value g and the class mean, while τ is the fuzziness index. Notice that for $\tau=1$, one obtains the K-means clustering. In our experiments we used $\tau=2$. In a second method proposed by Jawahar, Biswas, and Ray[47] (Cluster_Jawahar_2), the distance function and the membership function are defined as (for $k=f,b$):

$$d(g,m_k)=\frac{1}{2}\left(\frac{g-m_k}{\sigma_k}\right)^2+\log\sigma_k-\log\beta_k,$$

$$\beta_k=\frac{\Sigma_{g=0}^{G}p(g)\mu_k^\tau(g)}{\Sigma_{g=0}^{G}p(g)[\mu_f^\tau(g)+\mu_b^\tau(g)]},$$

$$\sigma_k^2=\frac{\Sigma_{g=0}^{G}p(g)\mu_k^\tau(g)(g-m_k)^2}{\Sigma_{g=0}^{G}\mu_k^\tau(g)p(g)}.$$

In either method, the threshold is established as the crossover point of membership functions, i.e., $T_{\text{opt}}=\arg\operatorname{equal}_g\{\mu_f^\tau(g)=\mu_b^\tau(g)\}$.

5 Entropy-Based Thresholding Methods

This class of algorithms exploits the entropy of the distribution of the gray levels in a scene. The maximization of the entropy of the thresholded image is interpreted as indicative of maximum information transfer.[51–55] Other authors try to minimize the cross-entropy between the input gray-level image and the output binary image as indicative of preservation of information,[56–59] or a measure of fuzzy entropy.[60,61] Johannsen and Bille[62] and Pal, King, and Hashim[63] were the first to study Shannon entropy-based thresholding. See Table 3 for these algorithms.

Entropic thresholding. Kapur, Sahoo, and Wong[53] (Entropy_Kapur) consider the image foreground and background as two different signal sources, so that when the sum of the two class entropies reaches its maximum, the image is said to be optimally thresholded. Following this idea, Yen, Chang, and Chang[54] (Entropy_Yen) define the entropic correlation as

$$TC(T)=C_b(T)+C_f(T)=-\log\left\{\sum_{g=0}^{T}\left[\frac{p(g)}{P(T)}\right]^2\right\}-\log\left\{\sum_{g=T+1}^{G}\left[\frac{p(g)}{1-P(T)}\right]^2\right\}$$

and obtain the threshold that maximizes it. This method corresponds to the special case of the following method in

Table 3 Thresholding functions for entropy-based algorithms.

Entropy_Kapur[53]	$T_{opt}=\arg\max[H_f(T)+H_b(T)]$ with $H_f(T)=-\sum_{g=0}^{T}\frac{p(g)}{P(T)}\log\frac{p(g)}{P(T)}$ and $H_b(T)=-\sum_{g=T+1}^{G}\frac{p(g)}{P(T)}\log\frac{p(g)}{P(T)}$
Entropy_Yen[54]	$T_{opt}(T)=\arg\max\{C_b(T)+C_f(T)\}$ with $C_b(T)=-\log\left\{\sum_{g=0}^{T}\left[\frac{p(g)}{P(T)}\right]^2\right\}$ and $C_f(T)=-\log\left\{\sum_{g=T+1}^{G}\left[\frac{p(g)}{1-P(T)}\right]^2\right\}$
Entropy_Sahoo[55]	$T_{opt}=T_{[1]}[P_{T_{[1]}}+\frac{1}{4}\cdot w\cdot B_1]+\frac{1}{4}\cdot T_{[2]}\cdot w\cdot B_2+T_{[3]}\cdot[1-P_{T_{[3]}}+\frac{1}{4}\cdot w\cdot B_3]$ where $P[T_{(k)}]=\sum_{g=0}^{T_{[k]}}p(g)$, $k=1,2,3$, $w=P[T_{(3)}]-P[T_{(1)}]$ and $B_1,B_2,B_3=\begin{cases}(1,2,1) & \text{if } \lvert T_{[1]}-T_{[2]}\rvert\le 5 \text{ and } \lvert T_{[2]}-T_{[3]}\rvert\le 5 \text{ or } \lvert T_{[1]}-T_{[2]}\rvert>5 \text{ and } \lvert T_{[2]}-T_{[3]}\rvert>5\\ (0,1,3) & \text{if } \lvert T_{[1]}-T_{[2]}\rvert\le 5 \text{ and } \lvert T_{[2]}-T_{[3]}\rvert>5\\ (3,1,0) & \text{if } \lvert T_{[1]}-T_{[2]}\rvert>5 \text{ and } \lvert T_{[2]}-T_{[3]}\rvert\le 5\end{cases}$
Entropy_Pun_a[51]	$T_{opt}=\arg\text{equal}[H_f(T)=\alpha H(T)]$ where $\alpha=\arg\max\left[\frac{\alpha\log P(T)}{\log\{\max[p(1),\ldots,p(T)]\}}+(1-\alpha)\frac{\log[1-P(T)]}{\log\{\max[p(T+1),\ldots p(G)]\}}\right]$
Entropy_Pun_b[52]	$T_{opt}=\arg\left\{\sum_{g=0}^{T}p(g)=(0.5+\lvert 0.5-\alpha\rvert)\right\}$ optimizing histogram symmetry by tuning α.
Entropy_Li[56,57]	$T_{opt}=\arg\min\left[\sum_{g=0}^{T}gp(g)\log\frac{g}{m_f(T)}+\sum_{g=T+1}^{G}gp(g)\log\frac{g}{m_b(T)}\right]$ where $\sum_{g\le T}g=\sum_{g\le T}m_f(T)$ and $\sum_{g\ge T}g=\sum_{g\ge T}m_b(T)$.
Entropy_Brink[58]	$T_{opt}=\arg\min\{H(T)\}$ where $H(T)$ is $\sum_{g=0}^{T}p(g)\left[m_f(T)\log\frac{m_f(T)}{g}+g\log\frac{g}{m_f(T)}\right]+\sum_{T+1}^{G}p(g)\left[m_b(T)\log\frac{m_b(T)}{g}+g\log\frac{g}{m_b(T)}\right]$
Entropy_Pal[59]	$T_{opt}=\arg\max\{H_f(T)+H_b(T)\}$ where $H_f(T)=\sum_{g=0}^{T}\left[p_f(g)\log\frac{p_f(g)}{q_f(g)}+q_f(g)\log\frac{q_f(g)}{p_f(g)}\right]$ $H_b(T)=\sum_{g=T+1}^{G}\left[p_b(g)\log\frac{p_b(g)}{q_b(g)}+q_b(g)\log\frac{q_b(g)}{p_b(g)}\right]$ and $q_f(g)=\exp[-m_f(T)]\frac{m_f(T)^g}{g!}$, $g=0,\ldots T$, $q_b(g)=\exp[-m_b(T)]\frac{m_b(T)^g}{g!}$, $g=T+1,\ldots G$.
Entropy_Shanbag[60]	$T_{opt}=\arg\min\{\lvert H_f(T)-H_b(T)\rvert\}$ where $H_f(T)=-\sum_{g=0}^{T}\frac{p(g)}{P(T)}\log[\mu_f(g)]$, $H_b(T)=-\sum_{g=T+1}^{G}\frac{p(g)}{1-P(T)}\log[\mu_b(g)]$
Entropy_Cheng[61]	$\max_{a,b,c}\left\{H(A,\mu_A)=-\frac{1}{\log 2}[Q(A_f)\log Q(A_f)+Q(A_b)\log Q(A_b)]\right\}$ where μ_A is Zadeh's membership with parameters a,b,c, and $Q(A_f)=\sum_{\mu(g)\in A_f}p(g)$, $Q(A_b)=\sum_{\mu(g)\in A_b}p(g)$

Ref. 55 (Entropy_Sahoo) for the Renyi power $\rho=2$. Sahoo, Wilkins, and Yeager[55] combine the results of three different threshold values. The Renyi entropy of the foreground and background sources for some parameter ρ are defined as:

$$H_f^{\rho}=\frac{1}{1-\rho}\ln\left\{\sum_{g=0}^{T}\left[\frac{p(g)}{P(T)}\right]^{\rho}\right\} \quad \text{and}$$

$$H_b^{\rho}=\frac{1}{1-\rho}\ln\left\{\sum_{g=T+1}^{G}\left[\frac{p(g)}{1-P(T)}\right]^{\rho}\right\}.$$

They then find three different threshold values, namely, T_1, T_2, and T_3, by maximizing the sum of the foreground and background Renyi entropies for the three ranges of ρ, 0

$<\rho<1$, $\rho>1$, and $\rho=1$, respectively. For example T_2 for $\rho=1$ corresponds to the Kapur, Sahoo and Wong[53] threshold value, while for $\rho>1$, the threshold corresponds to that found in Yen, Chang, and Chang.[54] Denoting $T_{[1]}$, $T_{[2]}$, and $T_{[3]}$ as the rank ordered T_1, T_2, and T_3 values, ''optimum'' T is found by their weighted combination.

Finally, although the two methods of Pun[51,52] have been superseded by other techniques, for historical reasons, we have included them (Entropy_Pun1, Entropy_Pun2). In Ref. 51, Pun considers the gray-level histogram as a G-symbol source, where all the symbols are statistically independent. He considers the ratio of the *a posteriori* entropy $H'(T)=-P(T)\log[P(T)]-[1-P(T)]\log[1-P(T)]$ as a function of the threshold T to that of the source entropy

$$H(T)=-\sum_{g=0}^{T} p(g)\log[p(g)]-\sum_{g=T+1}^{G} p(g)\log[p(g)].$$

This ratio is lower bounded by

$$\frac{H'(T)}{H}\geqslant\left[\frac{\alpha\log P(T)}{\log\{\max[p(1),\ldots,p(T)]\}}+(1-\alpha)\frac{\log(1-P(T))}{\log\{\max[p(T+1),\ldots p(G)]\}}\right].$$

In a second method, the threshold[52] depends on the anisotropy parameter α, which depends on the histogram asymmetry.

Cross-entropic thresholding. Li, Lee, and Tam[56,57] (Entropy_Li) formulate the thresholding as the minimization of an information theoretic distance. This measure is the Kullback-Leibler distance

$$D(q,p)=\sum q(g)\log\frac{q(g)}{p(g)}$$

of the distributions of the observed image $p(g)$ and of the reconstructed image $q(g)$. The Kullback-Leibler measure is minimized under the constraint that observed and reconstructed images have identical average intensity in their foreground and background, namely the condition

$$\sum_{g\leqslant T} g=\sum_{g\leqslant T} m_f(T) \quad\text{and}\quad \sum_{g\geqslant T} g=\sum_{g\geqslant T} m_b(T).$$

Brink and Pendock[58] (Entropy_Brink) suggest that a threshold be selected to minimize the cross-entropy, defined as

$$H(T)=\sum_{g=0}^{T} q(g)\log\frac{q(g)}{p(g)}+\sum_{g=T+1}^{G} p(g)\log\frac{p(g)}{q(g)}.$$

The cross-entropy is interpreted as a measure of data consistency between the original and the binarized images. They show that this optimum threshold can also be found by maximizing an expression in terms of class means. A variation of this cross-entropy approach is given by specifically modeling the *a posteriori* PMF of the foreground and background regions, as in Pal[59] (Entropy_Pal). Using the maximum entropy principle in Shore and Johanson,[64] the corresponding PMFs are defined in terms of class means:

$$q_f(g)=\exp[-m_f(T)]\frac{m_f(T)^g}{g!},\quad g=0,\ldots T,$$

$$q_b(g)=\exp[-m_b(T)]\frac{m_b(T)^g}{g!},\quad g=T+1,\ldots G.$$

Wong and Sahoo[65] have also presented a former study of thresholding based on the maximum entropy principle.

Fuzzy entropic thresholding. Shanbag[60] (Entropy_Shanbag) considers the fuzzy memberships as an indication of how strongly a gray value belongs to the background or to the foreground. In fact, the farther away a gray value is from a presumed threshold (the deeper in its region), the greater becomes its potential to belong to a specific class. Thus, for any foreground and background pixel, which is i levels below or i levels above a given threshold T, the membership values are determined by

$$\mu_f(T-i)=0.5+\frac{p(T)+\ldots+p(T-1-i)+p(T-i)}{2P(T)},$$

that is, its measure of belonging to the foreground, and by

$$\mu_b(T+i)=0.5+\frac{p(T+1)+\ldots+p(T-1+i)+p(T+i)}{2(1-P(T))},$$

respectively. Obviously on the gray value corresponding to the threshold, one should have the maximum uncertainty, such that $\mu_f(T)=\mu_b(T)=0.5$. The optimum threshold is found as the T that minimizes the sum of the fuzzy entropies,

$$T_{\text{opt}}=\arg\min_T\{|H_0(T)-H_1(T)|\},$$

$$H_0(T)=-\sum_{g=0}^{T}\frac{p(g)}{P(T)}\log[\mu_0(g)],$$

$$H_1(T)=-\sum_{g=T+1}^{G}\frac{p(g)}{1-P(T)}\log[\mu_1(g)],$$

since one wants to get equal information for both the foreground and background. Cheng, Chen, and Sun's method[61] (Entropy_Cheng) relies on the maximization of fuzzy event entropies, namely, the foreground A_f and background A_b subevents. The membership function is assigned using Zadeh's S function in Ref. 66. The probability of the foreground subevent $Q(A_f)$ is found by summing those gray-value probabilities that map into the A_f subevent:

$$Q(A_f)=\sum_{\mu(g)\in A_f} p(g)$$

Table 4 Thresholding functions for the attribute-based algorithms.

Attribute_Tsai[70]	$T_{\text{opt}}=\arg\text{equal}[m_1=b_1(T),m_2=b_2(T),m_1=b_3(T)]$ where $m_k=\sum_{g=0}^{G}p(g)g^k$ and $b_k=P_f m_f^k+P_b m_b^k$
Attribute_Hertz[67]	$T_{\text{opt}}=\arg\max[E_{\text{gray}}\cap E_{\text{binary}}(T)]$, where E_{gray}: gray-level edge field, $E_{\text{binary}}(T)$ binary edge field
Attribute_Huang[82]	$T_{\text{opt}}=\arg\min\left\{-\frac{1}{N^2\log 2}\sum_{g=0}^{G}[\mu_f(g,T)\log(\mu_f(g,T)]+[1-\mu_f(g,T)]\log[1-\mu_f(g,T)]p(g)\right\}$ where $\mu_f[I(i,j),T]=\frac{G}{G+\lvert I(i,j)-m_f(T)\rvert}$
Attribute_Pikaz[76]	$T_{\text{opt}}=$Arg[most stable point of s-sized object function $N_s(T)$]
Attribute_Leung[81]	$T_{\text{opt}}=\arg\max\{-p_f\log p_f-(1-p_f)\log(1-p_f)+p_f(\pi_{ff}\log\pi_{ff}+\pi_{bf}\log\pi_{bf})+(1-p_f)(\pi_{fb}\log\pi_{fb}+\pi_{bb}\log\pi_{bb})\}$ where $\pi_{bf}=$Prob(segmented as background|belongs to the foreground) etc.
Attribute_Pal[68]	$T_{\text{opt}}=\arg\max\{\text{Compactness}[\mu(T)]\}=\frac{\text{Area}[\mu(T)]}{\text{Perim}[\mu(T)]^2}$, over all foreground regions.

and similarly for the background.

$$Q(A_b)=\sum_{\mu(g)\in A_b}p(g)\ \text{ as }\ H(A,\mu_A)$$
$$=-\frac{1}{\log 2}[Q(A_f)\log Q(A_f)+Q(A_b)\log Q(A_b)].$$

In other words, $Q(A_i)$, $i=f,b$ corresponds to the probabilities summed in the g domain for all gray values mapping into the A_i subevent. One maximizes this entropy of the fuzzy event over the parameters (a, b, c) of the S function. The threshold T is the value g satisfying the partition for $\mu_A(g)=0.5$.

6 Thresholding Based on Attribute Similarity

These algorithms select the threshold value based on some attribute quality or similarity measure between the original image and the binarized version of the image. These attributes can take the form of edge matching,[67,15] shape compactness,[68,69] gray-level moments,[70–72] connectivity,[73] texture,[74,75] or stability of segmented objects.[7,76] Some other algorithms evaluate directly the resemblance of the original gray-level image to binary image resemblance using fuzzy measure[77–79] or resemblance of the cumulative probability distributions,[80] or in terms of the quantity of information revealed as a result of segmentation.[81] See Table 4 for examples.

Moment preserving thresholding. Tsai[70] (Attribute_Tsai) considers the gray-level image as the blurred version of an ideal binary image. The thresholding is established so that the first three gray-level moments match the first three moments of the binary image. The gray-level moments m_k and binary image moments b_k are defined, respectively, as:

$$m_k=\sum_{g=0}^{G}p(g)g^k\quad\text{and}\quad b_k=P_f m_f^k+P_b m_b^k.$$

Cheng and Tsai[71] reformulate this algorithm based on neural networks. Delp and Mitchell[72] have extended this idea to quantization.

Edge field matching thresholding. Hertz and Schafer[67] (Attribute_Hertz) consider a multithresholding technique where a thinned edge field, obtained from the gray-level image E_{gray}, is compared with the edge field derived from the binarized image $E_{\text{binary}}(T)$. The global threshold is given by the value that maximizes the coincidence of the two edge fields based on the count of matching edges, and penalizes the excess original edges and the excess thresholded image edges. Both the gray-level image edge field and the binary image edge field have been obtained via the Sobel operator. In a complementary study, Venkatesh and Rosin[15] have addressed the problem of optimal thresholding for edge field estimation.

Fuzzy similarity thresholding. Murthy and Pal[77] were the first to discuss the mathematical framework for fuzzy thresholding, while Huang and Wang[82] (Attribute_Huang) proposed an index of fuzziness by measuring the distance between the gray-level image and its crisp (binary) version. In these schemes, an image set is represented as the double array $F=\{I(i,j),\mu_f[I(i,j)]\}$, where $0\leqslant\mu_f[I(i,j)]\leqslant 1$ represents for each pixel at location (i,j) its fuzzy measure to belong to the foreground. Given the fuzzy membership

value for each pixel, an index of fuzziness for the whole image can be obtained via the Shannon entropy or the Yager's measure.[78] The optimum threshold is found by minimizing the index of fuzziness defined in terms of class (foreground, background) medians or means $m_f(T)$, $m_b(T)$ and membership functions $\mu_f[I(i,j),T]$, $\mu_b[I(i,j),T]$. Ramar *et al.*[79] have evaluated various fuzzy measures for threshold selection, namely linear index of fuzziness, quadratic index of fuzziness, logarithmic entropy measure, and exponential entropy measure, concluding that the linear index works best.

Topological stable-state thresholding. Russ[7] has noted that experts in microscopy subjectively adjust the thresholding level at a point where the edges and shape of the object get stabilized. Similarly Pikaz and Averbuch[76] (Attribute_Pikaz) pursue a threshold value, which becomes stable when the foreground objects reach their correct size. This is instrumented via the size-threshold function $N_s(T)$, which is defined as the number of objects possessing at least s number of pixels. The threshold is established in the widest possible plateau of the graph of the $N_s(T)$ function. Since noise objects rapidly disappear with shifting the threshold, the plateau in effect reveals the threshold range for which foreground objects are easily distinguished from the background. We chose the middle value of the largest sized plateau as the optimum threshold value.

Maximum information thresholding. Leung and Lam[81] (Attribute_Leung) define the thresholding problem as the change in the uncertainty of an observation on specification of the foreground and background classes. The presentation of any foreground/background information reduces the class uncertainty of a pixel, and this information gain is measured by $H(p)-\alpha H(p_f)-(1-\alpha)H(p_b)$, where $H(p)$ is the initial uncertainty of a pixel and α is the probability of a pixel to belong to the foreground class. The optimum threshold is established as that generating a segmentation map that, in turn, minimizes the average residual uncertainty about which class a pixel belongs after the segmented image has been observed. Such segmentation would obviously minimize the wrong classification probability of pixels, in other words, the false alarms π_{fb} (pixel appears in the foreground while actually belonging to the background) and the miss probability π_{bf}. According to this notation π_{ff}, π_{bb} denote the correct classification conditionals. The optimum threshold corresponds to the maximum decrease in uncertainty, which implies that the segmented image carries as close a quantity of information as that in the original information.

Enhancement of fuzzy compactness thresholding. Rosenfeld generalized the concept of fuzzy geometry.[69] For example, the area of a fuzzy set is defined as

$$\mathrm{Area}(\mu)=\sum_{i,j=1}^{N}\mu[I(i,j)]$$

while its perimeter is given by

$$\mathrm{Perim}(\mu)=\sum_{i,j=1}^{N}|\mu[I(i,j)]-\mu[I(i,j+1)]|+\sum_{i,j=1}^{N}|\mu[I(i,j)]-\mu[I(i+1,j)]|,$$

where the summation is taken over any region of nonzero membership, and N is the number of regions in a segmented image. Pal and Rosenfeld[68] (Attribute_Pal) evaluated the segmentation output, such that both the perimeter and area are functions of the threshold T. The optimum threshold is determined to maximize the compactness of the segmented foreground sets. In practice, one can use the standard S function to assign the membership function at the pixel $I(i,j)$: $\mu[I(i,j)]=S[I(i,j);a,b,c]$, as in Kaufmann,[66] with crossover point $b=(a+c)/2$ and bandwidth $\Delta b=b-a=c-b$. The optimum threshold T is found by exhaustively searching over the $(b,\Delta b)$ pairs to minimize the compactness figure. Obviously the advantage of the compactness measure over other indexes of fuzziness is that the geometry of the objects or fuzziness in the spatial domain is taken into consideration.

Other studies involving image attributes are as follows. In the context of document image binarization, Liu, Srihari, and Fenrich[74,75] have considered document image binarization based on texture analysis, while Don[83] has taken into consideration noise attribute of images. Guo[84] develops a scheme based on morphological filtering and the fourth order central moment. Solihin and Leedham[85] have developed a global thresholding method to extract handwritten parts from low-quality documents. In another interesting approach, Aviad and Lozinskii[86] have introduced semantic thresholding to emulate the human approach to image binarization. The semantic threshold is found by minimizing measures of conflict criteria, so that the binary image resembles most to a verbal description of the scene. Gallo and Spinello[87] have developed a technique for thresholding and isocontour extraction using fuzzy arithmetic. Fernandez[80] has investigated the selection of a threshold in matched filtering applications in the detection of small target objects. In this application, the Kolmogorov-Smirnov distance between the background and object histograms is maximized as a function of the threshold value.

7 Spatial Thresholding Methods

This class of algorithms utilizes not only gray value distribution but also dependency of pixels in a neighborhood, for example, in the form of context probabilities, correlation functions, cooccurrence probabilities, local linear dependence models of pixels, 2-D entropy, etc. One of the first to explore spatial information was Kirby and Rosenfeld,[88] who considered local average gray levels for thresholding. Others have followed using relaxation to improve on the binary map as in Refs. 89 and 90, the Laplacian of the images to enhance histograms,[25] the quadtree thresholding,[91] and second-order statistics.[92] Cooccurrence probabilities have been used as indicators of spatial dependence as in Refs. 93–96. The characteristics of ''pixels jointly with their local average'' have been considered via their second-order entropy as in Refs. 97–100 and via the fuzzy partitioning as in Refs. 101, 102, and 103. The local

Table 5 Thresholding functions for spatial thresholding methods.

Spatial_Pal_1 and Spatial_Pal_2[94]	$T_{opt}=\arg\max[H_{bb}(T)+H_{ff}(T)]$ or $T_{opt}=\arg\max[H_{fb}(T)+H_{bf}(T)]$ where $H_{fb}(T)$, $H_{bf}(T)$, $H_{ff}(T)$, $H_{bb}(T)$ are the co-occurrence entropies
Spatial_Abutaleb[97]	$(T_{opt},\bar{T}_{opt})=\arg\min\{\log[P(T,\bar{T})[1-P(T,\bar{T})]+H_f/P(T,\bar{T})+H_b/[1-P(T,\bar{T})]\}$ where $H_f=-\sum_{i=1}^{T}\sum_{j=1}^{\bar{T}}\frac{p(g,\bar{g})}{P(T,\bar{T})}\log\frac{p(g,\bar{g})}{P(T,\bar{T})}$ and $H_b=-\sum_{i=T+1}^{G}\sum_{j=\bar{T}+1}^{G}\frac{p(g,\bar{g})}{[1-P(T,\bar{T})]}\log\frac{p(g,\bar{g})}{[1-P(T,\bar{T})]}$
Spatial_Beghdadi[104]	$T_{opt}=\arg\min\left[-\sum_{k=0}^{s\times s}p_k^{block}(T)\cdot\log p_k^{block}(T)\right]$ where $p_k^{block}(T)$ is the probability of $s\times s$ size blocks containing k whites and s^2-k blacks ($s=2,4,8,16$)
Spatial_Cheng[101]	$T_{opt}=\max_{a,b,c}\{H_{fuzzy}(\text{foreground})+H_{fuzzy}(\text{background})\}$ where $H_{fuzzy}(A)=-\sum_{x,y}\mu_A(x,y)p(x,y)\log p(x,y)$, $\{a,b,c\}$, the S-function parameters; $\{A=$ foreground, background$\}$; $\{x,y\}=\{$ pixel value, local average value within 3×3 region$\}$.

spatial dependence of pixels is captured in Ref. 104 as binary block patterns. Thresholding based on explicit *a posteriori* spatial probability estimation was analyzed in Ref. 105, and thresholding as the max-min distance to the extracted foreground object was considered in Ref. 106. See Table 5 for these methods.

Cooccurrence thresholding methods. Chanda and Majumder[96] have suggested the use of cooccurrences for threshold selection, and Lie[93] has proposed several measures to this effect. In this vein, Pal[94] (Spatial_Pal), realizing that two images with identical histograms can yet have different n'th order entropies due to their spatial structure, considered the cooccurrence probability of the gray values g_1 and g_2 over its horizontal and vertical neighbors. Thus the pixels, first binarized with threshold value T, are grouped into background and foreground regions. The cooccurrence of gray levels k and m is calculated as

$$c_{k,m}=\sum_{\text{all pixels}}\delta, \text{ where } \delta=1 \text{ if}$$
$$\{[I(i,j)=k]\wedge[I(i,j+1)=m]\vee[I(i,j)=k]\wedge[I(i+1,j)=m]\},$$

and $\delta=0$ otherwise. Pal proposes two methods to use the cooccurrence probabilities. In the first expression, the binarized image is forced to have as many background-to-foreground and foreground-to-background transitions as possible. In the second approach, the converse is true, in that the probability of the neighboring pixels staying in the same class is rewarded.

Chang, Chen, Wang and Althouse[95] establish the threshold in such a way that the cooccurrence probabilities of the original image and of the binary image are minimally divergent. As a measure of similarity, the directed divergence or the Kullback-Leibler distance is used. More specifically, consider the four quadrants of the cooccurrence matrix as illustrated in Fig. 1, where the first quadrant denotes the background-to-background (bb) transitions, and the third quadrant to the foreground-to-foreground (ff) transitions. Similarly, the second and fourth quadrants denote, respectively, the background-to-foreground (bf) and the foreground-to-background (fb) transitions. Using the cooccurrence probabilities pij, (that is, the score of i to j gray level transitions normalized by the total number of transitions) the quadrant probabilities are obtained as:

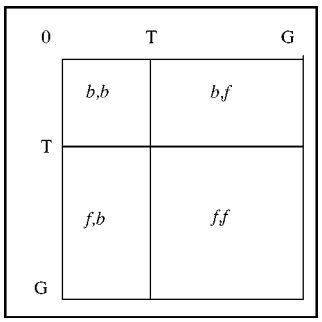

Fig. 1 Co-occurrence matrix.

$$P_{bb}(T)=\sum_{i=0}^{T}\sum_{j=0}^{T} p_{ij},\quad P_{bf}(T)=\sum_{i=0}^{T}\sum_{j=T+1}^{G} p_{ij},$$

$$P_{fb}(T)=\sum_{i=T+1}^{G}\sum_{j=0}^{T} p_{ij},\quad P_{ff}(T)=\sum_{i=T+1}^{G}\sum_{j=T+1}^{G} p_{ij},$$

and similarly for the thresholded image, one finds the quantities

$$Q_{bb}(T)=\sum_{i=0}^{T}\sum_{j=0}^{T} q_{ij},\quad Q_{bf}(T)=\sum_{i=0}^{T}\sum_{j=T+1}^{G} q_{ij},$$

$$Q_{\mathrm{fb}}(T)=\sum_{i=T+1}^{G}\sum_{j=0}^{T} q_{ij},\quad Q_{ff}(T)=\sum_{i=T+1}^{G}\sum_{j=T+1}^{G} q_{ij},$$

$$T_{\mathrm{opt}}=\mathrm{argmin}[P_{bb}(T)\log Q_{bb}(T)+P_{bf}(T)\log Q_{bf}(T)+P_{ff}(T)\log Q_{ff}(T)+P_{fb}(T)\log Q_{fb}(T)].$$

Higher-order entropy thesholding. Abutaleb[97] (Spatial_Abutaleb) considers the joint entropy of two related random variables, namely, the image gray value g at a pixel, and the average gray value $\overline{g}$ of a neighborhood centered at that pixel. Using the 2-D histogram $p(g,\overline{g})$, for any threshold pair $(T,\overline{T})$, one can calculate the cumulative distribution $P(T,\overline{T})$, and then define the foreground entropy as

$$H_f=-\sum_{i=1}^{T}\sum_{j=1}^{\overline{T}}\frac{p(g,\overline{g})}{P(T,\overline{T})}\log\frac{p(g,\overline{g})}{P(T,\overline{T})}.$$

Similarly, one can define the background region's second order entropy. Under the assumption that the off-diagonal terms, that is the two quadrants $[(0,T),(\overline{T},G)]$ and $[(T,G),(0,\overline{T})]$ are negligible and contain elements only due to image edges and noise, the optimal pair $(T,\overline{T})$ can be found as the minimizing value of the 2-D entropy functional. In Wu, Songde, and Hanqing,[10] a fast recursive method is suggested to search for the $(T,\overline{T})$ pair. Cheng and Chen[98] have presented a variation of this theme by using fuzzy partitioning of the 2-D histogram of the pixels and their local average. Li, Gong, and Chen[99] have investigated the Fisher linear projection of the 2-D histogram. Brink[100] has modified Abutaleb's expression by redefining class entropies and finding the threshold as the value that maximizes the minimum (maximin) of the foreground and background entropies: more explicitly, $(T_{\mathrm{opt}},\overline{T}_{\mathrm{opt}})=\max\{\min[H_f(T,\overline{T}),H_b(T,\overline{T})]\}$.

Beghdadi, Negrate, and Lesegno[104] (Spatial_Beghdadi), on the other hand, exploit the spatial correlation of the pixels using entropy of block configurations as a symbol source. For any threshold value T, the image becomes a set of juxtaposed binary blocks of size $s\times s$ pixels. Letting B_k represent the subset of $(s\times s)$ blocks out of $N=2^{s^2}$ containing k whites and K-k blacks, their relative population becomes the binary source probabilities $p_k^{\mathrm{block}}=\mathrm{Prob}\{\mathrm{block}\in B_k\}$. Here p_k^{block} represents the probability of the block containing k $(0\leqslant k\leqslant s\times s)$ whites irrespective of the binary pixel configurations. An optimum gray level threshold is found by maximizing the entropy function of the block probabilities. The choice of the block size is a compromise between image detail and computational complexity. As the block size becomes large, the number of configurations increases rapidly; on the other hand, small blocks may not be sufficient to describe the geometric content of the image. The best block size is determined by searching over 2×2, 4×4, 8×8, and 16×16 block sizes.

Thresholding based on random sets. The underlying idea in the method is that each grayscale image gives rise to the distribution of a random set. Friel and Mulchanov[106] consider that each choice of threshold value gives rise to a set of binary objects with differing distance property, denoted by F_T (the foreground according to the threshold T). The distance function can be taken as Chamfer distance.[107] Thus, the expected distance function at a pixel location (i,j), $\overline{d}(i,j)$ is obtained by averaging the distance maps $d(i,j;F_T)$ for all values of the threshold values from 0 to G, or alternately by weighting them with the corresponding histogram value. Then for each value of T, the L_∞ norm of the signed difference function between the average distance map and the individual distance maps corresponding to the threshold values is calculated. Finally, the threshold is defined as that gray value that generates a foreground map most similar in their distance maps to the distance-averaged foreground.

$$T_{\mathrm{opt}}=\min\{\max_{i,j}|\overline{d}(i,j)-d(i,j;F_T)|\},$$

where $d(i,j;E_T)$, Chamfer distance to the foreground object F_T, and $\overline{d}(i,j)$ is the average distance.

2-D fuzzy partitioning. Cheng and Chen[101] (Spatial_Cheng), combine the ideas of fuzzy entropy and the 2-D histogram of the pixel values and their local 3×3 averages. Given a 2-D histogram, it is partitioned into fuzzy dark and bright regions according to the S function given also in Kaufmann.[66] The pixels x_i are assigned to A (i.e., background or foreground) according to the fuzzy rule $\mu_A(x_i)$, which in turn is characterized by the three parameters (a,b,c). To determine the best fuzzy rule, the Zadeh's fuzzy entropy formula is used,

$$H_{\mathrm{fuzzy}}(A)=-\sum_{x,y}\mu_A(x,y)p(x,y)\log p(x,y),$$

where x and y are, respectively, pixel values and pixel average values, and where A can be foreground and background events. Thus, optimum threshold is established by exhaustive searching over all permissible (a,b,c) values using the genetic algorithm to maximize the sum of foreground and background entropies, or alternatively, as the crossover point which has membership 0.5, implying the largest fuzziness. Brink[102,103] has considered the concept of spatial entropy that indirectly reflects the cooccurrence statistics. The spatial entropy is obtained using the 2-D PMF

Table 6 Thresholding functions for locally adaptive methods.

Local_Niblack[110]	$T(i,j)=m(i,j)+k.\sigma(i,j)$ where $k=-0.2$ and local window size is $b=15$
Local_Sauvola[111]	$T(i,j)=m(i,j)+\left\{1+k.\left[\frac{\sigma(i,j)}{R}-1\right]\right\}$ where $k=0.5$ and $R=128$
Local_White[112]	$B(i,j)=\begin{cases}1 & \text{if } m_{w\times w}(i,j)<I(i,j)*\text{bias}\\ 0 & \text{otherwise}\end{cases}$ where $m_{w\times w}(i,j)$ is the local mean over a $w=15$-sized window and bias$=2$.
Local_Bernsen[113]	$T(i,j)=0.5\{\max_w[I(i+m,j+n)]+\min_w[I(i+m,j+n)]\}$ where $w=31$, provided contrast $C(i,j)=I_{\text{high}}(i,j)-I_{\text{low}}(i,j)\geqslant 15$.
Local_Palumbo[21]	$B(i,j)=1$ if $I(i,j)\leqslant T_1$ or $m_{\text{neigh}}T_3+T_5>m_{\text{center}}T_4$ where $T_1=20$, $T_2=20$, $T_3=0.85$, $T_4=1.0$, $T_5=0$, neighborhood size is 3×3.
Local_Yanowitz[115]	$\lim_{n\to\infty} T_n(i,j)=T_{n-1}(i,j)+R_n(i,j)/4$ where $R_n(i,j)$ is the thinned Laplacian of the image.
Local_Kamel[1]	$B(i,j)=1$ if $\{[L(i+b,j)\wedge L(i-b,j)]\vee[L(i,j+b)\wedge L(i,j-b)]\}$ $\{[L(i+b,j+b)\wedge L(i-b,j-b)]\vee[L(i+b,j-b)\wedge L(i-b,j+b)]\}$ where $L(i,j)=\begin{cases}1 & \text{if } [m_{w\times w}(i,j)-I(i,j)]\geqslant T_0\\ 0 & \text{otherwise}\end{cases}$, $w=17$, $T_0=40$
Local_Oh[13]	Define the optimal threshold value (T_{opt}) by using a global thresholding method, such as the Kapur[53] method, then locally fine tune the pixels between $[T_0-T_1]$ considering local covariance ($T_0<T_{\text{opt}}<T_1$).
Local_Yasuda[114]	$B(i,j)=1$ if $m_{w\times w}(i,j)<T_3$ or $\sigma_{w\times w}(i,j)>T_4$ where $w=3$, $T_1=50$, $b=16$, $T_2=16$, T_3 128, $T_4=35$

$p(g,g')$, where g and g' are two gray values occurring at a lag λ, and where the spatial entropy is the sum of bivariate Shannon entropy over all possible lags.

8 Locally Adaptive Thresholding

In this class of algorithms, a threshold is calculated at each pixel, which depends on some local statistics like range, variance, or surface-fitting parameters of the pixel neighborhood. In what follows, the threshold $T(i,j)$ is indicated as a function of the coordinates (i,j) at each pixel, or if this is not possible, the object/background decisions are indicated by the logical variable $B(i,j)$. Nakagawa and Rosenfeld,[108] and Deravi and Pal,[109] were the early users of adaptive techniques for thresholding. Niblack[110] and Sauvola and Pietaksinen[111] use the local variance, while the local contrast is exploited by White and Rohrer,[112] Bernsen,[113] and Yasuda, Dubois, and Huang.[114] Palumbo, Swaminathan, and Srihari,[21] and Kamel and Zhao[1] built a center-surround scheme for determining the threshold. A surface fitted to the gray-level landscape can also be used as a local threshold, as in Yanowitz and Bruckstein,[115] and Shen and Ip.[116] See Table 6 for these methods.

Local variance methods. The method from Niblack[110] (Local_Niblack) adapts the threshold according to the local mean $m(i,j)$ and standard deviation $\sigma(i,j)$ and calculated a window size of $b\times b$. In Trier and Jain,[3] a window size of $b=15$ and a bias setting of $k=-0.2$ were found satisfactory. Sauvola and Pietaksinen's method[111] (Local_Sauvola) is an improvement on the Niblack method, especially for stained and badly illuminated documents. It adapts the contribution of the standard deviation. For example, in the case of text on a dirty or stained paper, the threshold is lowered.

Local contrast methods. White and Rohrer[112] (Local_White) compares the gray value of the pixel with the average of the gray values in some neighborhood (15×15 window suggested) about the pixel, chosen approximately to be of character size. If the pixel is significantly darker than the average, it is denoted as character; otherwise, it is classified as background. A comparison of various local adaptive methods, including White and Rohrer's, can be found in Venkateswarluh and Boyle.[117] In the local method of Bernsen[113] (Local_Bernsen), the threshold is set at the midrange value, which is the mean of the minimum $I_{\text{low}}(i,j)$ and maximum $I_{\text{high}}(i,j)$ gray values in a local window of suggested size $w=31$. However, if the contrast $C(i,j)=I_{\text{high}}(i,j)-I_{\text{low}}(i,j)$ is below a certain threshold (this contrast threshold was 15), then that neighborhood is said to consist only of one class, print or background, depending on the value of $T(i,j)$.

In Yasuda, Dubois, and Huang's method[114] (Local_Yasuda), one first expands the dynamic range of the image, followed by a nonlinear smoothing, which preserves

the sharp edges. The smoothing consists of replacing each pixel by the average of its eight neighbors, provided the local pixel range (defined as the span between the local maximum and minimum values) is below a threshold T_1. An adaptive threshold is applied, whereby any pixel value is attributed to the background (i.e., set to 255) if the local range is below a threshold T_2 or the pixel value is above the local average, both computed over $w \times w$ windows. Otherwise, the dynamic range is expanded accordingly. Finally, the image is binarized by declaring a pixel to be an object pixel if its minimum over a 3×3 window is below T_3 or its local variance is above T_4.

Center-surround schemes. Palumbo, Swaminathan, and Srihari's algorithm[21] (Local_Palumbo), based on an improvement of a method in Giuliano, Paitra, and Stringer,[118] consists in measuring the local contrast of five 3×3 neighborhoods organized in a center-surround scheme. The central 3×3 neighborhood A_{center} of the pixel is supposed to capture the foreground (background), while the four 3×3 neighborhoods, called in ensemble A_{neigh}, in diagonal positions to A_{center}, capture the background. The algorithm consists of a two-tier analysis: if $I(i,j) < T_1$, then $B(i,j) = 1$. Otherwise, one computes the average m_{neigh} of those pixels in A_{neigh} that exceed the threshold T_2, and compares it with the average m_{center} of the A_{center} pixels. The test for the remaining pixels consists of the inequality, such that, if $m_{\text{neigh}} T_3 + T_5 > m_{\text{center}} T_4$, then $B(i,j) = 1$. In Palumbo, Swaminathan, and Srihari,[21] the following threshold values have been suggested: $T_1 = 20$, $T_2 = 20$, $T_3 = 0.85$, $T_4 = 1.0$, and $T_5 = 0$.

The idea in Kamel and Zhao's[1] (Local_Kamel) method is to compare the average gray value in blocks proportional to the object width (e.g., stroke width of characters) to that of their surrounding areas. If b is the estimated stroke width, averages are calculated over a $w \times w$ window, where $w = 2b + 1$. Their approach, using the comparison operator $L(i,j)$ is somewhat similar to smoothed directional derivatives. The following parameter settings have been found appropriate: $b = 8$, and $T_0 = 40$ (T_0 is the comparison value). Recently, Yang and Yan have improved on the method of Kamel and Zhao by considering various special conditions.[119]

Surface-fitting thresholding. In Yanowitz and Bruckstein's[115] (Local_Yanowitz) method, edge, and gray level information is combined to construct a threshold surface. First, the image gradient magnitude is thinned to yield local gradient maxima. The threshold surface is constructed by interpolation with potential surface functions using a successive overrelaxation method. The threshold is obtained iteratively using the discrete Laplacian of the surface. A recent version of surface fitting by variational methods is provided by Chan, Lam, and Zhu.[120] Shen and Ip[116] used a Hopfield neural network for an active surface paradigm. There have been several other studies for local thresholding, specifically for badly illuminated images, as in Parker.[121] Other local methods involve Hadamard multiresolution analysis,[122] foreground and background clustering,[123] and joint use of horizontal and vertical derivatives.[124]

Kriging method. Oh's method (Local_Oh) is a two-pass algorithm. In the first pass, using an established nonlocal thresholding method, such as Kapur, Sahoo, and Wong,[53] the majority of the pixel population is assigned to one of the two classes (object and background). Using a variation of Kapur's technique, a lower threshold T_0 is established, below which gray values are surely assigned to class 1, e.g., an object. A second higher threshold T_1 is found, such that any pixel with gray value $g > T_1$ is assigned to class 2, i.e., the background. The remaining undetermined pixels with gray values $T_0 < g < T_1$ are left to the second pass. In the second pass, called the indicator kriging stage, these pixels are assigned to class 1 or class 2 using local covariance of the class indicators and the constrained linear regression technique called kriging, within a region with $r = 3$ pixels radius (28 pixels).

Among other local thresholding methods specifically geared to document images, one can mention the work of Kamada and Fujimoto,[125] who develop a two-stage method, the first being a global threshold, followed by a local refinement. Eikvil, Taxt, and Moen[126] consider a fast adaptive method for binarization of documents, while Pavlidis[127] uses the second-derivative of the gray level image. Zhao and Ong[128] have considered validity-guided fuzzy c-clustering to provide thresholding robust against illumination and shadow effects.

9 Thresholding Performance Criteria

Automated image thresholding encounters difficulties when the foreground object constitutes a disproportionately small (large) area of the scene, or when the object and background gray levels possess substantially overlapping distributions, even resulting in an unimodal distribution. Furthermore, the histogram can be noisy if its estimate is based on only a small sample size, or it may have a comb-like structure due to histogram stretching/equalization efforts. Consequently, misclassified pixels and shape deformations of the object may adversely affect the quality-testing task in NDT applications. On the other hand, thresholded document images may end up with noise pixels both in the background and foreground, spoiling the original character bitmaps. Thresholding may also cause character deformations such as chipping away of character strokes or conversely adding bumps and merging of characters among themselves and/or with background objects. Spurious pixels as well as shape deformations are known to critically affect the character recognition rate. Therefore, the criteria to assess thresholding algorithms must take into consideration both the noisiness of the segmentation map as well as the shape deformation of the characters, especially in the document processing applications.

To put into evidence the differing performance features of thresholding methods, we have used the following five performance criteria: misclassification error (ME), edge mismatch (EMM), relative foreground area error (RAE), modified Hausdorff distance (MHD), and region nonuniformity (NU). Obviously, these five criteria are not all inde-

pendent: for example, there is a certain amount of correlation between misclassification error and relative foreground area error, and similarly, between edge mismatch and Hausdorff distance, both of which penalize shape deformation. The region nonuniformity criterion is not based on ground-truth data, but judges the intrinsic quality of the segmented areas. We have adjusted these performance measures so that their scores vary from 0 for a totally correct segmentation to 1 for a totally erroneous case.

Misclassification error. Misclassification error (ME),[129] reflects the percentage of background pixels wrongly assigned to foreground, and conversely, foreground pixels wrongly assigned to background. For the two-class segmentation problem, ME can be simply expressed as:

$$\mathrm{ME}=1-\frac{|B_O\cap B_T|+|F_O\cap F_T|}{|B_O|+|F_O|}, \tag{5}$$

where B_O and F_O denote the background and foreground of the original (ground-truth) image, B_T and F_T denote the background and foreground area pixels in the test image, and $|.|$ is the cardinality of the set. The ME varies from 0 for a perfectly classified image to 1 for a totally wrongly binarized image.

Edge mismatch. This metric penalizes discrepancies between the edge map of the gray level image and the edge map obtained from the thresholded image. The edge mismatch metric is expressed as:[6]

$$\mathrm{EMM}=1-\frac{\mathrm{CE}}{\mathrm{CE}+\omega[\Sigma_{k\in\{\mathrm{EO}\}}\delta(k)+\alpha\Sigma_{k\in\{\mathrm{ET}\}}\delta(k)]}, \quad \text{with}$$

$$\delta(k)=\begin{cases}|d_k| & \text{if } |d_k|<\mathrm{maxdist}\\ D_{\max} & \text{otherwise}\end{cases}, \tag{6}$$

where CE is the number of common edge pixels found between the ground-truth image and the thresholded image, EO is the set of excess ground-truth edge pixels missing in the threshold image, ET is the set of excess thresholded edge pixels not taking place in the ground truth, ω is the penalty associated with an excess original edge pixel, and finally α is the ratio of the penalties associated with an excess threshold edge pixel to an excess original edge pixel. Here d_k denotes the Euclidean distance of the k'th excess edge pixel to a complementary edge pixel within a search area determined by the maxdist parameter. It has been suggested[6] to select the parameter as maxdist $=0.025N$, where N is the image dimension, $D_{\max}=0.1N$, $\omega=10/N$, and $\alpha=2$.

Region nonuniformity. This measure,[130,131] which does not require ground-truth information, is defined as

$$\mathrm{NU}=\frac{|F_T|}{|F_T+B_T|}\frac{\sigma_f^2}{\sigma^2}, \tag{7}$$

where σ^2 represents the variance of the whole image, and σ_f^2 represents the foreground variance. It is expected that a well-segmented image will have a nonuniformity measure close to 0, while the worst case of NU$=1$ corresponds to an image for which background and foreground are indistinguishable up to second order moments.

Relative foreground area error. The comparison of object properties such as area and shape, as obtained from the segmented image with respect to the reference image, has been used in Zhang[31] under the name of relative ultimate measurement accuracy (RUMA) to reflect the feature measurement accuracy. We modified this measure for the area feature A as follows:

$$\mathrm{RAE}=\begin{cases}\dfrac{A_0-A_T}{A_0} & \text{if } A_T<A_0\\ \dfrac{A_T-A_0}{A_T} & \text{if } A_T\leqslant A_0\end{cases}, \tag{8}$$

where A_0 is the area of reference image, and A_T is the area of thresholded image. Obviously, for a perfect match of the segmented regions, RAE is zero, while if there is zero overlap of the object areas, the penalty is the maximum one.

Shape distortion penalty via Hausdorff distance. The Hausdorff distance can be used to assess the shape similarity of the thresholded regions to the ground-truth shapes. Recall that, given two finite sets of points, say ground-truth and thresholded foreground regions, their Hausdorff distance is defined as

$$H(F_O,F_T)=\max\{d_H(F_O,F_T),d_H(F_T,F_O)\},$$

$$\text{where } d_H(F_O,F_T)=\max_{f_O\in F_O} d(f_O,F_T)$$

$$=\max_{f_O\in F_O f_T\in F_T}\min\|f_O-f_T\|,$$

and $\|f_O-f_T\|$ denotes the Euclidean distance of two pixels in the ground-truth and thresholded objects.

Since the maximum distance is sensitive to outliers, we have measured the shape distortion via the average of the modified Hausdorff distances (MHD)[132] over all objects. The modified distance is defined as:

$$\mathrm{MHD}(F_O,F_T)=\frac{1}{|F_O|}\sum_{f_O\in F_O} d(f_O,F_T). \tag{9}$$

For example, the MHDs are calculated for each 19 $\times$19 pixel character box, and then the MHDs are averaged over all characters in a document. Notice that, since an upper bound for the Hausdorff distance cannot be established, the MHD metric could not be normalized to the interval [0, 1], and hence is treated separately (by dividing each MHD value to the highest MHD value over the test image set NMHD).

Combination of measures. To obtain an average performance score from the previous five criteria, we have considered two methods. The first method was the arithmetic averaging of the normalized scores obtained from the ME, EMM, NU, NMHD, and RAE criteria. In other words,

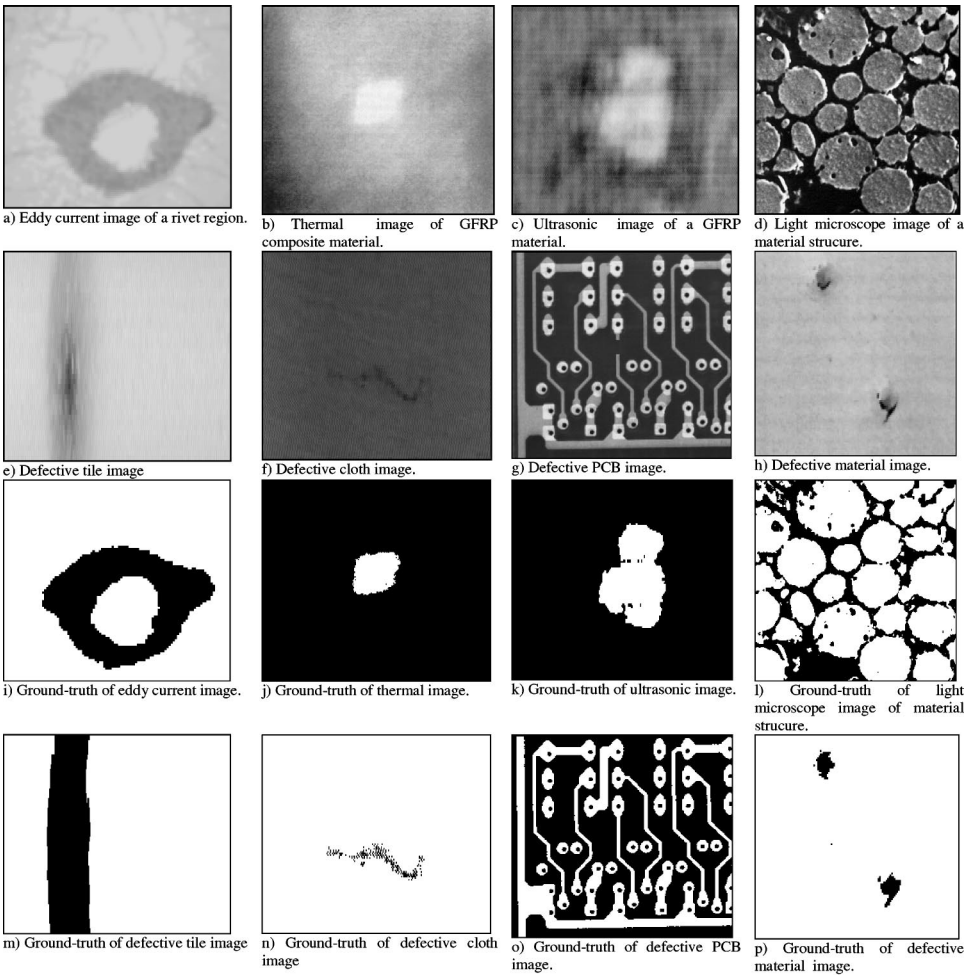

Fig. 2 Sample NDT images and their ground truths.

given a thresholding algorithm, for each image the average of ME, EMM, NU, and RAE was an indication of its segmentation quality. In turn, the sum of these image quality scores determines the performance of the algorithm. In the second method, we used rank averaging, so that, for each test image, we ranked the thresholding algorithms from 1 to 40 according to each criterion separately. Then the ranks (not the actual scores) were averaged over both the images and the five criteria ME, EMM, NU, RAE, and NMHD. A variation of this scheme could be to first take the arithmetic average over the criteria for each image, then rank the thresholding methods, and finally average the ranks. We experimented with the previous two score averaging methods (arithmetic and rank averaging of quality measures), and we found that they resulted in fairly similar ranking of the thresholding algorithms. Consequently, we chose the arithmetic averaging method, as it was more straightforward. Thus the performance measure for the i'th image is written in terms of the scores of the five metrics as:

$$S(i)=[\mathrm{ME}(i)+\mathrm{EMM}(i)+\mathrm{NU}(i)+\mathrm{RAE}(i)+\mathrm{NMHD}(i)]/5. \quad (10)$$

10 Dataset and Experimental Results

10.1 *Dataset*

Our test data consisted of a variety of 40 NDT and 40 document images.

Nondestructive testing images. The variety of NDT we considered consisted of eight eddy current, for thermal, two ultrasonic, six light microscope, four ceramic, six material, two PCB, and eight cloth images. A few samples of the NDT image set are given in Fig. 2.

Eddy current image inspection is frequently used for the detection of invisible small cracks and defects of different materials including aircraft fuselages. A defective eddy current image and its ground-truth segmentation map are illustrated in Figs. 2(a) and 2(i), where the dark region, representing the rivet surroundings, must be circular in a healthy case. Infrared thermographs are used, among other applications, for surface defect detection in aircraft materials, such as carbon fiber reinforced composites. A defective thermal image specimen and its segmentation ground truth are illustrated in Figs. 2(b) and 2(j). An ultrasonic image of a defective glass-fiber reinforced plastics (GFRP) image and

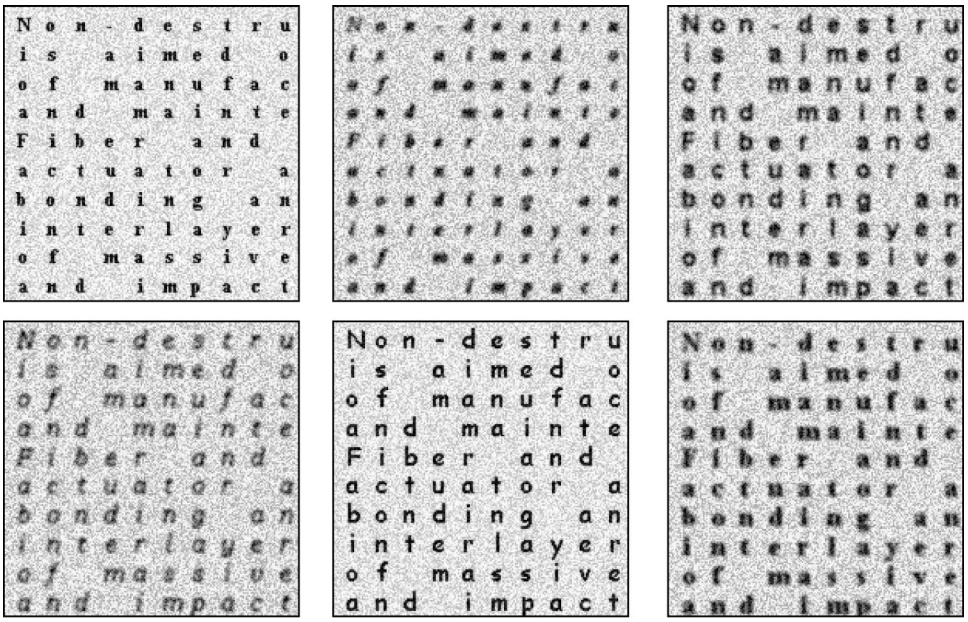

Fig. 3 Sample degraded document images.

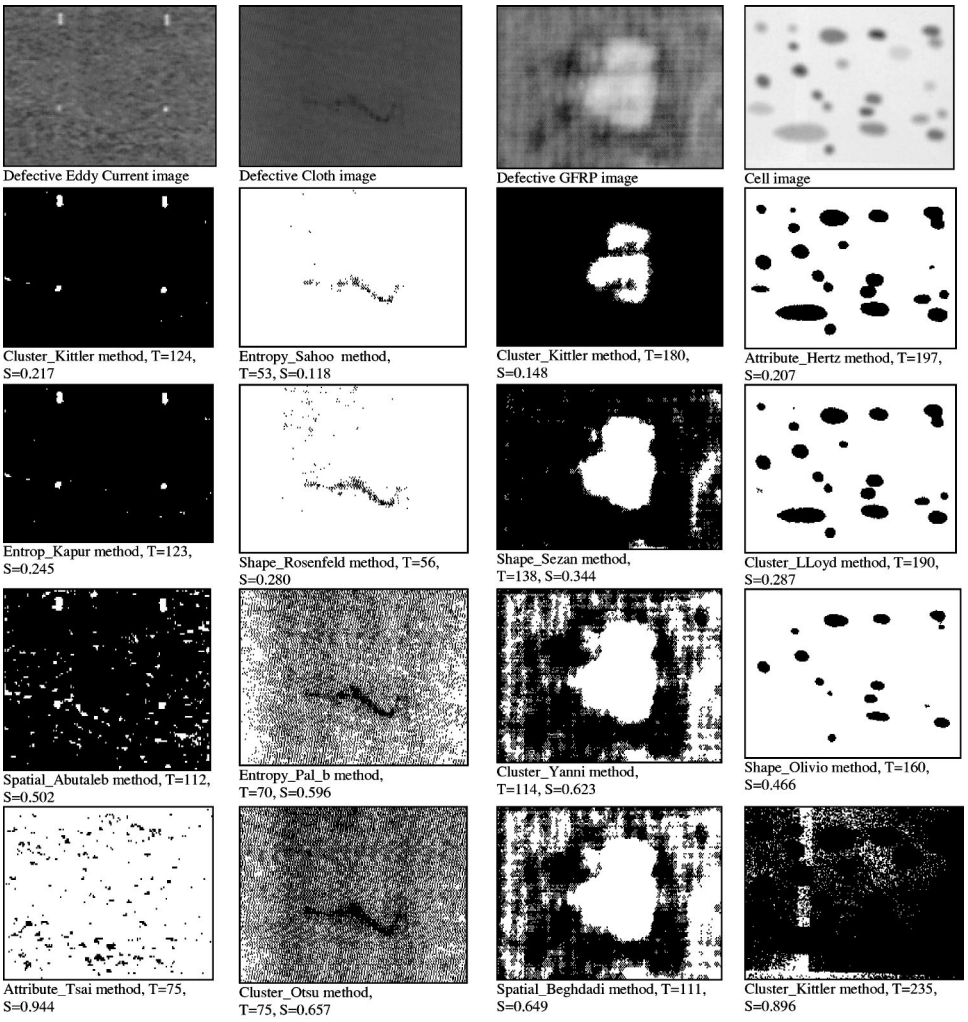

Fig. 4 Thresholding results of sample NDT images.

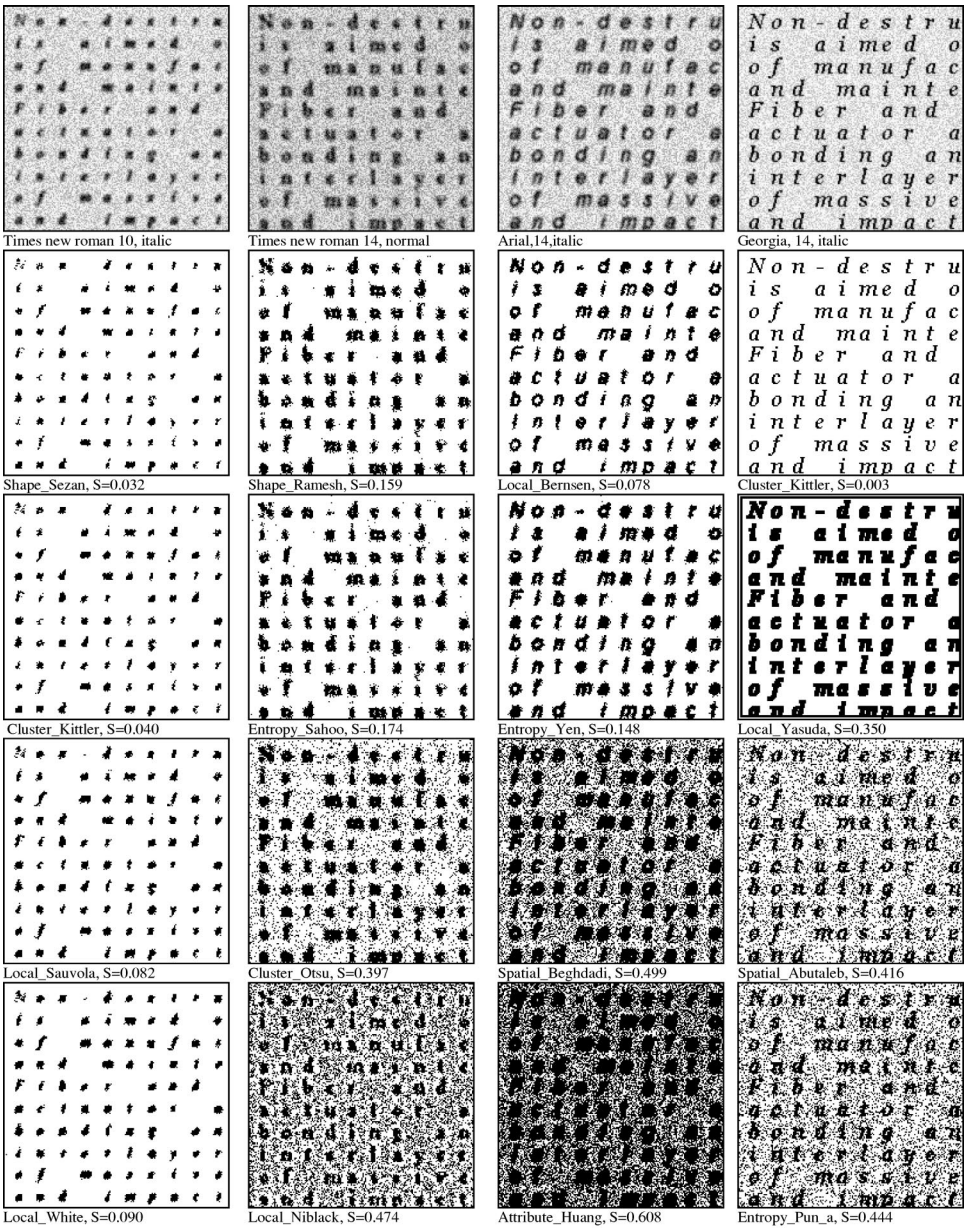

Fig. 5 Thresholding results of sample document images.

its ground truth are given in Figs. 2(c) and 2(k). In material science applications, the microstructure of materials is frequently inspected by light microscopy. These observations yield information on the phases of the material, as well as on porosity, particle sizes, uniformity of distribution, etc. A sample image of light microscopy and its ground-truth segmentation map images are displayed in Figs. 2(d) and 2(l). Tile and cloth quality testing application images and their ground truths are given in Figs. 2(e), 2(m), 2(f), and 2(n), respectively. Visual inspection of PCB boards, as practiced today on the production line, is a tedious and error-prone task. Computer-vision-based automatic inspection schemes are increasingly being deployed,[133] where the first processing stage is again thresholding. The thresholding and subsequent processing aims to put into evidence such defects as broken lines, undrilled vias, etc. An example PCB image and ground truth are given in Figs. 2(g) and 2(c). Finally a defective fuselage material image and its ground truth are given in Figs. 2(h) and 2(p).

Document image applications. 40 documents containing ground-truth character images were created with different fonts (times new roman, arial, comics, etc.), sizes (10,12,14), and typefaces (normal, bold, italic, etc.). Furthermore, to simulate more realistic documents such as the effects of the poor quality of paper, photocopied and faxed documents, etc., degraded documents were set using Baird[134] degradation models. The simulated document defects were blur and speckle noise since, among the several other defects proposed in Baird,[134] these two had a direct bearing on thresholding. Three levels of document degradation, namely light, medium, and poor, were used. Sample degraded document images (a small part of real images) are

Table 7 Thresholding evaluation ranking of 40 NDT images according to the overall average quality score.

Rank	Method	Average score (AVE)	Rank	Method	Average score ($\bar{S}$)
1	Cluster_Kittler	0.256	21	Shape_Ramesh	0.460
2	Entropy_Kapur	0.261	22	Spatial_Cheng	0.481
3	Entropy_Sahoo	0.269	23	Attribute_Tsai	0.484
4	Entropy_Yen	0.289	24	Local_Bernsen	0.550
5	Cluster_Lloyd	0.292	25	Spatial_Pal_a	0.554
6	Cluster_Otsu	0.318	26	Local_Yasuda	0.573
7	Cluster_Yanni	0.328	27	Local_Palumbo	0.587
8	Local_Yanowitz	0.339	28	Entropy_Sun	0.588
9	Attribute_Hertz	0.351	29	Attribute_Leung	0.590
10	Entropy_Li	0.364	30	Entropy_Pun_a	0.591
11	Spatial_Abutaleb	0.370	31	Spatial_Beghdadi	0.619
12	Attribute_Pikaz	0.383	32	Local_Oh	0.619
13	Shape_Guo	0.391	33	Local_Niblack	0.638
14	Cluster_Ridler	0.401	34	Spatial_Pal_b	0.642
15	Cluster_Jawahar_b	0.423	35	Entropy_Pun_b	0.665
16	Attribute_Huang	0.427	36	Local_White	0.665
17	Shape_Sezan	0.431	37	Local_Kamel	0.697
18	Entropy_Shanbag	0.433	38	Local_Sauvola	0.707
19	Shape_Rosenfeld	0.442	39	Cluster_Jawahar_a	0.735
20	Shape_Olivio	0.458	40	Entropy_Brink	0.753

Table 8 Thresholding evaluation ranking of 40 degraded document images according to the overall average quality score.

Rank	Method	Average score (AV)	Rank	Method	Average score ($\bar{S}$)
1	Cluster_Kittler	0.046	21	Cluster_Yanni	0.300
2	Local_Sauvola	0.066	22	Attribute_Tsai	0.308
3	Local_White	0.08	23	Attribute_Hertz	0.317
4	Local_Bernsen	0.09	24	Spatial_Cheng	0.320
5	Shape_Ramesh	0.093	25	Local_Yasuda	0.336
6	Attribute_Leung	0.110	26	Entropy_Sun	0.39
7	Entropy_Li	0.114	27	Local_Kamel	0.391
8	Cluster_Ridler	0.136	28	Entropy_Pun_a	0.463
9	Entropy_Shanbag	0.144	29	Local_Niblack	0.475
10	Shape_Sezan	0.145	30	Local_Oh	0.514
11	Entropy_Shaoo	0.148	31	Spatial_Abutaleb	0.515
12	Entropy_Kapur	0.149	32	Spatial_Pal_a	0.533
13	Entropy_Yen	0.156	33	Spatial_Beghdadi	0.539
14	Entropy_Brink	0.17	34	Attribute_Huang	0.566
15	Cluster_Lloyd	0.18	35	Entropy_Pun_b	0.593
16	Local_Palumbo	0.195	36	Shape_Guo	0.596
17	Cluster_Otsu	0.197	37	Spatial_Pal_b	0.605
18	Cluster_Jawahar_b	0.251	38	Shape_Rosenfeld	0.663
19	Attribute_Pikaz	0.259	39	Shape_Olivio	0.711
20	Local_Yanowitz	0.288	40	Cluster_Jawahar_a	0.743

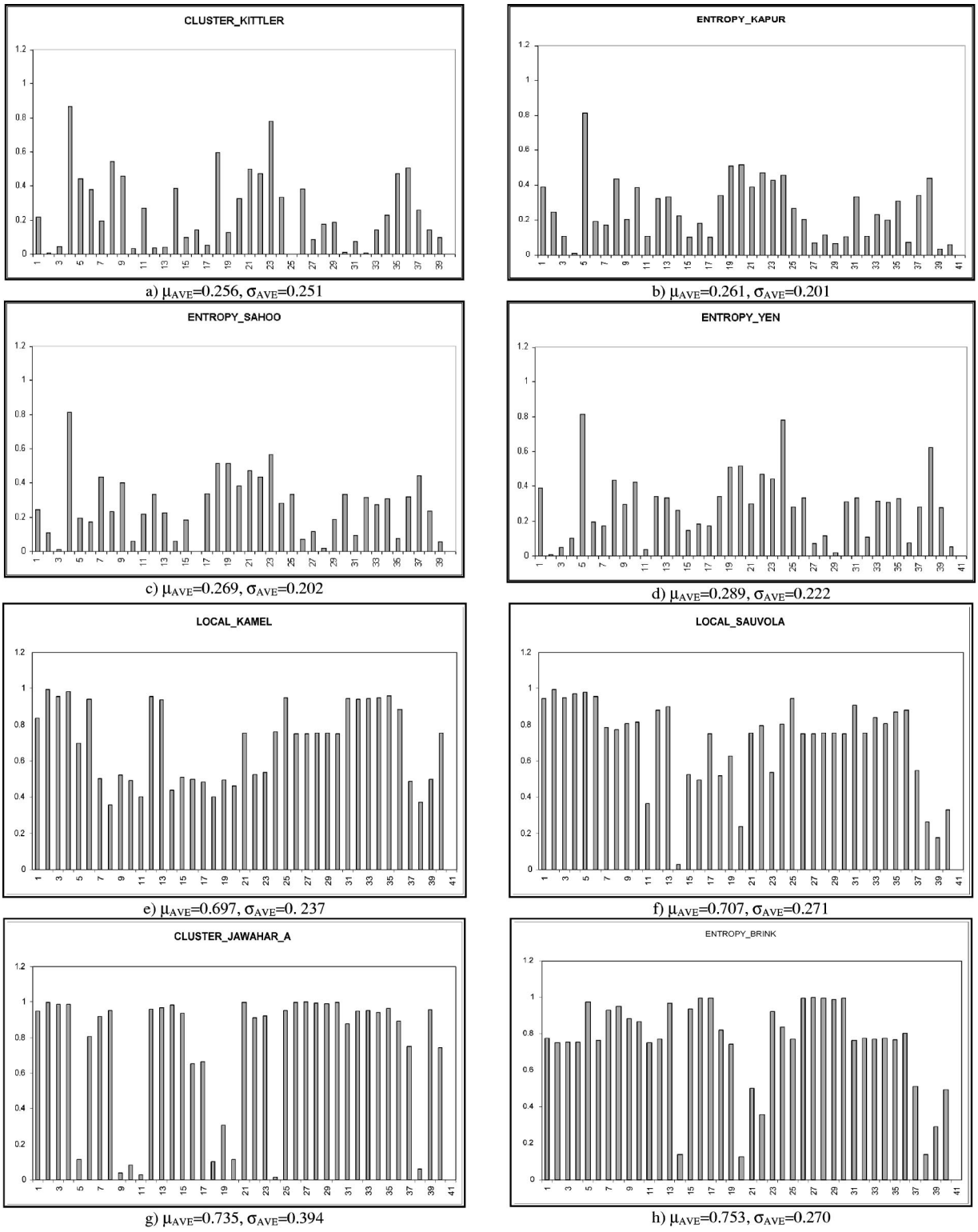

Fig. 6 Sample performance score distributions of best and worst thresholding methods over the NDT images.

shown in Fig. 3. The blur was modeled by a circularly symmetric Gaussian filter, representing the point spread function, with standard error of blur in units of output pixel size. The blur size was chosen, in pixel units, as 5×5.

10.2 *Experimental Results*

Representative visual results and their quantitative scores are given in Fig. 4 for NDT images and in Fig. 5 for document images. The overall evaluation result of NDT images

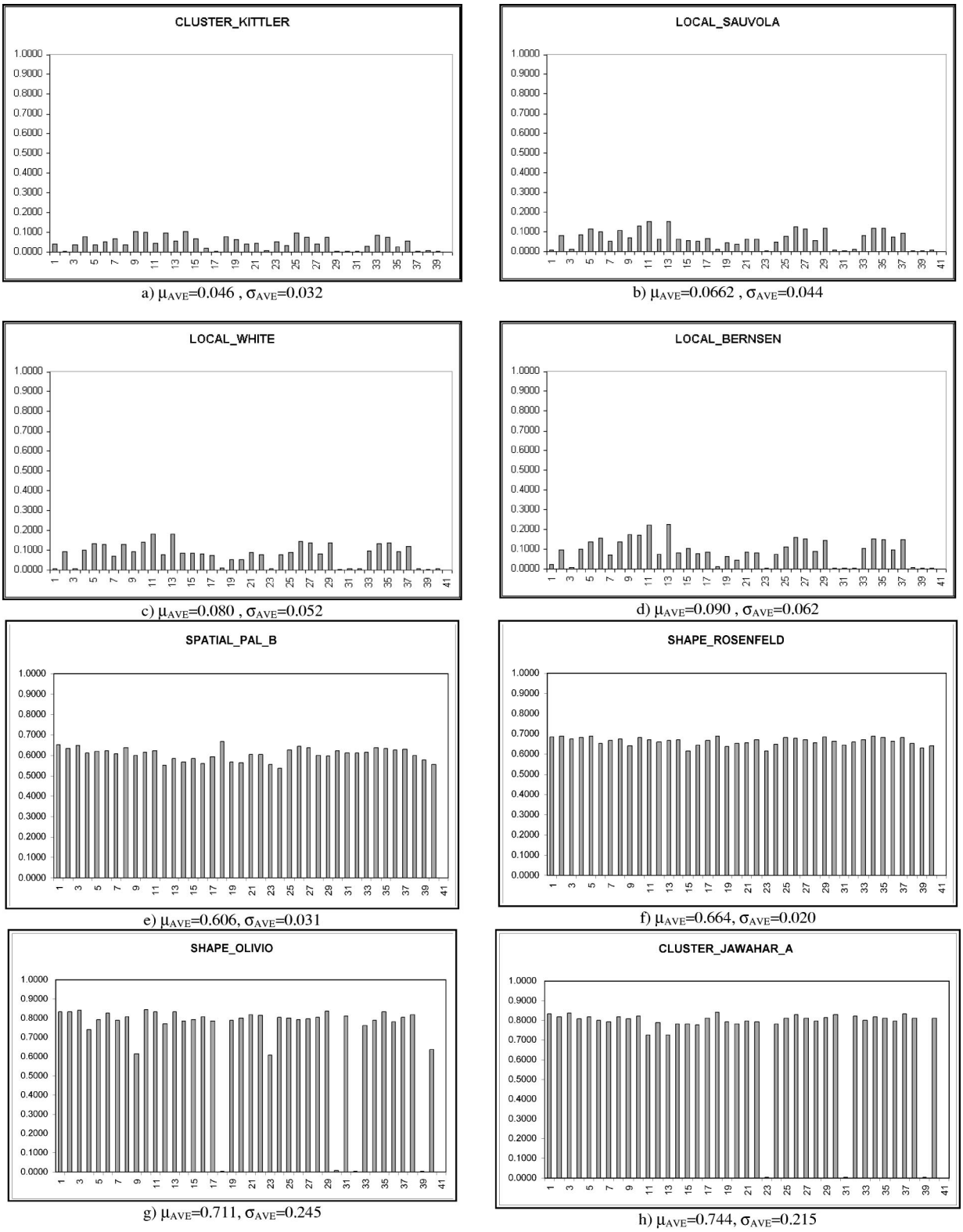

Fig. 7 Sample performance score distributions of best and worst thresholding methods over the document images.

is given in Table 7, while the performance scores for document images are shown in Table 8. In these tables, the final score of each thresholding method is calculated by taking the average of the ME, EMM, NU, NMHD, and RAE measures over 40 images. It is interesting to point out the following.

- As it was conjectured, the rankings of the thresholding

algorithms differ under NDT and document applications. In fact, only the Cluster_Kittler algorithm was common to both tables, if we only consider the top seven.

- For NDT applications, the highest-ranking seven techniques are all from the clustering and entropy category. These scores reflect also our subjective evaluation on the plausibility of the extracted object.
- For document applications, techniques from locally adaptive and shape categories appear in the top seven. These results are consistent with those of the Trier and Jain[3] study.

The performance variability of the algorithms from case to case are illustrated in Figs. 6 and 7. Here the horizontal axis denotes the index of the 40 images, and the vertical axis denotes the corresponding average score. Notice that all methods invariably performed poorly for at least one or two images.[135] Thus it was observed that any single algorithm could not be successful for all image types, even in a single application domain. To obtain the robustness of the thresholding method, we explored the combination of more than one thresholding algorithm based on the conjecture that they could be complementary to each other. The combination of thresholding algorithms can be done at the feature level or at the decision level. At the feature level, one could apply, for example, some averaging operation on the Topt values obtained from individual algorithms; on the decision level, one could do fusion of the foreground-background decisions, for example, by taking the majority decision. We could not obtain better results than the Kittler method.

11 Conclusions

In this study, we make a categorized survey of image thresholding methods. Furthermore, we select a subset of 40 bilevel image thresholding methods, which have been implemented and for which the thresholding formulas have been expressed in a streamlined fashion. Quantitative evaluation scores have been obtained using a database of 40 NDT and 40 document images. We have observed that the clustering-based method of Kittler and Illingworth[43] and the entropy-based methods of Kapur, Sahoo, and Wong,[53] and Sahoo, Wilkins, and Yeager,[55] are, in that order, the best performing thresholding algorithms in the case of NDT images. Similarly, the clustering-based method of Kittler and Illingworth[43] and the local-based methods of Sauvola and Pietaksinen,[111] and of White and Rohrer,[112] are, in that order, the best performing document binarization algorithms. Note, however, that these results apply only to text document images degraded with noise and blur. On the other hand, documents with patterned backgrounds like checks, documents degraded by nonuniform illumination and shadow, or mixed-type documents[85,136–138] are outside the scope of this comparison.

Several other issues remain to be addressed. For example, the increasing number of color documents becomes a new challenge for binarization and segmentation.[139,140] Multilevel thresholding, or simply multithresholding, is gaining more relevance in vision applications. A few authors have addressed this issue.[5,12,37,54,67,73,138,141] The much needed performance comparison of multithresholding algorithms would necessitate reformulation of the performance measures outlined in Sec. 9. To obtain better thresholding results, one should consider application-specific information, for example, an expected foreground region, whether it is on a dark or bright side, for guidance.

One should keep on eye on the fact that thresholding should be opted for two-class segmentation problems due to their simplicity whenever they achieve performance similar to more sophisticated methods, like Bayesian schemes and random Markov models.

Acknowledgment

We would like to thank the Tübitak Marmara Research Center for supporting the study.

References

1. M. Kamel and A. Zhao, ''Extraction of binary character/graphics images from grayscale document images,'' *Graph. Models Image Process.* **55**(3), 203–217 (1993).
2. T. Abak, U. Barış, and B. Sankur, ''The performance of thresholding algorithms for optical character recognition,'' *Intl. Conf. Document Anal. Recog. ICDAR'97*, pp. 697–700 (1997).
3. O. D. Trier and A. K. Jain, ''Goal-directed evaluation of binarization methods,'' *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-17**, 1191–1201 (1995).
4. B. Bhanu, ''Automatic target recognition: state of the art survey,'' *IEEE Trans. Aerosp. Electron. Syst.* **AES-22**, 364–379 (1986).
5. M. Sezgin and R. Tasaltin, ''A new dichotomization technique to multilevel thresholding devoted to inspection applications,'' *Pattern Recogn. Lett.* **21**, 151–161 (2000).
6. M. Sezgin and B. Sankur, ''Comparison of thresholding methods for non-destructive testing applications,'' *IEEE ICIP'2001, Intl. Conf. Image Process.*, pp. 764–767 (2001).
7. J. C. Russ, ''Automatic discrimination of features in gray-scale images,'' *J. Microsc.* **148**(3), 263–277 (1987).
8. M. E. Sieracki, S. E. Reichenbach, and K. L. Webb, ''Evaluation of automated threshold selection methods for accurately sizing microscopic fluorescent cells by image analysis,'' *Appl. Environ. Microbiol.* **55**, 2762–2772 (1989).
9. P. Bock, R. Klinnert, R. Kober, R. M. Rovner, and H. Schmidt, ''Gray-scale ALIAS,'' *IEEE Trans. Knowl. Data Eng.* **4**, 109–122 (1992).
10. L. U. Wu, M. A. Songde, and L. U. Hanqing, ''An effective entropic thresholding for ultrasonic imaging,'' *ICPR'98: Intl. Conf. Patt. Recog.*, pp. 1522–1524 (1998).
11. J. Moysan, G. Corneloup, and T. Sollier, ''Adapting an ultrasonic image threshold method to eddy current images and defining a validation domain of the thresholding method,'' *NDT & E Intl.* **32**, 79–84 (1999).
12. J. S. Chang, H. Y. M. Liao, M. K. Hor, J. W. Hsieh, and M. Y. Chern, ''New automatic multi-level thresholding technique for segmentation of thermal images,'' *Image Vis. Comput.* **15**, 23–34 (1997).
13. W. Oh and B. Lindquist, ''Image thresholding by indicator kriging,'' *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-21**, 590–602 (1999).
14. T. Srikanthan and K. V. Asari, ''Automatic segmentation algorithm for the extraction of lumen region and boundary from endoscopic images,'' *Med. Biol. Eng. Comput.* **39**(1), 8–14 (2001).
15. S. Venkatesh and P. L. Rosin, ''Dynamic threshold determination by local and global edge evaluation,'' *CVGIP: Graph. Models Image Process.* **57**, 146–160 (1995).
16. R. Kohler, ''A segmentation system based on thresholding,'' *Graph. Models Image Process.* **15**, 319–338 (1981).
17. A. Perez and T. Pavlidis, ''An iterative thresholding algorithm for image segmentation,'' *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-9**, 742–75 (1987).
18. J. Fan, J. Yu, G. Fujita, T. Onoye, L. Wu, and I. Shirakawa, ''Spatiotemporal segmentation for compact video representation,'' *Signal Process. Image Commun.* **16**, 553–566 (2001).
19. S. U. Le, S. Y. Chung, and R. H. Park, ''A comparative performance study of several global thresholding techniques for segmentation,'' *Graph. Models Image Process.* **52**, 171–190 (1990).
20. J. S. Weszka and A. Rosenfeld, ''Threshold evaluation techniques,'' *IEEE Trans. Syst. Man Cybern.* **SMC-8**, 627–629 (1978).
21. P. W. Palumbo, P. Swaminathan, and S. N. Srihari, ''Document *image* binarization: Evaluation of algorithms,'' *Proc. SPIE* **697**, 278–286 (1986).
22. P. K. Sahoo, S. Soltani, A. K. C. Wong, and Y. Chen, ''A survey of thresholding techniques,'' *Comput. Graph. Image Process.* **41**, 233–260 (1988).
23. C. A. Glasbey, ''An analysis of histogram-based thresholding algo-

rithms,'' *Graph. Models Image Process.* **55**, 532–537 (1993).
24. A. Rosenfeld and P. De la Torre, ''Histogram concavity analysis as an aid in threshold selection,'' *IEEE Trans. Syst. Man Cybern.* **SMC-13**, 231–235 (1983).
25. J. Weszka and A. Rosenfeld, ''Histogram modification for threshold selection,'' *IEEE Trans. Syst. Man Cybern.* **SMC-9**, 38–52 (1979).
26. L. Halada and G. A. Osokov, ''Histogram concavity analysis by quasicurvature,'' *Comp. Artif. Intell.* **6**, 523–533 (1987).
27. S. C. Sahasrabudhe and K. S. D. Gupta, ''A valley-seeking threshold selection technique,'' *Comput. Vis. Image Underst.* **56**, 55–65 (1992).
28. R. Guo and S. M. Pandit, ''Automatic threshold selection based on histogram modes and a discriminant criterion,'' *Mach. Vision Appl.* **10**, 331–338 (1998).
29. J. Cai and Z. Q. Liu, ''A new thresholding algorithm based on all-pole model,'' *ICPR'98, Intl. Conf. Patt. Recog.*, pp. 34–36 (1998).
30. N. Ramesh, J. H. Yoo, and I. K. Sethi, ''Thresholding based on histogram approximation,'' *IEE Proc. Vision Image Signal Process.* **142**(5), 271–279 (1995).
31. T. Kampke and R. Kober, ''Nonparametric optimal binarization,'' *ICPR'98, Intl. Conf. Patt. Recog.*, pp. 27–29 (1998).
32. M. I. Sezan, ''A peak detection algorithm and its application to histogram-based image data reduction,'' *Graph. Models Image Process.* **29**, 47–59 (1985).
33. M. J. Carlotto, ''Histogram analysis using a scale-space approach,'' *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-9**, 121–129 (1997).
34. J. C. Olivo, ''Automatic threshold selection using the wavelet transform,'' *Graph. Models Image Process.* **56**, 205–218 (1994).
35. R. J. Whatmough, ''Automatic threshold selection from a histogram using the exponential hull,'' *Graph. Models Image Process.* **53**, 592–600 (1991).
36. S. Boukharouba, J. M. Rebordao, and P. L. Wendel, ''An amplitude segmentation method based on the distribution function of an image,'' *Graph. Models Image Process.* **29**, 47–59 (1985).
37. D. M. Tsai, ''A fast thresholding selection procedure for multimodal and unimodal histograms,'' *Pattern Recogn. Lett.* **16**, 653–666 (1995).
38. T. W. Ridler and S. Calvard, ''Picture thresholding using an iterative selection method,'' *IEEE Trans. Syst. Man Cybern.* **SMC-8**, 630–632 (1978).
39. C. K. Leung and F. K. Lam, ''Performance analysis of a class of iterative image thresholding algorithms,'' *Pattern Recogn.* **29**(9), 1523–1530 (1996).
40. H. J. Trussel, ''Comments on picture thresholding using iterative selection method,'' *IEEE Trans. Syst. Man Cybern.* **SMC-9**, 311 (1979).
41. M. K. Yanni and E. Horne, ''A new approach to dynamic thresholding,'' *EUSIPCO'94: 9th European Conf. Sig. Process.* **1**, 34–44 (1994).
42. D. E. Lloyd, ''Automatic target classification using moment invariant of image shapes,'' Technical Report, RAE IDN AW126, Farnborough, UK (Dec. 1985).
43. J. Kittler and J. Illingworth, ''Minimum error thresholding,'' *Pattern Recogn.* **19**, 41–47 (1986).
44. S. Cho, R. Haralick, and S. Yi, ''Improvement of Kittler and Illingworths's minimum error thresholding,'' *Pattern Recogn.* **22**, 609–617 (1989).
45. J. Kittler and J. Illingworth, ''On threshold selection using clustering criteria,'' *IEEE Trans. Syst. Man Cybern.* **SMC-15**, 652–655 (1985).
46. N. Otsu, ''A threshold selection method from gray level histograms,'' *IEEE Trans. Syst. Man Cybern.* **SMC-9**, 62–66 (1979).
47. C. V. Jawahar, P. K. Biswas, and A. K. Ray, ''Investigations on fuzzy thresholding based on fuzzy clustering,'' *Pattern Recogn.* **30**(10), 1605–1613 (1997).
48. F. R. D. Velasco, ''Thresholding using the isodata clustering algorithm,'' *IEEE Trans. Syst. Man Cybern.* **SMC-10**, 771–774 (1980).
49. H. Lee and R. H. Park, ''Comments on an optimal threshold scheme for image segmentation,'' *IEEE Trans. Syst. Man Cybern.* **SMC-20**, 741–742 (1990).
50. J. Z. Liu and W. Q. Li, ''The automatic thresholding of gray-level pictures via two-dimensional Otsu method,'' *Acta Automatica Sin.* **19**, 101–105 (1993).
51. T. Pun, ''A new method for gray-level picture threshold using the entropy of the histogram,'' *Signal Process.* **2**(3), 223–237 (1980).
52. T. Pun, ''Entropic thresholding: A new approach,'' *Comput. Graph. Image Process.* **16**, 210–239 (1981).
53. 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,'' *Graph. Models Image Process.* **29**, 273–285 (1985).
54. J. C. Yen, F. J. Chang, and S. Chang, ''A new criterion for automatic multilevel thresholding,'' *IEEE Trans. Image Process.* **IP-4**, 370–378 (1995).
55. P. Sahoo, C. Wilkins, and J. Yeager, ''Threshold selection using Renyi's entropy,'' *Pattern Recogn.* **30**, 71–84 (1997).
56. C. H. Li and C. K. Lee, ''Minimum cross-entropy thresholding,'' *Pattern Recogn.* **26**, 617–625 (1993).
57. C. H. Li and P. K. S. Tam, ''An iterative algorithm for minimum cross-entropy thresholding,'' *Pattern Recogn. Lett.* **19**, 771–776 (1998).
58. A. D. Brink and N. E. Pendock, ''Minimum cross entropy threshold selection,'' *Pattern Recogn.* **29**, 179–188 (1996).
59. N. R. Pal, ''On minimum cross-entropy thresholding,'' *Pattern Recogn.* **29**(4), 575–580 (1996).
60. A. G. Shanbag, ''Utilization of information measure as a means of image thresholding,'' *Comput. Vis. Graph. Image Process.* **56**, 414–419 (1994).
61. H. D. Cheng, Y. H. Chen, and Y. Sun, ''A novel fuzzy entropy approach to image enhancement and thresholding,'' *Signal Process.* **75**, 277–301 (1999).
62. G. Johannsen and J. Bille, ''A threshold selection method using information measures,'' *ICPR'82: Proc. 6th Intl. Conf. Patt. Recog.*, pp. 140–143 (1982).
63. S. K. Pal, R. A. King, and A. A. Hashim, ''Automatic gray level thresholding through index of fuzziness and entropy,'' *Pattern Recogn. Lett.* **1**, 141–146 (1980).
64. J. E. Shore and R. W. Johnson, ''Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy,'' *IEEE Trans. Inf. Theory* **IT-26**, 26–37 (1980).
65. A. K. C. Wong and P. K. Sahoo, ''A gray-level threshold selection method based on maximum entropy principle,'' *IEEE Trans. Syst. Man Cybern.* **SMC-19**, 866–871 (1989).
66. A. Kaufmann, *Introduction to the Theory of Fuzzy Sets: Fundamental Theoretical Elements*, Vol. 1, Academic Press, New York (1980).
67. L. Hertz and R. W. Schafer, ''Multilevel thresholding using edge matching,'' *Comput. Vis. Graph. Image Process.* **44**, 279–295 (1988).
68. S. K. Pal and A. Rosenfeld, ''Image enhancement and thresholding by optimization of fuzzy compactness,'' *Pattern Recogn. Lett.* **7**, 77–86 (1988).
69. A. Rosenfeld, ''The fuzzy geometry of image subsets,'' *Pattern Recogn. Lett.* **2**, 311–317 (1984).
70. W. H. Tsai, ''Moment-preserving thresholding: A new approach,'' *Graph. Models Image Process.* **19**, 377–393 (1985).
71. S. C. Cheng and W. H. Tsai, ''A neural network approach of the moment-preserving technique and its application to thresholding,'' *IEEE Trans. Comput.* **C-42**, 501–507 (1993).
72. E. J. Delp and O. R. Mitchell, ''Moment-preserving quantization,'' *IEEE Trans. Commun.* **39**, 1549–1558 (1991).
73. L. O'Gorman, ''Binarization and multithresholding of document images using connectivity,'' *Graph. Models Image Process.* **56**, 494–506 (1994).
74. Y. Liu and S. N. Srihari, ''Document image binarization based on texture analysis,'' *Proc. SPIE* **2181**, 254–263 (1994).
75. Y. Liu, R. Fenrich, and S. N. Srihari, ''An object attribute thresholding algorithm for document image binarization,'' *ICDAR'93: Proc. 2nd Intl. Conf. Document Anal. Recog.*, pp. 278–281 (1993).
76. A. Pikaz and A. Averbuch, ''Digital image thresholding based on topological stable state,'' *Pattern Recogn.* **29**, 829–843 (1996).
77. C. A. Murthy and S. K. Pal, ''Fuzzy thresholding: A mathematical framework, bound functions and weighted moving average technique,'' *Pattern Recogn. Lett.* **11**, 197–206 (1990).
78. R. Yager, ''On the measure of fuzziness and negation. Part I: Membership in the unit interval,'' *Int. J. Gen. Syst.* **5**, 221–229 (1979).
79. K. Ramar, S. Arunigam, S. N. Sivanandam, L. Ganesan, and D. Manimegalai, ''Quantitative fuzzy measures for threshold selection,'' *Pattern Recogn. Lett.* **21**, 1–7 (2000).
80. X. Fernandez, ''Implicit model oriented optimal thresholding using Kolmogorov-Smirnov similarity measure,'' *ICPR'2000: Intl. Conf. Patt. Recog.*, pp. 466–469, Barcelona (2000).
81. C. K. Leung and F. K. Lam, ''Maximum segmented image information thresholding,'' *Graph. Models Image Process.* **60**, 57–76 (1998).
82. L. K. Huang and M. J. J. Wang, ''Image thresholding by minimizing the measures of fuzziness,'' *Pattern Recogn.* **28**, 41–51 (1995).
83. H. S. Don, ''A noise attribute thresholding method for document image binarization,'' *IEEE Conf. Image Process.*, pp. 231–234 (1995).
84. S. Guo, ''A new threshold method based on morphology and fourth order central moments,'' *Proc. SPIE* **3545**, 317–320 (1998).
85. Y. Solihin and C. G. Leedham, ''Integral ratio: A new class of global thresholding techniques for handwriting images,'' *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-21**, 761–768 (1999).
86. Z. Aviad and E. Lozinskii, ''Semantic thresholding,'' *Pattern Recogn. Lett.* **5**, 321–328 (1987).
87. G. Gallo and S. Spinello, ''Thresholding and fast iso-contour extraction with fuzzy arithmetic,'' *Pattern Recogn. Lett.* **21**, 31–44 (2000).
88. R. L. Kirby and A. Rosenfeld, ''A note on the use of (gray level, local average gray level) space as an aid in threshold selection,'' *IEEE Trans. Syst. Man Cybern.* **SMC-9**, 860–864 (1979).
89. G. Fekete, J. O. Eklundh, and A. Rosenfeld, ''Relaxation: evaluation and applications,'' *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-3**(4), 459–469 (1981).
90. A. Rosenfeld and R. Smith, ''Thresholding using relaxation,'' *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-3**, 598–606 (1981).
91. A. Y. Wu, T. H. Hong, and A. Rosenfeld, ''Threshold selection using

quadtrees,'' *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-4**(1), 90–94 (1982).
92. N. Ahuja and A. Rosenfeld, ''A note on the use of second-order gray-level statistics for threshold selection,'' *IEEE Trans. Syst. Man Cybern.* **SMC-5**, 383–388 (1975).
93. W. N. Lie, ''An efficient threshold-evaluation algorithm for image segmentation based on spatial gray level cooccurrences,'' *Signal Process.* **33**, 121–126 (1993).
94. N. R. Pal and S. K. Pal, ''Entropic thresholding,'' *Signal Process.* **16**, 97–108 (1989).
95. C. Chang, K. Chen, J. Wang, and M. L. G. Althouse, ''A relative entropy based approach in image thresholding,'' *Pattern Recogn.* **27**, 1275–1289 (1994).
96. B. Chanda and D. D. Majumder, ''A note on the use of gray level co-occurrence matrix in threshold selection,'' *Signal Process.* **15**, 149–167 (1988).
97. A. S. Abutaleb, ''Automatic thresholding of gray-level pictures using two-dimensional entropy,'' *Comput. Vis. Graph. Image Process.* **47**, 22–32 (1989).
98. H. D. Cheng and Y. H. Chen, ''Thresholding based on fuzzy partition of 2D histogram,'' *Intl. Conf. Patt. Recog.*, pp. 1616–1618 (1998).
99. L. Li, J. Gong, and W. Chen, ''Gray-level image thresholding based on fisher linear projection of two-dimensional histogram,'' *Pattern Recogn.* **30**, 743–749 (1997).
100. A. D. Brink, ''Thresholding of digital images using two-dimensional entropies,'' *Pattern Recogn.* **25**, 803–808 (1992).
101. H. D. Cheng and Y. H. Chen, ''Fuzzy partition of two-dimensional histogram and its application to thresholding,'' *Pattern Recogn.* **32**, 825–843 (1999).
102. A. D. Brink, ''Gray level thresholding of images using a correlation criterion,'' *Pattern Recogn. Lett.* **9**, 335–341 (1989).
103. A. D. Brink, ''Minimum spatial entropy threshold selection,'' *IEE Proc. Vision Image Signal Process.* **142**, 128–132 (1995).
104. A. Beghdadi, A. L. Negrate, and P. V. De Lesegno, ''Entropic thresholding using a block source model,'' *Graph. Models Image Process.* **57**, 197–205 (1995).
105. C. K. Leung and F. K. Lam, ''Maximum a posteriori spatial probability segmentation,'' *IEE Proc. Vision Image Signal Process.* **144**, 161–167 (1997).
106. N. Friel and I. S. Molchanov, ''A new thresholding technique based on random sets,'' *Pattern Recogn.* **32**, 1507–1517 (1999).
107. G. Borgefors, ''Distance transformations in digital images,'' *Comput. Vis. Graph. Image Process.* **34**, 344–371 (1986).
108. Y. Nakagawa and A. Rosenfeld, ''Some experiments on variable thresholding,'' *Pattern Recogn.* **11**(3), 191–204 (1979).
109. F. Deravi and S. K. Pal, ''Grey level thresholding using second-order statistics,'' *Pattern Recogn. Lett.* **1**, 417–422 (1983).
110. W. Niblack, *An Introduction to Image Processing*, pp. 115–116, Prentice-Hall, Englewood Cliffs, NJ (1986).
111. J. Sauvola and M. Pietaksinen, ''Adaptive document image binarization,'' *Pattern Recogn.* **33**, 225–236 (2000).
112. J. M. White and G. D. Rohrer, ''Image thresholding for optical character recognition and other applications requiring character image extraction,'' *IBM J. Res. Dev.* **27**(4), 400–411 (1983).
113. J. Bernsen, ''Dynamic thresholding of gray level images,'' *ICPR'86: Proc. Intl. Conf. Patt. Recog.*, pp. 1251–1255 (1986).
114. Y. Yasuda, M. Dubois, and T. S. Huang, ''Data compression for check processing machines,'' *Proc. IEEE* **68**, 874–885 (1980).
115. S. D. Yanowitz and A. M. Bruckstein, ''A new method for image segmentation,'' *Comput. Graph. Image Process.* **46**, 82–95 (1989).
116. D. Shen and H. H. S. Ip, ''A Hopfield neural network for adaptive image segmentation: An active surface paradigm,'' *Pattern Recogn. Lett.* **18**, 37–48 (1997).
117. N. B. Venkateswarluh and R. D. Boyle, ''New segmentation techniques for document image analysis,'' *Image Vis. Comput.* **13**, 573–583 (1995).
118. E. Giuliano, O. Paitra, and L. Stringer, ''Electronic character reading system,'' U.S. Patent No. 4,047,15 (Sep. 1977).
119. Y. Yang and H. Yan, ''An adaptive logical method for binarization of degraded document images,'' *Pattern Recogn.* **33**, 787–807 (2000).
120. F. H. Y. Chan, F. K. Lam, and H. Zhu, ''Adaptive thresholding by variational method,'' *IEEE Trans. Image Process.* **IP-7**, 468–473 (1991).
121. J. Parker, ''Gray level thresholding on badly illuminated images,'' *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-13**, 813–891 (1991).
122. F. Chang, K. H. Liang, T. M. Tan, and W. L. Hwang, ''Binarization of document images using Hadamard multiresolution analysis,'' *ICDAR'99: Intl. Conf. Document Anal. Recog.*, pp. 157–160 (1999).
123. A. Savakis, ''Adaptive document image thresholding using foreground and background clustering,'' *ICIP'98: Intl. Conf. Image Process.*, Chicago, October 1998.
124. J. D. Yang, Y. S. Chen, and W. H. Hsu, ''Adaptive thresholding algorithm and its hardware implementation,'' *Pattern Recogn. Lett.* **15**, 141–150 (1994).
125. H. Kamada and K. Fujimoto, ''High-speed, high-accuracy binarization method for recognizing text in images of low spatial resolution,'' *ICDAR'99, Intl. Conf. Document Anal. Recog.*, pp. 139–142 (1999).
126. L. Eikvil, T. Taxt, and K. Moen, ''A fast adaptive method for binarization of document images,'' *ICDAR'91, Intl. Conf. Document Anal. Recog.*, pp. 435–443 (1991).
127. T. Pavlidis, ''Threshold selection using second derivatives of the gray-scale image,'' *ICDAR'93, Intl. Conf. Document Anal. Recog.*, pp. 274–277 (1993).
128. X. Zhao and S. H. Ong, ''Adaptive local thresholding with fuzzy-validity guided spatial partitioning,'' *ICPR'98, Intl. Conf. Patt. Recog.*, pp. 988–990 (1998).
129. W. A. Yasnoff, J. K. Mui, and J. W. Bacus, ''Error measures for scene segmentation,'' *Pattern Recogn.* **9**, 217–231 (1977).
130. M. D. Levine and A. M. Nazif, ''Dynamic measurement of computer generated image segmentations,'' *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-7**, 155–164 (1985).
131. Y. J. Zhang, ''A survey on evaluation methods for image segmentation,'' *Pattern Recogn.* **29**, 1335–1346 (1996).
132. M. P. Dubuisson and A. K. Jain, ''A modified Hausdorff distance for object matching,'' *ICPR'94, 12th Intl. Conf. Patt. Recog.*, A-566-569 (1994).
133. D. Demir, S. Birecik, F. Kurugöllü, M. Sezgin, İ. Ö. Bucak, B. Sankur, and E. Anarim, ''Quality inspection in PCBs and SMDs using computer vision techniques,'' *20th Intl. Conf. Industrial Elec. Control Instrum.*, pp. 857–860 (1994).
134. H. S. Baird, ''Document image defect models and their uses,'' *ICDAR'92, Proc. Intl. Conf. Document Anal. Recog.*, pp. 62–67 (1992).
135. M. Sezgin, ''Quantitative evaluation of image thresholding methods and application to nondestructive testing,'' PhD Thesis, Istanbul Technical University, Turkey (2002).
136. X. Ye, M. Cheriet, and C. Y. Suen, ''Stroke-model-based character extraction from gray-level document images,'' *IEEE Trans. Image Process.* **IP-10**(8), 1152–1161 (2001).
137. S. Djeziri, F. Nouboud, and R. Plamondon, ''Extraction of signatures from check background based on a filiformity criterion,'' *IEEE Trans. Image Process.* **IP-7**(10), 1425–1438 (1998).
138. C. Strouthopoulos and N. Papamarkos, ''Multithresholding of mixed type documents,'' *Eng. Applic. Artif. Intell.* **13**(3), 323–343 (2000).
139. A. H. Dekker, ''Kohonen neural networks for optimal colour quantization,'' *Network Comput. Neural Syst.* **5**, 351–367 (1994).
140. C. M. Tsai and H. H. Lee, ''Binarization of color document images via luminance and saturation color features,'' *IEEE Trans. Image Process.* **IP-11**, 434–451 (Apr. 2002).
141. N. Papamarkos and B. Gatos, ''A new approach for multithreshold selection,'' *Comput. Vis. Graph. Image Process* **56**(5), 357–370 (1994).

Mehmet Sezgin graduated from İstanbul Technical University in 1986 from the electronic and communication department. He completed his Msc and Phd at the same university in 1990 and 2002, respectively. His research areas are signal and image processing. He has been working at the Electronic Research Institute and Information Technologies Research Institute of the Turkish Scientific and Technical Research Council since 1991.

Bülent Sankur received his BS degree in electrical engineering at Robert College, Istanbul, and completed his MSc and PhD degrees at Rensselaer Polytechnic Institute, New York. He has been active at Boğaziçi University in the Department of Electric and Electronic Engineering in establishing curricula and laboratories, and guiding research in the areas of digital signal processing, image and video compression, and multimedia systems. He was the chairman of the International Telecommunications Conference and the technical co-chairman of ICASSP 2000. He has held visiting positions at the University of Ottawa, Canada: Istanbul Technical University: Technical University of Delft, The Netherlands: and Ecole Nationale Supérieure des Telecommunications, France.