56

IMAGE MODEL, POISSON DISTRIBUTION AND OBJECT EXTRACTION

NIKHIL R. PAL and SANKAR K. PAL
Electronics and Communication Sciences Unit
Indian Statistical Institute
203 B. T. Road, Calcutta–700035, India

The theory of formation of an ideal image has been described which shows that the gray level in an image follows the Poisson distribution. Based on this concept, various algorithms for object background classification have been developed. Proposed algorithms involve either the maximum entropy principle or the minimum χ^2 statistic. The appropriateness of the Poisson distribution is further strengthened by comparing the results with those of similar algorithms which use conventional normal distribution. A set of images with various types of histograms has been considered here as the test data.

Keywords: Ideal-image model; Poisson distribution; χ^2 statistic; Entropy; Thresholding.

1. INTRODUCTION

The extraction of an object from its background is an essential step of computer vision and scene analysis. The existing methods can broadly be classified into two categories, namely, methods based on global information (gray-level histogram) and methods based on local information (spatial details). Some of the methods[1–4] sharpen the histogram to facilitate the task of thresholding, while others optimize some objective function for the selection of the threshold. In the recent past there had been some research on entropic thresholding[5–9] in which entropy was considered as a measure of information in a gray tone image. Various types of entropy, e.g. global, local and conditional, have been defined[5–7] in this context. Global entropic thresholding depends only on the information in the histogram, whereas the local and conditional entropic methods take into account the information present in the co-occurrence matrix of an image. As expected, the latter is found to be more effective than the former.

It may be recalled that none of the above-mentioned techniques considers the theory of formation of an ideal image while formulating an algorithm. For example, the existing entropic thresholding algorithms estimate the probability of occurrence p_i of the ith level in a region without making use of the appropriate distribution which the gray-level variation may follow. There are some other algorithms[10] for object–background classification which use the well-known normal distribution for the gray level without any justification.

The work presented here is an attempt to formulate object extraction algorithms based on the theory of formation of an ideal image. An ideal imaging process has been described which shows that the gray-level distributions within the object and

International Journal of Pattern Recognition and Artificial Intelligence Vol. 5 No. 3 (1991) 459–483

background can be approximated with Poisson distributions characterized by two different parameters. In order to establish the validity of the proposed concept, two approaches for object–background classification have been adopted. The first approach is dependent on the maximum entropy principle whereas the other is based on the minimum χ^2 statistic. The χ^2 statistic is often used as a criterion for testing goodness of fit of a distribution. Therefore, if an appropriate distribution is assumed, the minimum χ^2 is likely to produce a good segmentation.

It has been demonstrated, on a set of various images, that algorithms which use the Poisson distribution perform better than similar algorithms which use the normal distribution or the empirical distributions. The appropriateness of the exponential entropy over the logarithmic entropy, to represent image information, has also been established here. The performance of the proposed methods has also been compared with that of some of the existing ones.[2,7,8]

Although the algorithms considered here are all based on global information, the thresholds obtained by the Poisson distribution-based algorithms conform well to those of local (spatial) information-based algorithms.[5,6]

2. AN IDEAL-IMAGE MODEL

An ideal imaging device can be thought of as a spatial array of individual photon receptors and counters, each with identical properties. Obviously, the spatial resolution of the image is governed by the spatial dimension of the receptors. It is assumed that each receptor can receive light quanta (photon) independent of its neighbouring receptors. The image state of a receptor is completely determined by the number of quanta it receives and records, and each receptor can be in one of a finite number of distinguishable image states. Since the possible number of states of a receptor is finite, after a receptor attains the final state, all other additional incident quanta will remain unrecorded. In other words, the receptor gets saturated.

If we feel the exposure level over the entire imaging device is uniform, the number of incident quanta is found to follow the Poisson distribution with a sufficient degree of validity.[11] In other words, if the uniform exposure level is such that each receptor receives, on the average, q quanta then

$$\begin{aligned} p_r &= \text{probability of a counter receiving } r \text{ quanta} \\ &= \frac{q^r e^{-q}}{r!} \qquad \text{for } r = 0, 1, \ldots, q, \ldots \qquad (1) \\ &= \text{proportion of counters receiving } r \text{ quanta.} \end{aligned}$$

If the number of received quanta exceeds the saturation level (S), the excess quanta will not be reflected in the recorded value. In fact in any photographic process, in addition to this upper limit, there is also a lower limit, i.e. threshold (T), to the number of recorded quanta. In other words, as long as the number of incident quanta is less than T, no quanta will be recorded and the Tth incident quanta will be recorded as one. The behaviour of the recorded quanta is shown in Fig. 1.

Though the number of incident quanta follows the Poisson distribution with average number λ, the average value of the recorded number of quanta will be different from λ because of the two limits mentioned above. The average value a of the recorded number of quanta is given by

$$a = \sum_{x=0}^{\infty} x' \frac{e^{-\lambda}\lambda^x}{x!} \tag{2}$$

where

$$\begin{aligned} x' &= 0 && \text{for } x < T \\ &= x - T + 1 && \text{for } T \le x < S \\ &= S - T + 1 && \text{for } x \ge S. \end{aligned}$$

On the other hand,

$$\lambda = \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!}\,. \tag{3}$$

Obviously, $a \neq \lambda$.

In the case of a digital image, each pixel can be viewed as a receptor. Like the ideal imaging system, the spatial resolution here depends also on the spatial size of the pixel and each pixel can have only a finite number of states with a saturation level. The

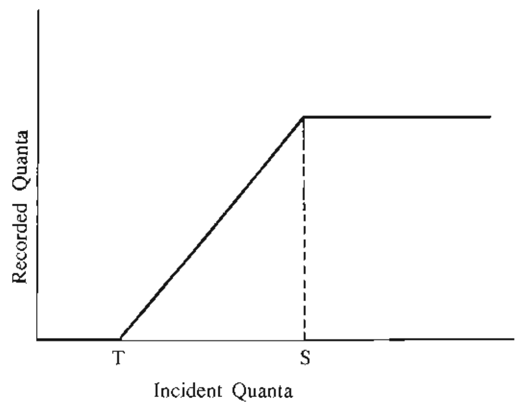


Fig. 1. Relationship between recorded and incident quanta.

observed gray level of a pixel is nothing but the effect of the received quanta by the corresponding receptor. The larger the number of recorded quanta, the higher is the gray value. For the sake of simplicity we assume the probability that the number of incident quanta is less than T or is greater than S is very small. Thus, under the above assumption, the average number of recorded quanta a (Eq. (2)) will be, for all practical purposes, equal to λ. Therefore, the number of recorded quanta will approximately follow the Poisson distribution with an average value of λ.

Let us now consider a scene consisting of an ideal object and an ideal background. An ideal object means that the entire object surface has uniform characteristics (i.e. constant coefficient of reflection, constant temperature distribution, made up of same material and so on). Similarly, the ideal background also has uniform characteristics, but they are obviously different from those of the object. When we take the photograph of such an ideal scene illuminated by uniform light, the scene itself acts as the source of light for the imaging system. Though the illumination of the entire scene is uniform, the object and background exposure levels for the imaging system will be of two different natures since they have different characteristics. Ignoring the interaction between the quanta emitted by the object with those emitted by the background, we can say that the recorded uniform image will have two uniform illuminations, one corresponding to the object and the other to the background.

We assume that the pixel value of a cell is equal to the number of recorded quanta measured by the corresponding receptor cell of the imaging system. In other words, this implies that the pixel value for uniform illumination follows the Poisson distribution. Thus the gray-level histogram of a digital image will be a combination of two Poisson distributions characterized by two different parameters λ_O and λ_B.

To establish the validity of the above model, the idea of the Poisson distribution has been used to develop various image segmentation algorithms. Two approaches, namely, entropy maximization and χ^2 statistic minimization are adopted in the aforesaid context.

The superiority of the Poisson distribution over the commonly used normal distribution for the gray level in a digital image has been demonstrated by considering the same algorithms with normal distribution.

3. MAXIMUM ENTROPIC THRESHOLDING (MAXET)

Before proceeding further let us define a digital image as follows. A digital image $\mathbf{F} = [f(x, y)]_{P \times Q}$ is a matrix of size $P \times Q$, where $f(x, y)$ is the gray level at (x, y) and $f(x, y) \in \{0, 1, \ldots, L-1\}$, the set of gray levels. Also define p_i, the probability of the gray level i as $p_i = N_i/N$, where N_i is the frequency of the gray level i and $N = P \times Q$.

Shannon[12,13] defined the entropy of an n-state system as

$$H = -\sum_{i=1}^{n} p_i \log p_i\,, \qquad \sum_{i=1}^{n} p_i = 1\,, \qquad 0 \le p_i \le 1 \tag{4}$$

where p_i is the probability of the ith state of the system. Such a measure is claimed to give information about the actual probability structure of the system. The drawback of this measure of information has recently been pointed out by Pal and Pal[5] and the entropy of an n-state system as suggested by them is

$$H = \sum_{i=1}^{n} p_i e^{1-p_i} , \qquad \sum_{i=1}^{n} p_i = 1 , \qquad 0 \le p_i \le 1 . \tag{5}$$

The term $\{-\log(p_i)\}$, i.e. $\log(1/p_i)$ in Eq. (4) or e^{1-p_i} in Eq. (5) is called the gain in information from the occurrence of the ith event. Let us denote it by $\Delta I(p_i)$, i.e. $\Delta I(p_i)$ is the gain in information from the occurrence of the ith state of the n-state system. Thus, in general we can write the expression of entropy as

$$H = \sum_{i=1}^{n} p_i \Delta I(p_i) ,$$

where $\Delta I(p_i)$ can either be $\log(1/p_i)$ or e^{1-p_i} depending on the definition used.

In this context certain points in support of the exponential entropy are in order.[5]

(i) It is to be noted from the logarithmic entropic measure that as $p_i \to 0$, $\Delta I(p_i) \to \infty$ and $\Delta I(p_i) = -\log(p_i)$ is not defined for $p_i = 0$. On the other hand, as $p_i \to 1$, $\Delta I(p_i) \to 0$ and $\Delta I(p_i = 1) = 0$.

Thus we see that information gain from an event is neither bounded at both ends nor defined at all points. In practice, the gain in information from an event, whether highly probable or highly unlikely, is expected to lie between two finite limits. For example, as more and more pixels in an image are analysed, the gain in information increases, and when all the pixels are inspected the gain attains its maximum value, irrespective of the content of the image.

(ii) In Shannon's theory the measure of ignorance or the gain in information is taken as $\log(1/p_i)$, i.e. ignorance is inversely related to p_i. But mathematically, it is possible to arrive at a more sound expression. If u_i is the uncertainty of the ith event, then using the knowledge of probability one can write $u_i = 1 - p_i$. Since u_i is the unlikeliness (i.e. the probability of non-occurrence), statistically, ignorance can be better represented by $(1 - p_i)$ than $(1/p_i)$.

Now if we define the gain in information corresponding to the occurrence of the ith event as

$$\Delta I(p_i) = \log(1 - p_i) ,$$

then $\Delta I < 0$ which is intuitively unappealing. Furthermore, consideration of $-\log(1 - p_i)$ as the gain in information leads to the fact that $\Delta I(p_i)$ increases with p_i, which is again not desirable.

The above problem is circumvented by considering an exponential function of $(1 - p_i)$ instead of the logarithmic behaviour. This is also appropriate while considering the information gain in an image.

For example, consider Fig. 2. Suppose the images (Figs. 2(a)–(e)) have only two gray levels: one corresponding to the lines (black portion) and the other corresponding to the white portion. In the case of the first image we have analysed only a few black pixels and from it we cannot speak firmly about the content of the image. At this stage we see that it can be either a curtain or hair surrounding a face or something else. From image 2(b) we can say that it is not a curtain (i.e. there is some gain in knowledge) while from image 2(c) we realize that it is a face. Image 2(d) shows a face with a mouth. However, image 2(e) does not say anything more than what is described by image 2(d), though the number of black pixels (hence probability) has increased.

Let $\Delta I(a)$, $\Delta I(b)$, $\Delta I(c)$, $\Delta I(d)$ and $\Delta I(e)$ be the information content of the images 2(a) through 2(e) respectively. Now define the following quantities representing the changes in gain.

$$G1 = \Delta I(b) - \Delta I(a) \tag{6.1}$$

$$G2 = \Delta I(c) - \Delta I(b) \tag{6.2}$$

$$G3 = \Delta I(d) - \Delta I(c) \tag{6.3}$$

$$G4 = \Delta I(e) - \Delta I(d) \tag{6.4}$$

Obviously, $G1 > G2 > G3 > G4 \cong 0$.

The above observation and the fact that information gain approaches a finite limit when more and more pixels (increase in N_i and hence p_i) are analysed, led us to accept e^{1-p} as the gain in information from the occurrence of the ith event and Eq. (5) as the entropy of the system.

Let us now go back to our original problem of segmentation. Suppose $g_O(x)$ is the probability function (or probability density function) of the gray level x over the object and $g_B(x)$ is the same over the background. The maximum entropic thresholding principle may now be stated as follows. Partition the image into two non-intersecting regions (say, object and background) such that the total entropy of $g_O(x)$ and $g_B(x)$ is maximized. In other words, when $g_O(x)$ and $g_B(x)$ are discrete, maximize

$$\sum_{x \in \text{object}} g_O(x) \cdot \Delta I(g_O(x)) + \sum_{x \in \text{background}} g_B(x) \cdot \Delta I(g_B(x)) \, . \tag{7}$$

On the other hand, when $g_O(x)$ and $g_B(x)$ are continuous, one needs to maximize

$$\int_{x \in \text{object}} g_O(x) \cdot \Delta I(g_O(x))dx + \int_{x \in \text{baekground}} g_B(x) \cdot \Delta I(g_B(x))dx \, . \tag{8}$$

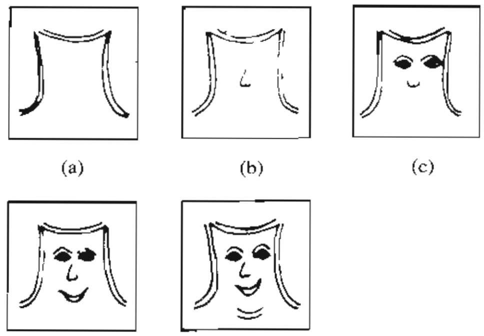


Fig. 2. Variation of the gain in information in images.

In the previous section we justified that the gray levels within the object and the background follow Poisson distributions with two different parameters λ_O and λ_B. Thus,

$$g_O(x) = \frac{e^{-\lambda_O}\lambda_O^x}{x!} \tag{9}$$

and

$$g_B(x) = \frac{e^{-\lambda_B}\lambda_B^x}{x!} \tag{10}$$

Therefore, the problem is to find a gray level l, $0 < l < L - 1$, and $\lambda_O < l < \lambda_B$, such that the gray levels in the range 0 to l represent the object with expected uniform illumination λ_O, and the gray levels in the range $l + 1$ to $L - 1$ constitute the background with average illumination λ_B. In order to achieve this we maximize Eq. (7). Maximization of Eq. (7) requires the evaluation of $g_O(x)$ and $g_B(x)$ and hence requires the values of λ_O and λ_B which are to be estimated from the input digital image. There are various methods of estimation of the parameter λ of a Poisson distribution. We use here the maximum likelihood (ML) estimate of λ. For some hypothetical boundary l, the ML estimates of λ_O and λ_B are given as follows

$$\hat{\lambda}_O = \frac{\sum_{i=0}^{l} iN_i}{\sum_{i=0}^{l} N_i} \tag{11}$$

$$\hat{\lambda}_B = \frac{\sum_{i=l+1}^{L-1} iN_i}{\sum_{i=l+1}^{L-1} N_i} . \tag{12}$$

($\hat{}$ is used to indicate the estimated values.)

Thus, the estimated probability of a gray level x, $x \in$ object, is given by

$$\hat{p}_x^{O} = \frac{e^{-\hat{\lambda}_O}\hat{\lambda}_O^x}{x!} \tag{13}$$

and that for an $x \in$ background is

$$\hat{p}_x^{B} = \frac{e^{-\hat{\lambda}_B}\hat{\lambda}_B^x}{x!} . \tag{14}$$

Therefore, for an assumed boundary l, $0 < l < L-1$, the total entropy of the partitioned image can be written as

$$H_l = \sum_{x=0}^{l} \hat{p}_x^{O}\Delta I(\hat{p}_x^{O}) + \sum_{x=l+1}^{L-1} \hat{p}_x^{B}\Delta I(\hat{p}_x^{B}) . \tag{15}$$

Based on H_l (Eq. (15)), the following two algorithms can be formulated.

3.1. Algorithm 1

In this method we take $\Delta I(p_x) = e^{1-p_x}$ and maximize Eq. (15) with respect to l.

3.2. Algorithm 2

This algorithm uses Shannon's formula of entropy, i.e. it takes $\Delta I(p_x) = \log(1/p_x)$.

In order to strengthen the concept of the Poisson distribution we have also experimented with MAXET using normal distribution to describe the probability distribution of gray levels within a homogeneous region. Thus, we assume that the gray level x is continuous and

$$g_O(x) \approx \mathrm{N}(\mu_O, \sigma_O^2) \tag{16}$$

and

$$g_B(x) \approx \mathrm{N}(\mu_B, \sigma_B^2) , \tag{17}$$

i.e.

$$g_O(x) = \frac{1}{\sigma_O\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu_O}{\sigma_O}\right)^2\right\} \tag{18}$$

and

$$g_B(x) = \frac{1}{\sigma_B\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu_B}{\sigma_B}\right)^2\right\} \tag{19}$$

where μ_O, μ_B are means of the two normal distributions and σ_O, σ_B are the standard deviations of the same. With the above forms of probability density functions, we maximize the total entropy of the partitioned image as expressed by Eq. (8). However, the use of Eqs. (18) and (19) demands the knowledge of μ_O, μ_B, σ_O and σ_B. Like the previous method, we use here the ML estimators to estimate the μ and σ of the distributions. The ML estimates of the parameters are given by

$$\hat{\mu}_O = \frac{\sum_{i=0}^{l} iN_i}{\sum_{i=0}^{l} N_i} \tag{20}$$

$$\hat{\mu}_B = \frac{\sum_{i=l+1}^{L-1} iN_i}{\sum_{i=l+1}^{L-1} N_i} \tag{21}$$

$$\hat{\sigma}_O^2 = \frac{\sum_{i=0}^{l} (i-\hat{\mu}_O)^2 N_i}{\sum_{i=0}^{l} N_i} \tag{22}$$

$$\hat{\sigma}_B^2 = \frac{\sum_{i=l+1}^{L-1} (i-\hat{\mu}_B)^2 N_i}{\sum_{i=l+1}^{L-1} N_i}. \tag{23}$$

With the above parameter values, $g_O(x)$ and $g_B(x)$ can be determined and the following algorithms can be formulated.

3.3. Algorithm 3

In this algorithm, we use $\Delta I(g(x)) = e^{1-g(x)}$ and maximize, for an assumed threshold l,

$$H_l^e = \int_{x \in \text{object}} g_O(x)\Delta I(g_O(x))dx + \int_{x \in \text{background}} g_B(x)\Delta I(g_B(x))dx \,. \quad (24)$$

In this method x (here l) is assumed to be continuous. While selecting the optimum threshold, we compute H_l^e for different l at unit intervals over the entire range of the gray levels. Since an image can possess only discrete levels, one-dimensional grid search with unit interval suffices.

3.4. Algorithm 4

This algorithm is the same as Algorithm 3, except that it uses Shannon's entropy, i.e. it uses $\Delta I(g(x)) = \log(1/g(x))$.

4. MINIMUM CHI-SQUARE THRESHOLDING (MINCST)

Let us describe here another set of thresholding algorithms based on the minimum chi-square (χ^2) statistic to demonstrate the appropriateness of the Poisson distribution for gray levels. Let $N_1, N_2, \ldots, N_k$ be the observed frequencies of k different classes and $p_1, p_2, \ldots, p_k$ be the hypothetical probabilities of a normal distribution. In order to test the goodness of fit of the hypothetical probabilities to the observed ones, K. Pearson[14] suggested the criterion,

$$\chi^2 = \sum_{i=1}^{k} \frac{(N_i - Np_i)^2}{Np_i} = \sum_{i=1}^{k} \frac{(\text{observed} - \text{expected})^2}{\text{expected}} \quad (25)$$

where N is the total number of observations and the asymptotic distribution (ad) of χ^2 as defined in Eq. (25) is a $\chi^2(k-1)$, i.e. a chi-square with $(k-1)$ degrees of freedom.

The more general problem may be to test whether the class probabilities are some specific functions of a few parameters which may be known. If the class probabilities are some specified functions, $p_1(\theta), p_2(\theta), \ldots, p_k(\theta)$, where θ has q (say) components, i.e. $\theta = (\theta_1, \theta_2, \ldots, \theta_q)$ and $\hat{\theta}$ is an efficient estimator of θ, then under suitable conditions[14] it can be shown that

$$\chi^2 = \sum_{i=1}^{k} \frac{(N_i - Np_i(\hat{\theta}))^2}{Np_i(\hat{\theta})}$$

is a $\chi^2(k - 1 - q)$.

This statistic is known as the χ^2 statistic, and is very often used to measure the goodness of fit of some hypothetical distribution to an observed data. The lower the value of χ^2, the better is the fit. The expression for χ^2 may also be viewed as a weighted sum of the squared deviation of the expected frequencies from the observed ones. For the image segmentation problem, the minimum χ^2 thresholding (MINCST) principle may thus be stated as follows: "Given the distribution that the gray levels may follow, partition the image into two non-intersecting regions, such that the sum of

χ^2 (over the object and the background) is minimized". Based on this principle, two algorithms, one using the Poisson distribution and the other using normal distribution, are suggested.

4.1. Algorithm 5

Let $g_O(x)$ and $g_B(x)$ (Eqs. (9) and (10)) be the probability functions of the gray levels in the object and background regions, respectively. For an arbitrary threshold l, the observed frequencies of gray levels over the object are given by

$$N_i^O = N_i, \qquad i = 0, 1, 2, \ldots, l \tag{26}$$

and those over the background by

$$N_i^B = N_i, \qquad i = l+1, l+2, \ldots, L-1. \tag{27}$$

On the other hand, the estimated probability of a gray level i, $0 \le i \le l$, i.e. i is in the object, is given by $g_O(i)$, $i = 0, 1, 2, \ldots, l$, and that over the background is given by $g_B(i)$, $i = l+1, l+2, \ldots, L-1$. Normalising $g_O(i)$ and $g_B(i)$ over the object and the background regions respectively, we get the expected probabilities as follows.

$$\hat{p}_i^O = \frac{g_O(i)}{\sum_{j=0}^{l} g_O(j)}, \qquad i = 0, 1, 2, \ldots, l \tag{28}$$

$$\hat{p}_i^B = \frac{g_B(i)}{\sum_{i=l+1}^{L-1} g_B(j)}, \qquad i = l+1, l+2, \ldots, L-1. \tag{29}$$

Thus, the expected frequencies of the gray levels over the object is given by

$$\hat{N}_i^O = \hat{p}_i^O N_O, \qquad i = 0, 1, 2, \ldots, l \tag{30}$$

$$\hat{N}_i^B = \hat{p}_i^B N_B, \qquad i = l+1, l+2, \ldots, L-1 \tag{31}$$

where N_O and N_B are the total number of observations within the object and the background, respectively. The parameters λ_O and λ_B which are required for the evaluation of $g_O(x)$ and $g_B(x)$ are estimated using the ML estimator (described in Sec. 3). Therefore, for a threshold l, the total chi-square is given by

$$\chi_l^2 = \sum_{i=0}^{l} \frac{(N_i^O - \hat{N}_i^O)^2}{\hat{N}_i^O} + \sum_{i=l+1}^{L-1} \frac{(N_i^B - \hat{N}_i^B)^2}{\hat{N}_i^B}. \tag{32}$$

The optimum threshold is obtained by minimizing χ_l^2 with respect to l.

4.2. Algorithm 6

This algorithm is also based on the principle of MINCST, but it assumes normal distributions (Eqs. (18) and (19)) for the gray levels. In order to compute the χ^2 for an arbitrary threshold l, we proceed as follows.

For each gray level y, $y = 1, 2, \ldots,$ consider a band $(y - 0.5, y + 0.5)$ and compute I_y as follows.

$$I_y = \int_{y-0.5}^{y+0.5} g_{\mathrm{O}}(x)dx \qquad \text{for } 0 < y < l \tag{33}$$

and

$$I_y = \int_{y-0.5}^{y+0.5} g_{\mathrm{B}}(x)dx \qquad \text{for } l + 1 < y < L - 1 \,. \tag{34}$$

I_y gives the probability of the gray levels lying in the range $(y - 0.5, y + 0.5)$. The expected frequency in the band $(y - 0.5, y + 0.5)$ will then be

$$\hat{N}_y^{\mathrm{O}} = N_{\mathrm{O}} I_y \qquad \text{for } 0 < y < l \tag{35}$$

or

$$\hat{N}_y^{\mathrm{B}} = N_{\mathrm{B}} I_y \qquad \text{for } l + 1 < y < L - 1 \tag{36}$$

while the observed frequency of the band around y will be N_y, $y = 1, 2, \ldots,$.

Note that we have not considered the gray levels $y < 1$ and $y \geq l$ for the object; similarly, gray levels $y \leq l + 1$ and $y > L - 2$ for the background. In order to account for this, in the case of the object, we compute

$$I_0 = \int_{-\infty}^{0.5} g_{\mathrm{O}}(x)dx \tag{37}$$

and

$$I_l = \int_{l-0.5}^{\infty} g_{\mathrm{O}}(x)dx \tag{38}$$

so that

$$\hat{N}_0^{\mathrm{O}} = N_{\mathrm{O}} I_0 \tag{39}$$

and

$$\hat{N}_l^{\mathrm{O}} = N_{\mathrm{O}} I_l \,. \tag{40}$$

It is to be noted here that the sum of all the band integrals for $g_O(x)$ will be equal to 1, i.e.

$$\sum_{y=0}^{l} I_y = 1 .$$

Similar computation can also be done for the background probability distribution. Once we know the observed and expected frequencies in different classes, the χ^2 (Eq. (32)) can be computed and minimized with respect to l for selecting the threshold.

Before presenting the performance of the proposed methods, let us review some of the relevant methods which will be used for comparison of results.

5. REVIEW OF SOME RELATED METHODS

Pun[7] and Kapur *et al.*[8] considered the gray-level histogram of F as the outcome of an L-symbol source, independently of the underlying image. In addition to this, they also assumed that these symbols were statistically independent.

Following Shannon's definition of entropy (Eq. (4)), Pun[7] defined the entropy of the image (histogram) as

$$H = -\sum_{i=0}^{L-1} p_i \log_2 p_i ; \qquad p_i = \frac{N_i}{N} , \tag{41}$$

for the image segmentation problem.

5.1. Evaluation Function of Pun[7]

Let s be the threshold which classifies the image into object and background. Let N_B and N_W be the numbers of pixels in the black and white portions of the image. Then the *a posteriori* probability of a black pixel is $P_B = N_B/N$ and that of a white pixel is $P_W = N_W/N$. Thus, the *a posteriori* entropy of the image is:

$$\begin{aligned} H'_L(s) &= -P_B \log_2 P_B - P_W \log_2 P_W \\ &= -P_s \log_2 P_s - (1 - P_s) \log_2(1 - P_s) \end{aligned} \tag{42}$$

as

$$P_s = \sum_{i=0}^{s} p_i \qquad \text{and } P_W = 1 - P_s . \tag{43}$$

Since the maximization of H'_L gives the trivial result of $P_s = 1/2$, Pun maximized an upper bound $g(s)$ of $H'_L(s)$, where

$$g(s) = \frac{H_{\mathrm{B}}^{s} \log_2 P_s}{H_{\mathrm{L}} \log_2[\max(p_0, p_1, \ldots, p_s)]} + \frac{(H_{\mathrm{L}} - H_{\mathrm{B}}^{s}) \log_2(1 - P_s)}{H_{\mathrm{L}} \log_2[\max(p_{s+1}, p_{s+2}, \ldots, p_{L-1})]} \tag{44}$$

where

$$H_{\mathrm{L}} = -\sum_{i=0}^{L-1} p_i \log_2 p_i \tag{45}$$

and

$$H_{\mathrm{B}}^{s} = -\sum_{i=0}^{s} p_i \log_2 p_i \,. \tag{46}$$

The value of s which maximizes $g(s)$ can be taken as the threshold for object and background classification.

5.2. Method of Kapur, Sahoo and Wong[8]

Kapur *et al.* have also used Shannon's concept of entropy but from a different point of view. Instead of considering one probability distribution for the entire image, they considered two probability distributions, one for the object and the other for the background. The sum of the individual entropies of the object and the background is then maximized. In other words, this will result in equiprobable gray levels in each region, thus maximizing the sum of homogeneities in gray levels within the object and the background.

If s is an assumed threshold, then the probability distribution of the gray levels over the black portion of the image is

$$\frac{p_0}{P_s}, \frac{p_1}{P_s}, \ldots, \frac{p_s}{P_s},$$

and that of the white portion is

$$\frac{p_{s+1}}{1 - P_s}, \frac{p_{s+2}}{1 - P_s}, \ldots, \frac{p_{L-1}}{1 - P_s}.$$

The entropy of the black portion (object) of the image is

$$H_{\mathrm{B}}^{(s)} = -\sum_{i=0}^{s} \frac{p_i}{P_s} \log_2\left(\frac{p_i}{P_s}\right) \tag{47}$$

and that of the white portion is

$$H_{\mathrm{W}}^{(s)} = -\sum_{i=s+1}^{L-1} \frac{p_i}{1-P_s} \log_2 \left(\frac{p_i}{1-P_s}\right). \tag{48}$$

The total entropy of the image is then defined as

$$H_{\mathrm{T}}^{(s)} = H_{\mathrm{B}}^{(s)} + H_{\mathrm{W}}^{(s)}. \tag{49}$$

In order to select the threshold they maximized $H_{\mathrm{T}}^{(s)}$. In other words, the value of s which maximizes $H_{\mathrm{T}}^{(s)}$ gives the threshold for object and background classification.

5.3. Method of Otsu[2]

In order to evaluate the "goodness" of the threshold (at level s), Otsu maximized the between-class variance $\sigma_{\mathrm{B}}^2(s)$ defined by the following equation.

$$\sigma_{\mathrm{B}}^2(s) = P_s(\mu_0 - \mu_{\mathrm{T}})^2 + (1 - P_s)(\mu_{\mathrm{B}} - \mu_{\mathrm{T}})^2 \tag{50}$$

where

$$\mu_0 = \sum_{i=0}^{s} i \frac{p_i}{P_s},$$

$$\mu_{\mathrm{B}} = \sum_{i=s+1}^{L-1} i \frac{p_i}{1-P_s}$$

and

$$\mu_{\mathrm{T}} = \sum_{i=0}^{L-1} i\, p_i.$$

The expression for $\sigma_{\mathrm{B}}^2(s)$ can be simplified to the following

$$\sigma_{\mathrm{B}}^2(s) = \frac{(\mu_{\mathrm{T}} P_s - \mu_s)^2}{P_s(1-P_s)}. \tag{51}$$

Otsu maximized Eq. (51) over the range of s for which $P_s(1 - P_s) > 0$ or $0 < P_s < 1$, i.e. over the effective range of gray level.

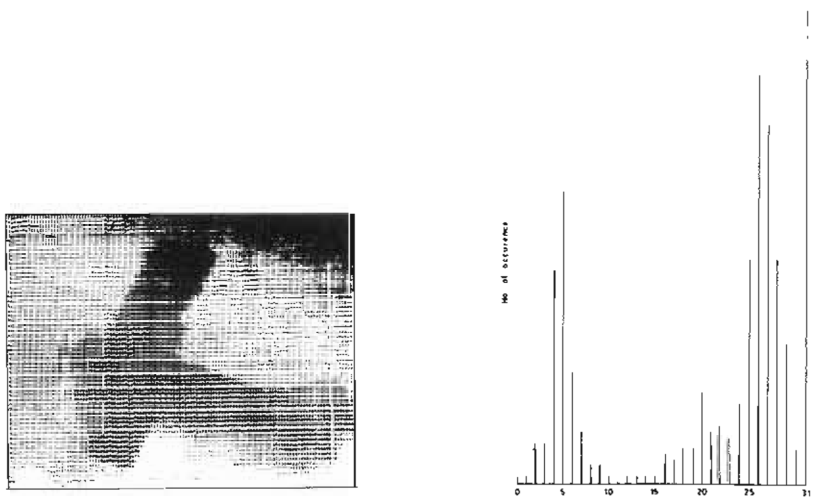
6. RESULTS

All the six algorithms discussed in Secs. 3 and 4 have been implemented on a set of four images (32-gray levels) with widely different types of histograms. In addition to this, the methods of Pun,[7] Kapur *et al.*[8] and Otsu[2] have also been implemented for comparison of results with those of the new methods suggested in this paper.

Table 1 displays the thresholds produced by different methods for various images. Figures 3(a) and 3(b) represent the input image of a biplane and its gray-level histogram, respectively. Figures 3(c)–3(k) give the segmented images produced by different methods.

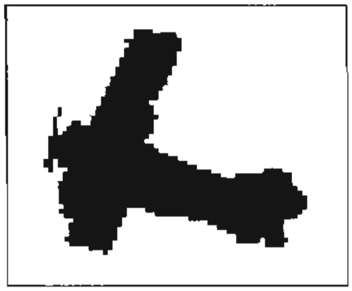
Table 1. Thresholds produced by different methods.

Methods			Images			
			Plane (Fig. 3)	Lincoln (Fig. 4)	Boy (Fig. 5)	Chromosome (Fig. 6)
MAXET (Maximum Entropy)	Alg. 1	Poisson distribution	12	11	9	12
	Alg. 2		13	12	11	16
	Alg. 3	Normal distribution	16	9	9	19
	Alg. 4		19	15	14	19
MINCST (Minimum χ^2)	Alg. 5	Poisson distribution	12	9	12	16
	Alg. 6	Normal distribution	12	6	5	—
Algorithm of Pun			24	16	15	27
Algorithm of Kapur *et al.*			21	15	13	20
Algorithm of Otsu			15	14	15	20

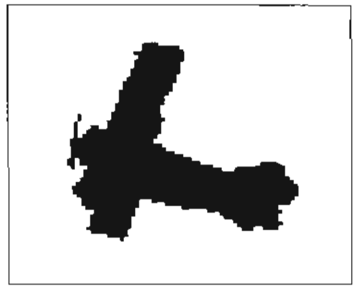


(a) (b)

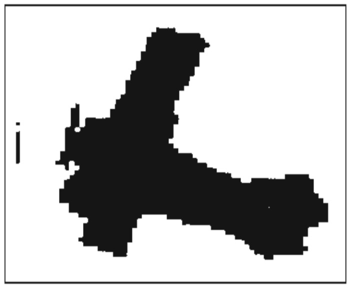
Fig. 3. Biplane image. (a) Input, (b) histogram. Segmentation result of (e) Algorithm 1, (d) Algorithm 2, (e) Algorithm 3, (f) Algorithm 4, (g) Algorithm 5, (h) Algorithm 6, (i) algorithm of Pun, (j) algorithm of Kapnr *et al.*, (k) algorithm of Otsu.



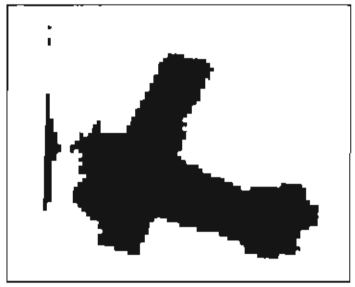
(c), (g), (h)



(d)

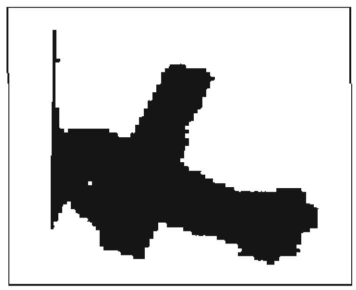


(e)

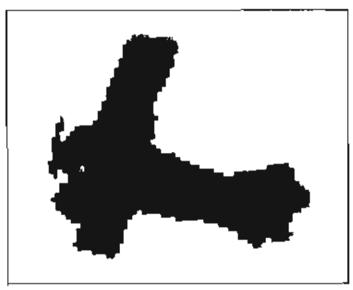


(f)

(i)



(j)



(k)

Fig. 3. Cont'd.

From the results, one can observe that Algorithms 1, 5 and 6 produced the same threshold and, in fact, are the best segmentations of the biplane image. Algorithm 2 produced almost identical results to that of the above three methods. On the other hand, the remaining algorithms could not extract the propeller in front of the biplane. It is to be noticed that for this image all the algorithms that used the Poisson distribution obtained good and comparable segments. Though the biplane has a bimodal histogram with a deep valley, Algorithms 3 and 4 (which use the normal distribution), the algorithm of Kapur *et al.*, the method of Pun and the algorithm of Otsu could not make the appropriate segmentation.

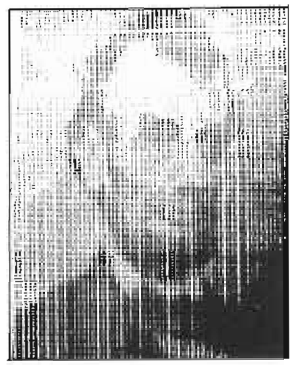
The Lincoln image (Fig. 4(a)) has a multimodal histogram (Fig. 4(b)). In this case, the best segmentation is produced by Algorithm 1 (Fig. 4(c)). Algorithms 2, 3 and 5 more or less correctly extracted the object (Figs. 4(d), (e) and (g)), while the remaining algorithms, which gave more or less the same threshold values, could not separate the beard of Lincoln at the lower right corner of the image (Figs. 4(f), 4(i)–(k)). Algorithm 6 produced a very low threshold. Here also all the algorithms which use the Poisson distribution give better results than the normal distribution-based algorithms. Another point to observe is that the exponential entropy gives a better threshold than the logarithmic entropy. In this context it is to be mentioned that for a multimodal image the proposed methods will make a bimodal approximation of the histogram, hence the threshold obtained by them may not always make an appropriate segmentation. Normally in this situation the image will have more than two regions and object background segmentation will not be possible.

Figures 5(a) and (b) depict the input image and the gray-level histogram of a boy. Here, except for Algorithm 6, which resulted in a low threshold, all other methods produced acceptable results (Figs. 5(c)–(k)). However, a critical analysis of the results (Figs. 5(c)–(k)) shows that the best result is produced by Algorithms 2 and 5 (Figs. 5(d) and (g)), as they preserved the features of the face and also extracted the ring in the left ear of the image.

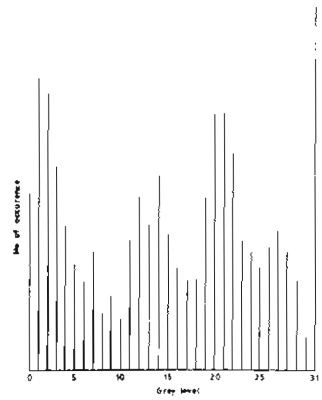
The proposed algorithms are also tested on an image of a set of three chromosomes (Fig. 6(a)) with a unimodal histogram (Fig. 6(b)). In this case, except for Algorithms 1 and 6 and the method of Pun, all the others produced comparable results (Figs. 6(c)–(j)). Algorithm 1 produced a thinned version of the chromosomes (Fig. 6(c)) while Algorithm 6 failed to make any segmentation. It is to be noted that the method of Pun could not extract one of the chromosomes.

Although the algorithms described here are all based on global information of an image, the thresholds obtained by the Poisson distribution-based methods conform well to those obtained by the local information-based methods (e.g. conditional entropic segmentation technique[5]). Based on the above observation, the following points can be mentioned.

All the algorithms which use the Poisson distribution (Algorithms 1, 2 and 5) consistently result in a better segmentation than the others. This observation establishes the validity of our digital image model based on the Poisson distribution. It reveals that the Poisson distribution is more appropriate than normal distribution for

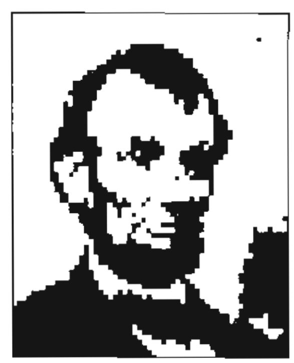


(a)

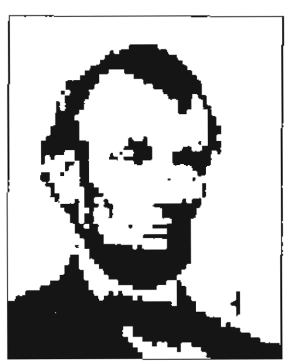


(b)

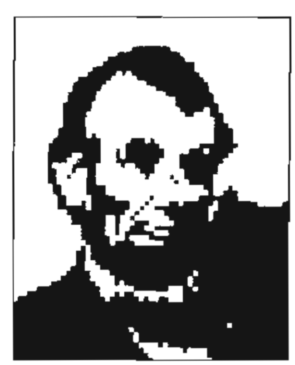
(c)



(d)

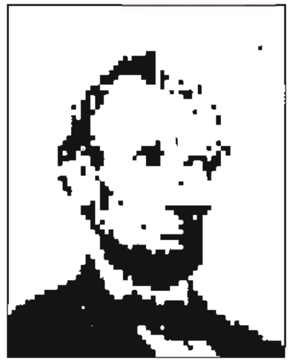


(e), (g)

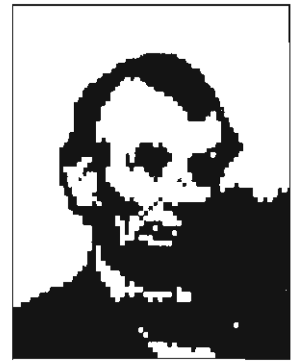


(f), (j)

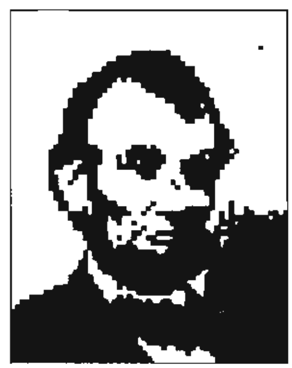
Fig. 4. Lincoln image. (a) Input, (b) histogram Segmentation result of (c) Algorithm 1, (d) Algorithm 2, (e) Algorithm 3, (f) Algorithm 4, (g) Algorithm 5, (h) Algorithm 6, (i) algorithm of Pun, (j) algorithm of Kapur *et al.*, (k) algorithm of Otsu.



(h)



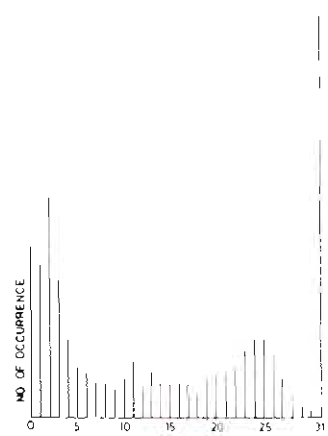
(i)



(k)

Fig. 4. Cont'd.

(a)



(b)

Fig. 5. Boy image. (a) Input, (b) histogram. Segmentation result of (c) Algorithm 1, (d) Algorithm 2, (e) Algorithm 3, (f) Algorithm 4, (g) Algorithm 5, (h) Algorithm 6, (i) algorithm of Pun, (j) algorithm of Kapur *et al.*, (k) algorithm of Otsu.

(c), (e)

(d)

(f)

(g)

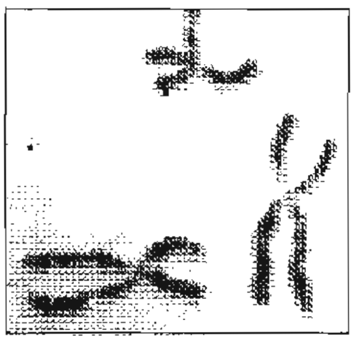
(h)

(i), (k)

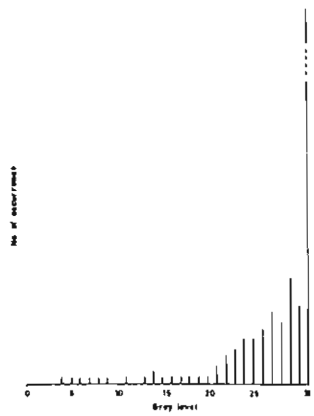
Fig. 5. Cont'd.

(j)

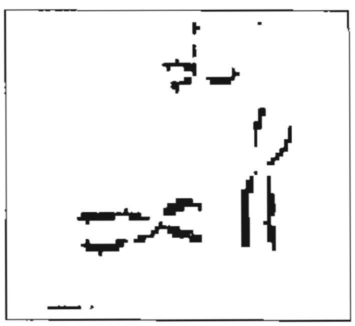
Fig. 5. Cont'd.



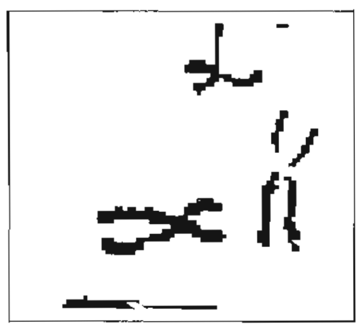
(a)



(b)

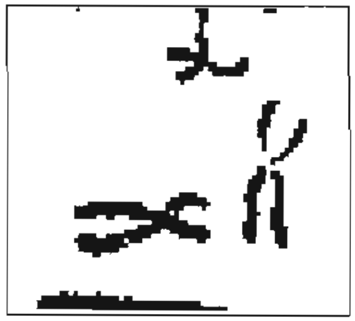


(c)



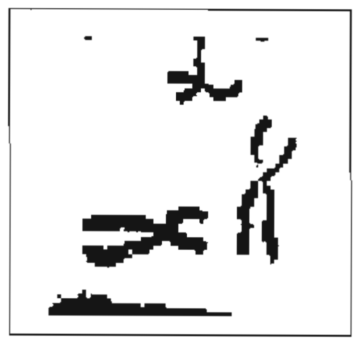
(d), (g)

Fig. 6. Chromosomes image. (a) Input, (b) histogram. Segmentation result of (c) Algorithm 1, (d) Algorithm 2, (e) Algorithm 3, (f) Algorithm 4, (g) Algorithm 5, (h) algorithm of Pun, (i) algorithm of Kapur *et al.*, (j) algorithm of Otsu.



(e), (f)

(h)



(i), (j)

Fig. 6. Cont'd.

the gray levels in a digital image.

The exponential entropic methods (Algorithms 1 and 3) are found to be much better than all other methods. This shows that for a digital image the exponential entropy is possibly a more appropriate measure of information than Shannon's entropy.

The minimum chi-square thresholding principle is an excellent tool for segmentation provided an appropriate distribution is assumed.

7. CONCLUSION

An ideal-image model for a gray tone digital image has been proposed. It has been found that the distribution of gray levels in an image follows the Poisson distribution. Based on this concept, various parametric algorithms for object background classification have been formulated. Two principles, namely, entropy maximization and χ^2 minimization have been adopted here while formulating the algorithms. The thresholds obtained are found to be highly satisfactory for a wide range of input images.

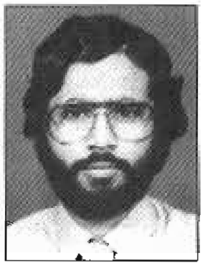
In order to strengthen the appropriateness of the Poisson distribution, the same algorithms have been implemented with the normal distribution. In a part of the experiment, the results have also been compared with those of two entropic thresholding algorithms which do not assume any parametric distribution, and the method of Otsu. In all cases, the algorithms based on the Poisson distribution consistently resulted in a better segmentation. It has further been revealed that the exponential entropic measure is more effective than the logarithmic (Shannon's) entropy in extracting thresholds. The χ^2 statistic can be regarded as a very effective tool for segmentation when the Poisson distribution is used.

It is to be mentioned that the algorithms described here are all based on the global information (i.e. histogram) of an image. Usually, histogram-based algorithms are less effective than local (spatial) information-based algorithms. However, the segmentation produced here by the Poisson distribution-based methods is no way worse than that of the spatial information-based methods described in Refs. 5 and 6.

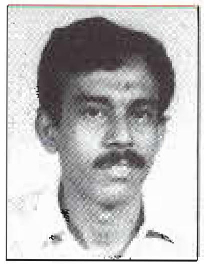
REFERENCES

1. J. S. Weszka and A. Rosenfeld, "Histogram modifieation for threshold selection", *IEEE Trans. Syst. Man Cybern.* **9** (1979) 38–52.
2. N. Otsu, "A threshold selection method from gray-level histogram", *IEEE Trans. Syst. Man Cybern.* **9**, 1 (1979) 62–66.
3. J. Kittler and J. Illingworth, "Minimum error thresholding", *Pattern Recogn.* **19**, 1 (1986) 41–47.
4. P. K. Sahoo, S. Soltani, A. K. C. Wong and Y. C. Chen, "A survey of thresholding techniques", *Comput. Graph. Vision Image Process.* **41** (1988) 233–260.
5. N. R. Pal and S. K. Pal, "Object background segmentation using new definitions of entropy", *IEE Proc.*, Vol. 136, No. 4, Pt. E, 1989, pp. 284–295.
6. N. P. Pal and S. K. Pal, "Entropic thresholding", *Signal Process.* **16** (1989) 97–108.
7. T. Pun, "A new method for gray-level picture thresholding using the entropy of the histogram", *Signal Process.* **2** (1980) 223–237.
8. J. N. Kapur, P. K. Sahoo and A. K. C. Wong, "A new method for gray level picture thresholding using the entropy of the histogram", *Comput. Graph. Vision Image Process.* **29** (1985) 273–285.
9. A. K. C. Wong and P. K. Sahoo, "A gray-level threshold selection method based on maximum entropy principle", *IEEE Trans. Syst. Man Cybern.* **19** (1989) 866–871.
10. E. L. Hall, *Computer Image Processing and Recognition*, Academic Press, New York, 1979, Chap. 7, pp. 381–382.
11. J. C. Dainty and R. Shaw, *Image Sciences*, Academic Press, New York, 1974, Chap. 1, pp. 1–31.
12. C. E. Shannon, "A mathematical theory of eommunication", *Bell Syst. Tech. J.* **27**, July (1948) 379–423.
13. C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, The University of Illinois Press, Urbana, 1949.
14. C. R. Rao, *Linear Statistical Inference and Applications*, Wiley Eastern Pte. Ltd., New Delhi, 1974.

Received 2 June 1989; revised 23 October 1990.



Nikhil R. Pal obtained the B.Sc. (Hons.) in physics and the M.B.M. (Operations Research) in 1979 and 1982, respectively, from the University of Calcutta, India. In 1984 he received the M.Tech. in eompnter seience from the Indian Statistical Institute, Calcutta. At present he is associated with the Computer Science Unit of the Indian Statistieal Institute, Calcutta. He is also a guest lecturer at the University of Calcutta. His researeh interests include image processing, artificial intelligence, fuzzy sets and systems, and neural networks.



Sankar K. Pal obtained the B.Sc. (Hons.) in physics, and the B.Tech., M.Tech. and Ph.D. in radiophysies and electronics in 1969, 1972, 1974, and 1979 respectively, from the University of Calcutta, India. In 1982 he received the Ph.D. in electrical engineering along with the DIC from Imperial College, University of London, England. He was a recipient of the Commonwealth Seholarship in 1979 and the MRC (UK) Post-Doetoral Award in 1981 to work at Imperial College, London. In 1986 he was awarded the Fulbright Post-Doetoral Visiting Fellowship to work at the University of California, Berkeley and the University of Maryland, College Park. In 1989, he received the NRC-NASA Senior Research Award to work at NASA Johnson Space Center, Houston, Texas.

He is at present a Professor in the Electronics and Communication Seiences Unit and the Professor-in-Charge of the Physical and Earth Sciences Division, at the Indian Statistical Institute, Calcutta. His research interests are mainly pattern reeognition, image processing, artifieial intelligence, and fuzzy sets and systems. He is co-author of the book *Fuzzy Mathematical Approach to Pattern Recognition* (John Wiley & Sons, Halsted Press, 1986) which received the Best Production Award in the 7th World Book Fair, New Delhi. He has more than a hundred research papers including nine in edited books and about sixty in international journals to his credit. He has lectured at different U.S. and Japanese universities/laboratories on his research work.

Dr. Pal is a reviewer for *Mathematical Reviews* in the fields of Fnzzy Sets, logic and applications, a Senior Member of the IEEE, a Fellow of the IETE (India) and Treasurer of the Indian Society for Fuzzy Mathematics and Information Processing (ISFUMIP).