Signal Processing 16 (1989) 97-108
North-Holland


# ENTROPIC THRESHOLDING

Nikhil R. PAL
*Computer Science Unit, Indian Statistical Institute, Calcutta 700 035, India*

Sankar K. PAL
*Electronics and Communication Sciences Unit, Indian Statistical Institute, Calcutta 700 035, India*

Received 25 November 1987
Revised 29 February 1988

**Abstract.** Entropy of order $q$ (depending on the information contained in a sequence of gray levels of length $q$) and the conditional entropy of an image are defined. The existing definition of global entropy is found to be its special case when $q = 1$.
Using these definitions, two algorithms for image thresholding (object-background classification) are formulated and implemented with the help of its co-occurrence matrix. The superiority of these algorithms is experimentally demonstrated for a set of images.

**Zusammenfassung.** Die Entropie der Ordnung $q$ (abhängig von der Information, die in einer Folge von Grauwerten der Länge $q$ enthalten ist) und die bedingte Entropie eines Bildes werden definiert. Die bekannte Definition der globalen Entropie erscheint als Spezialfall mit der Ordnung $q = 1$.
Unter Verwendung dieser Definitionen werden zwei Algorithmen zur Klassifikation von Bildobjekt und Hintergrundrauschen durch Anlegen von Schwellen im Bild formuliert und realisiert mit Hilfe der Matrix der Verbund-Ereignisse ("co-occurence matrix"). Die Überlegenheit dieser Algorithmen wird dur eine Reihe von Bildern experimentell nachgewiesen.

**Résumé.** On définit dans cet article l'entropie d'ordre $q$ (c'est-à-dire dépendant d'une séquence de niveaux de gris de longueur $q$) et l'entropie conditionnelle d'une image. On montre que la définition d'entropie globale, telle qu'elle existe actuellement, correspond au cas particulier $q = 1$.
A partir de ces définitions sont formulés deux algorithmes pour le seuillage d'images (classification objet-fond); ils sont implémentés à l'aide de la matrice de co-oceurence. La supériorité de ces algorithmes est démontrée expérimentalement sur une série d'images.

**Keywords.** Entropy, segmentation, co-occurrence matrix

## 1. Introduction

The entropy of a system as defined by Shannon [9, 10] gives a measure of our ignorance about its actual structure. Shannon's function is based on the concept that information gained from an event is inversely related to its probability of occurrence.

Several authors [3, 7, 8] have used this concept to image processing problems. Pun [7, 8] used Shannon's concept to define the entropy of an image assuming that an image is entirely represented by its gray level histogram only. He used this concept to derive an expression for an upper bound of the a posteriori entropy and finally used it to segment an image into object and background. Kapur et al. [3] have found some flaws in Pun's derivation and also used a similar concept to partition the image into object and background. They, instead of considering one probability distribution for the entire histogram, used two separate probability distributions; one for the object and the other for the background. The total entropy of the image is then maximised to arrive at the threshold for segmentation.

It should be mentioned here that these entropy-based methods were developed without highlighting the adequateness of the concept of Shannon's entropy in the case of an image. For example, the dependency of pixel intensities in an image and hence its spatial distribution are not taken into account in defining its entropy. As a result, different images with identical histograms will always result in the same entropic value and same threshold. This is, of course, not intuitively acceptable. Moreover, in the algorithm of Pun [7] the maximisation of an upper bound of the a posteriori entropy, to avoid the trivial result with the a posteriori entropy, is not justified.

The present work attempts to formulate two other definitions of entropy, namely entropy of order $q$, $q = 1, 2, \ldots,$ and the conditional entropy of an image. Entropy of order unity ($q = 1$) corresponds to the global entropy as used by Pun [7, 8] and Kapur et al. [3]. Higher order ($q > 1$) entropies take into account the information contained in a sequence of gray levels of length $q$ and thus denote the various local entropies of an image. The conditional entropy, on the other hand, gives an average amount of information that may be obtained from a region, when one has viewed the remaining portion of the image.

These new concepts are then introduced to develop two algorithms for the object-background image classification (segmentation) problem.

The effectiveness of these algorithms is demonstrated for a set of images having different histograms and their superiority in performance over those of Pun [7] and Kapur et al. [3] is also established.

## 2. Entropic measures for image processing

Shannon [9, 10] defined the entropy of a system as a function of the probability of occurrence of different states of the system. If a system has $n$ different states with probability of occurrence $p_i$, $i = 1, 2, \ldots, n$, $\sum_{i=1}^{n} p_i = 1$, then the gain in information from the occurrence of the event $i$ is defined as $\Delta I = -\log_2 p_i$.

The expected value of such a gain in information is defined as the entropy of the system. Thus, the entropy $H$ of the system is:

$$H = -\sum_{i=1}^{n} p_i \log_2 p_i. \tag{1}$$

Based on the concept of Shannon's entropy, different authors have defined the entropy for an image. Let us discuss here those measures and the associated problems when applied to image processing.

Let $F = [f(x, y)]_{P \times Q}$ be an image of size $P \times Q$, where $f(x, y)$ is the gray value at $(x, y)$; $f(x, y) \in G_L = \{0, 1, \ldots, L-1\}$, the set of gray levels. Let $N_i$ be the frequency of the gray level $i$. Then $\sum_{i=0}^{L-1} N_i = P \times Q = N$ (say).

Pun [7, 8] and Kapur et al. [3] considered the gray level histogram of $F$ an $L$-symbol source, independently from the underlying image. In addition to this, they also assumed that these symbols are statistically independent.

Following Shannon's definition of entropy (equation (1)), Pun [7] defined the entropy of the image (histogram) as

$$H = -\sum_{i=0}^{L-1} p_i \log_2 p_i; \qquad p_i = N_i / N, \tag{2}$$

for the image segmentation problem.

### 2.1. 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:

$$\begin{aligned} H'_L(s) &= -P_B \log_2 P_B - P_W \log_2 P_W \\ &= -P_S \log_2 P_S \\ &\quad -(1 - P_S) \log_2 (1 - P_S) \end{aligned} \tag{3}$$

as

$$P_S = \sum_{i=0}^{s} p_i = P_B \quad \text{and} \quad P_W = 1 - P_S. \tag{4}$$

Since the maximisation of $H'_L$ gives the trivial result of $P_S = 1/2$, Pun [7] maximised an upper bound $g(s)$ of $H'_L(s)$, where

$$g(s) = \frac{H'_B \cdot \log_2 P_S}{H_L \cdot \log_2[\max(p_0, p_1, \ldots, p_s)]} + \frac{(H_L - H'_B)\log_2(1 - P_s)}{H_L \cdot \log_2[\max(p_{s+1}, p_{s+2}, \ldots, p_{L-1})]}, \tag{5}$$

where

$$H_L = -\sum_{i=0}^{L-1} p_i \log_2 p_i \tag{6}$$

and

$$H'_B = -\sum_{i=0}^{s} p_i \log_2 p_i. \tag{7}$$

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

### 2.2. *Method of Kapur, Sahoo and Wong* [3]

Recently, Kapur et al. have also used Shannon's concept of entropy but from a different point of view. They, instead of considering one probability distribution of the entire image, considered two probability distributions, one for the object and the other for the background. The sum of the individual entropy of the object and background is then maximised. In other words, this will result in equiprobable gray levels in each region and thus maximises the sum of homogeneities in gray levels within object and background by making the gray levels equiprobable in either region.

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

$$H_B^{(s)} = -\sum_{i=0}^{s} \frac{p_i}{P_S} \log_2(p_i/P_S) \tag{8}$$

and that of the white portion is

$$H_W^{(s)} = -\sum_{i=s+1}^{L-1} \frac{p_i}{1-P_S} \log_2(p_i/(1-P_S)). \tag{9}$$

The total entropy of the image is then defined as

$$H_T^{(s)} = H_B^{(s)} + H_W^{(s)}. \tag{10}$$

In order to select the threshold they maximised $H_T^{(s)}$. In other words, the value of $s$ which maximises $H_T^{(s)}$ gives the threshold for object and background classification.

### 2.3. *Some remarks*

All the methods [3, 7, 8] discussed so far virtually assume that an image is entirely represented only by its histogram. Thus, different images with identical histograms will result in same entropic value in spite of their different spatial distributions of gray levels. This is, of course, not intuitively appealing. For example, consider Figs. 1 and 2 where "dotted", "shaded" and "blank" pixels represent the gray levels $l_1$, $l_2$ and $l_3$ ($l_1 < l_2 < l_3$) respectively. Both Fig. 1 and Fig. 2 have identical

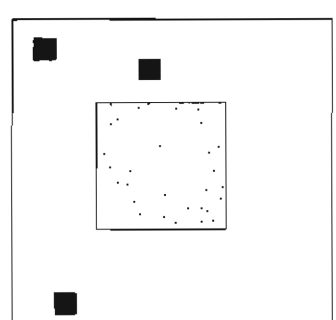

Fig. 1. Two-tone image

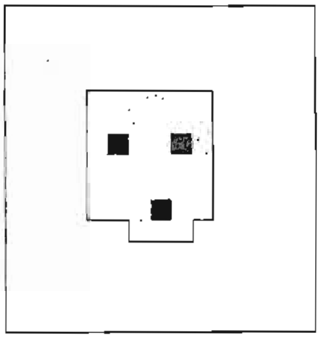

Fig. 2. Two-tone image [histogram identical to that of Fig. 1 but different spatial distribution].

histograms but different spatial distributions of gray levels. As a result, the entropy (information content) of Fig. 1 and Fig. 2 is expected to be different.

Under those definitions all images with identical histograms but different spatial distributions of gray levels will therefore give rise to the same threshold value. Our experience and intuition also do not support this. For example, $l_1$ should be the threshold in Fig. 1, whereas it is $l_2$ in Fig. 2 for object-background classification.

In the algorithm of Pun [7], the concept of maximisation of the upper bound of the evaluation function $g(s)$ (equation (5)) for object-background classification is not justified. For example, the maximum value of equation (5) may even correspond to a minimum value of the a posteriori entropy (equation (3)).

Moreover, all these methods have used Shannon's concept of entropy in image processing without highlighting its adequateness in the case of an image.

## 3. Entropy of an image

### 3.1. Global and local entropy

We know that in an image pixel intensities are not independent of each other. This dependency of pixel intensities can be incorporated by considering sequences of pixels to estimate the entropy. In order to arrive at the expression of entropy of an image the following theorem due to Shannon [2, 10] can be stated.

**Theorem.** *Let $p(s_i)$ be the probability of a sequence $s_i$ of gray levels of length $q$, where a sequence $s_i$ of length $q$ is defined as a permutation of $q$ gray levels. Let us define*

$$H^{(q)} = -\frac{1}{q}\sum_i p(s_i)\log_2 p(s_i), \tag{11}$$

*where the summation is taken over all gray level sequences of length $q$. Then $H^{(q)}$ is a monotonic decreasing function of $(q)$ and $\lim_{q\to\infty} H^{(q)} = H$, the entropy of the image.*

For different values of $q$ we get various orders of entropy.

*Case 1*: $q = 1$, i.e. sequence of length one. If $q = 1$ we get

$$H^{(1)} = -\sum_{i=0}^{L-1} p_i \log_2 p_i, \tag{12}$$

where $p_i$ is the probability of occurrence of the gray level $i$.

Such an entropy is a function of the histogram only and it may be called the "global entropy" of the image. Therefore, different images with identical histograms would have same $H^{(1)}$ value irrespective of their contents. The definitions used by Pun [7] and Kapur et al. [3] in fact belong to Case 1.

*Case 2*: $q = 2$, i.e. sequences of length two. Hence,

$$H^{(2)} = -\tfrac{1}{2}\sum_i p(s_i)\log_2 p(s_i),$$

where $s_i$ is a sequence of gray level of length two,

$$= -\tfrac{1}{2}\sum_i\sum_j p_{ij}\log_2 p_{ij}, \tag{13}$$

where $p_{ij}$ is the probability of co-occurrence of the gray levels $i$ and $j$. Therefore, $H^{(2)}$ can be obtained from the co-occurrence matrix, which will be defined in Section 4.1.

$H^{(2)}$ takes into account the spatial distribution of gray levels. Therefore, two images with identical histograms but different spatial distributions will result in different entropy, $H^{(2)}$ values. Expressions for higher order entropies ($q > 2$) can also be deduced in a similar manner. $H^{(i)}$, $i \geq 2$, may be called the "local entropy" of order $i$ of an image.

### 3.2. *Conditional entropy*

Suppose an image has two distinct portions, the object $X$ and the background $Y$. Suppose the object consists of the gray levels $\{x_i\}$ and the background contains the gray levels $\{y_i\}$. The conditional entropy of the object $X$ given the background $Y$, i.e. the average amount of information that may be obtained from $X$ given that one has viewed the background $Y$, can be defined as

$$H(X/Y) = -\sum_{x_i \in X}\sum_{y_j \in Y} p(x_i/y_j) \times \log_2 p(x_i/y_j). \tag{14}$$

Similarly, the conditional entropy of the background $Y$ given the object $X$ is defined as

$$H(Y/X) = -\sum_{y_j \in Y}\sum_{x_i \in X} p(y_j/x_i) \times \log_2 p(y_j/x_i). \tag{15}$$

The pixel $y_j$, in general, can be an $m$th order neighbour of the pixel $x_i$, i.e. $y_j$ can be the $m$th pixel after $x_i$. Since the estimation of such a probability is very difficult, we impose another constraint on $x_i$ and $y_j$ of equations (14) and (15). In addition to $x_i \in X$ and $y_j \in Y$, we also impose the restriction that $x_i$ and $y_j$ must be adjacent pixels. Thus, equations (14) and (15) can be rewritten as

$$H(X/Y) = -\sum_{\substack{x_i \in X \\ (x_i, y_j)\ \text{adjacent}}}\sum_{y_j \in Y} p(x_i/y_j) \times \log_2 p(x_i/y_j) \tag{16}$$

and

$$H(Y/X) = -\sum_{\substack{y_j \in Y \\ (y_j, x_i)\ \text{adjacent}}}\sum_{x_i \in X} p(y_j/x_i) \times \log_2 p(y_j/x_i). \tag{17}$$

The conditional entropy of the image can, therefore, be defined as

$$H^{(c)} = (H(X/Y) + H(Y/X))/2, \tag{18}$$

where $X$ and $Y$ represent the object and background of the image, respectively.

## 4. Application to image segmentation

Based on the new definitions of entropy of an image, the following two algorithms for object background classification are proposed.

### 4.1. *Algorithm 1*

We are now going to describe an algorithm based on equation (13), which takes into account the spatial details of an image. Since such a method is dependent on the probability of co-occurrence of pixel intensities, let us define, first of all, the co-occurrence matrix before proceeding further.

The co-occurrence matrix of the image $F$ is an $L \times L$ dimensional matrix $T = [t_{ij}]_{L \times L}$ that gives an idea about the transition of intensities between adjacent pixels. In other words, $t_{ij}$, the $(i,j)$th entry of the matrix, gives the number of times the gray level $j$ follows the gray level $i$ in some particular fashion. Depending upon the ways in which the gray level $i$ follows gray level $j$, different definitions of co-occurrence matrix are possible.

Here, we made the co-occurrence matrix asymmetric by considering the horizontally right and vertically lower transitions. It has been found experimentally [1, 4] that the consideration of a symmetric co-occurrence matrix by taking into account both right-left and upper-lower transitions needs some additional computations without changing much the information content in it. Furthermore, consideration of both horizontal and vertical transitions allows all the edges to participate in threshold selection.

Thus, $t_{ij}$ is defined as follows [4]:

$$t_{ij} = \sum_{l=1}^{P}\sum_{k=1}^{Q} \delta, \tag{19a}$$

where

$$\delta = 1, \quad \text{if} \begin{cases} f(l,k)=i \quad \text{and} \quad f(l,k+1)=j \\ \qquad\qquad \text{or} \\ f(l,k)=i \quad \text{and} \quad f(l+1,k)=j; \end{cases}$$

$\delta = 0$, otherwise.

The probability of co-occurrence $p_{ij}$ of gray levels $i$ and $j$ can therefore be written as

$$p_{ij} = t_{ij} \Big/ \left( \sum_i \sum_j t_{ij} \right), \tag{19b}$$

where obviously $0 \leqslant p_{ij} \leqslant 1$.

If $s$, $0 \leqslant s \leqslant L-1$, is a threshold, then $s$ partitions the co-occurrence matrix into four quadrants, namely A, B, C and D (Fig. 3).

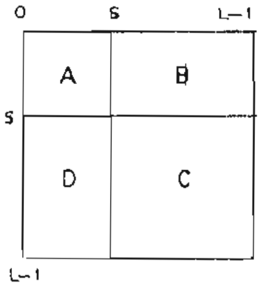


Fig. 3. Quadrants of co-occurrence matrix.

Let us define the following quantities:

$$P_A = \sum_{i=0}^{s} \sum_{j=0}^{s} p_{ij}, \qquad P_B = \sum_{i=0}^{s} \sum_{j=s+1}^{L-1} p_{ij},$$
$$P_C = \sum_{i=s+1}^{L-1} \sum_{j=s+1}^{L-1} p_{ij}, \qquad P_D = \sum_{i=s+1}^{L-1} \sum_{j=0}^{s} p_{ij}. \tag{20}$$

Normalising the probabilities within each individual quadrant, such that the sum of the probabilities of each quadrant equals one, we get the following cell probabilities for different quadrants:

$$p_{ij}^{A} = \frac{p_{ij}}{P_A} = \frac{t_{ij} \Big/ \left( \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} t_{ij} \right)}{\sum_{i=0}^{s} \sum_{j=0}^{s} \left\{ t_{ij} \Big/ \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} t_{ij} \right\}}$$
$$= \frac{t_{ij}}{\sum_{i=0}^{s} \sum_{j=0}^{s} t_{ij}},$$
$$\text{for } 0 \leqslant i \leqslant s,\ 0 \leqslant j \leqslant s. \tag{21}$$

Similarly,

$$p_{ij}^{B} = \frac{p_{ij}}{P_B} = \frac{t_{ij}}{\sum_{i=0}^{s} \sum_{j=s+1}^{L-1} t_{ij}},$$
$$\text{for } 0 \leqslant i \leqslant s,\ s+1 \leqslant j \leqslant L-1; \tag{22}$$

$$p_{ij}^{C} = \frac{p_{ij}}{P_C} = \frac{t_{ij}}{\sum_{i=s+1}^{L-1} \sum_{j=s+1}^{L-1} t_{ij}},$$
$$\text{for } s+1 \leqslant i \leqslant L-1,$$
$$s+1 \leqslant j \leqslant L-1, \tag{23}$$

and

$$p_{ij}^{D} = \frac{p_{ij}}{P_D} = \frac{t_{ij}}{\sum_{i=s+1}^{L-1} \sum_{j=0}^{s} t_{ij}},$$
$$\text{for } s+1 \leqslant i \leqslant L-1,\ 0 \leqslant j \leqslant s. \tag{24}$$

Now with the help of equations (13) and (21), the second-order local entropy of the object can be defined as

$$H_A^{(2)}(s) = -\tfrac{1}{2} \sum_{i=0}^{s} \sum_{j=0}^{s} p_{ij}^{A} \log_2 p_{ij}^{A}. \tag{25}$$

Similarly, the second-order entropy of the background can be written as

$$H_C^{(2)}(s) = -\tfrac{1}{2} \sum_{i=s+1}^{L-1} \sum_{j=s+1}^{L-1} p_{ij}^{C} \log_2 p_{ij}^{C}. \tag{26}$$

Hence the total second-order local entropy of the object and the background can be written as

$$H_T^{(2)}(s) = H_A^{(2)}(s) + H_C^{(2)}(s). \tag{27}$$

The gray level corresponding to the maximum of $H_T^{(2)}(s)$ gives the threshold for object-background classification.

### 4.2. Algorithm 2

This algorithm is based on the concept of conditional entropy (equations (16)-(18)). Suppose $s$ is an assumed threshold. Then pixels with gray level values ranging from 0 to $s$ constitute the object, while the remaining pixels with gray values lying between $s+1$ and $L-1$ correspond to the background. Let $t_{ij}$ be an entry of the quadrant B (Fig. 3), then $t_{ij}$ gives the number of transitions, such that $i$ belongs to the object and $j$ belongs to the background, and $i$ and $j$ are adjacent. Therefore, $p_{ij}^{B}$ as defined in equation (22) gives the probability that gray levels $i$ and $j$ belong to the object and background, respectively, and they are adjacent. Thus $p_{ij}^{B}$s of equation (22) give the probabilities required by equation (16). Similarly, $p_{ij}^{D}$s of equation (24) correspond to the probabilities of equation (17).

Therefore,

$$H\text{ (object/background)}$$
$$= H(\text{O/B}) = -\sum_{i=0}^{s}\sum_{j=s+1}^{L-1} p_{ij}^{B}\log_2 p_{ij}^{B} \qquad (28)$$

and

$$H\text{ (background/object)}$$
$$= H(\text{B/O}) = -\sum_{i=s+1}^{L-1}\sum_{j=0}^{s} p_{ij}^{D}\log_2 p_{ij}^{D}. \qquad (29)$$

Now the conditional entropy of the image is

$$H_T^{(c)} = (H(\text{O/B}) + H(\text{B/O}))/2. \qquad (30)$$

In order to get the threshold for object-background classification $H_T^{(c)}$ is maximised with respect to $s$.

## 5. Implementation and results

The segmentation (object-background classification) algorithms described in Sections 2 and 4 are implemented on a set of four images with widely different types of histogram. Figs. 4(a), 5(a), 6(a) and 7(a) represent the input images while Figs. 4(b), 5(b), 6(b) and 7(b) represent the corresponding gray level histograms. The input images are produced on a line printer by over printing different character combinations for different gray levels. The threshold levels produced by different methods are presented in Table 1.

Fig. 4(a) represents the image of a biplane with two dominant modes in its gray level histogram (Fig. 4(b)). The segmented images produced by Algorithm 1, Algorithm 2, the method of Pun [7] and the method of Kapur et al. [3] are shown in Figs. 4(c)–(f), respectively. From the results one can see that, except for the conditional entropic method (Algorithm 2), the propeller in front of the biplane is lost. In all but Algorithm 2, some portion of the background got mixed up with the object, though the image has two dominant modes. The methods of Pun [7] (Fig. 4(e)) and that of Kapur et al. [3] (Fig. 4(f)) have produced comparable results to that of Algorithm 1.

Figs. 5(a) and 5(b) represent the input image of Abraham Lincoln and its gray level histogram, respectively. The histogram has a number of deep valleys. The thresholds produced by different methods are shown in Table 1 and the corresponding segmented images are shown in Figs. 5(c)–(f). In this case too, all the methods except the conditional entropic method (Algorithm 2) have produced comparable result. The best result is produced by Algorithm 2 (equation (30)) which has

Table 1
Thresholds for object-background classification

| Images | Thresholds | | | |
|---|---|---|---|---|
| | Proposed Algorithm 1 (equation (27)) | Proposed Algorithm 2 (equation (30)) | Algorithm of Pun (equation (5)) | Algorithm of Kapur et al. (equation (10)) |
| Biplane (Fig. 4) | 22 | 12 | 24 | 21 |
| Lincoln (Fig. 5) | 17 | 10 | 16 | 15 |
| Saturn (Fig. 6) | 0 | 17 | 13 | 5 |
| Blurred chromosome (Fig. 7) | 33 | 41 | 32 | 33 |

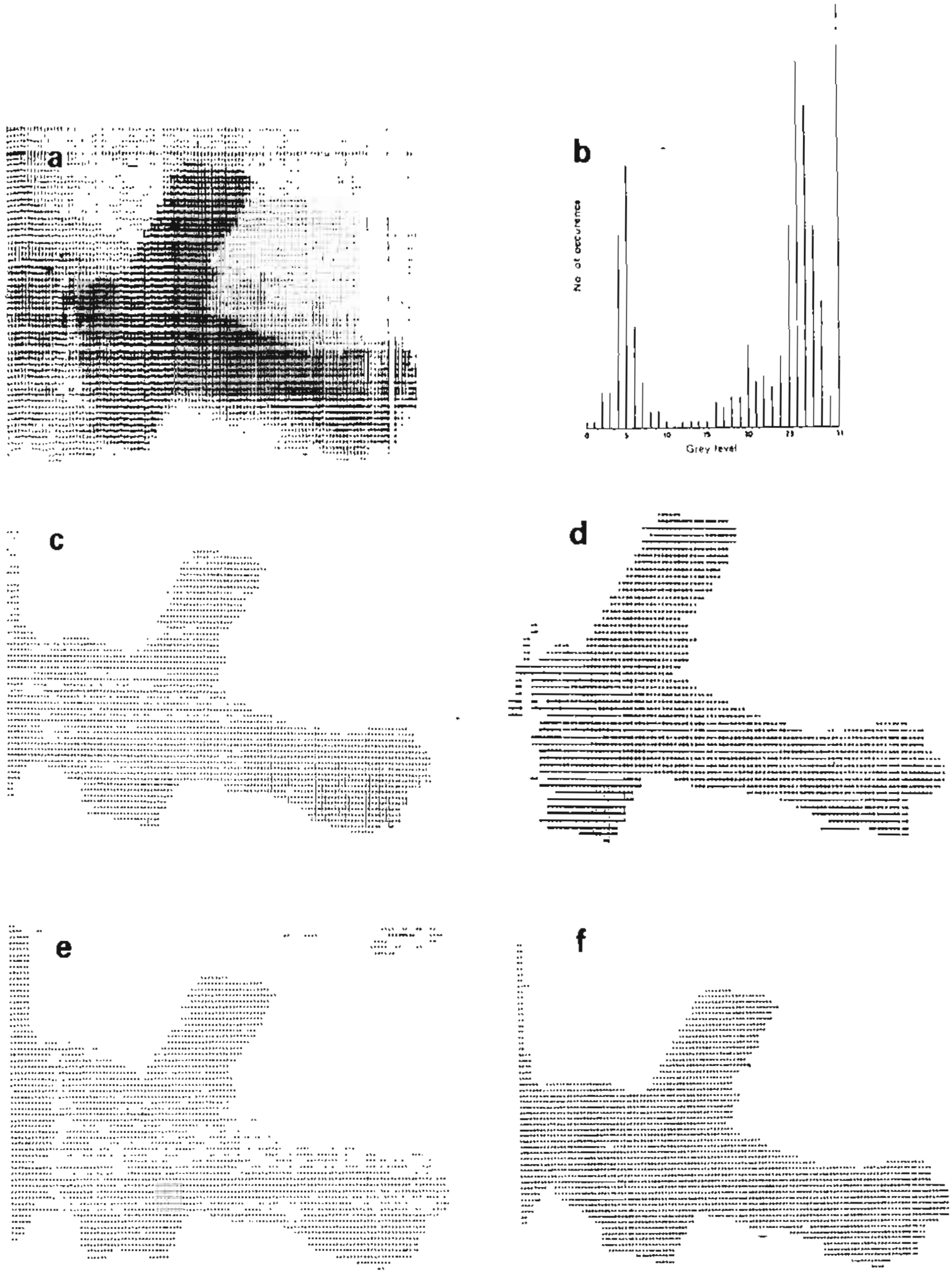


Fig. 4. Biplane image. (a) Input; (b) histogram; (c) proposed Algorithm 1; (d) proposed Algorithm 2; (e) algorithm of Pun; (f) algorithm of Kapur et al.

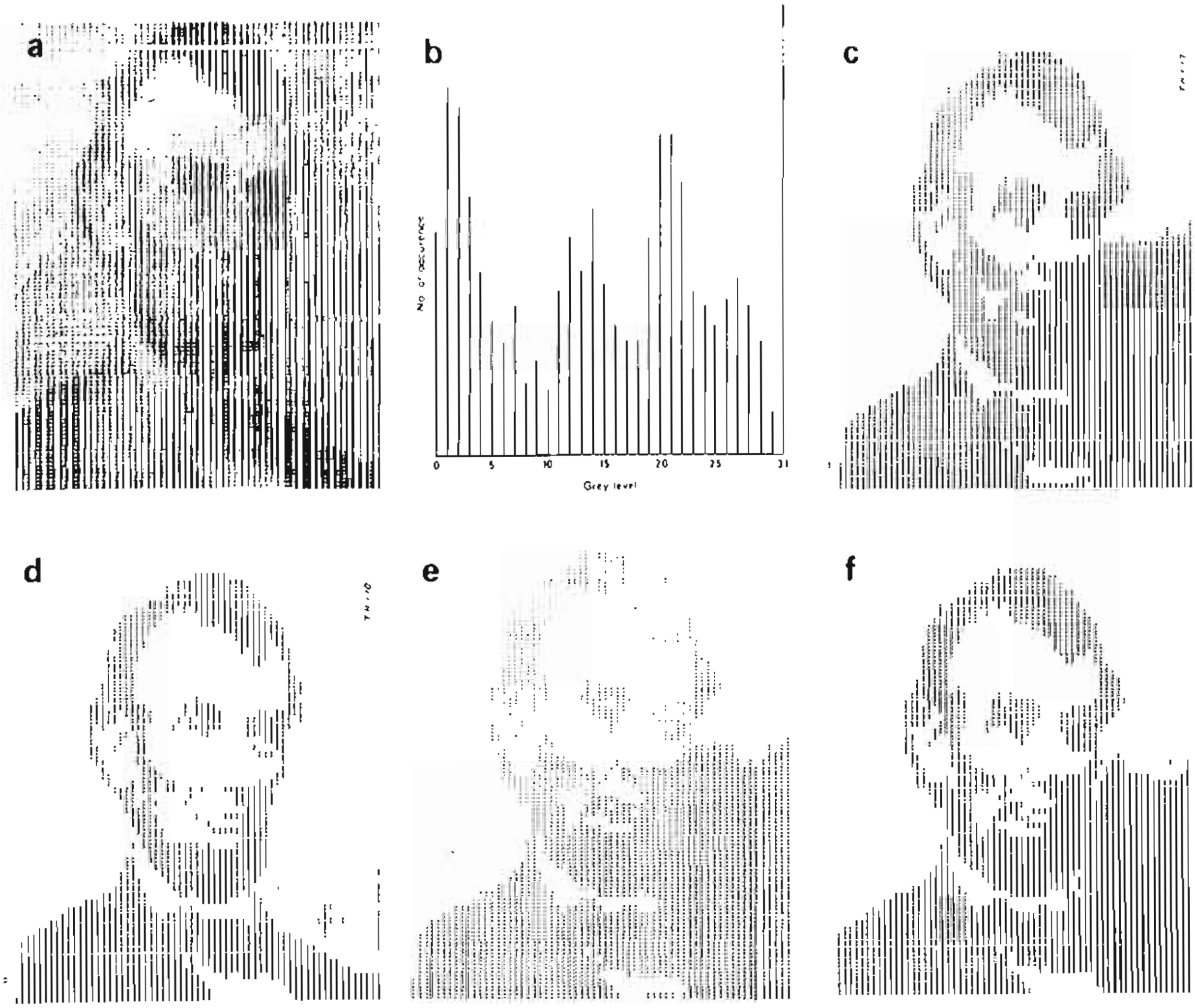


Fig. 5. Lincoln image. (a) Input; (b) histogram, (c) proposed Algorithm 1, (d) proposed Algorithm 2, (e) algorithm of Pun; (f) algorithm of Kapur et al.

clearly separated the object from the background. All other methods failed to discriminate between the beard and the background at the south-east corner of the image.

Let us now consider an image of Saturn (Fig. 6(a)). The characteristics of its histogram are depicted in Fig. 6(b). More or less, all the algorithms are found here to partition well Saturn from the background. For Algorithm 1 and the method of Kapur et al. [3] a small portion of the background got mixed up with the right side of the ring. On the other hand, the thin shade on the left side of the ring is made visible by Algorithm 2 and the method of Pun [7]. Of these two, the boundaries of Saturn produced by Algorithm 2 are more smooth and thus closely resemble that of the input.

The algorithms are also tested on an image of a blurred chromosome (Fig. 7(a)) having a bimodal histogram (Fig. 7(b)). Here too, all the methods except the conditional entropic method (Algorithm 2) have produced similar results. However, the

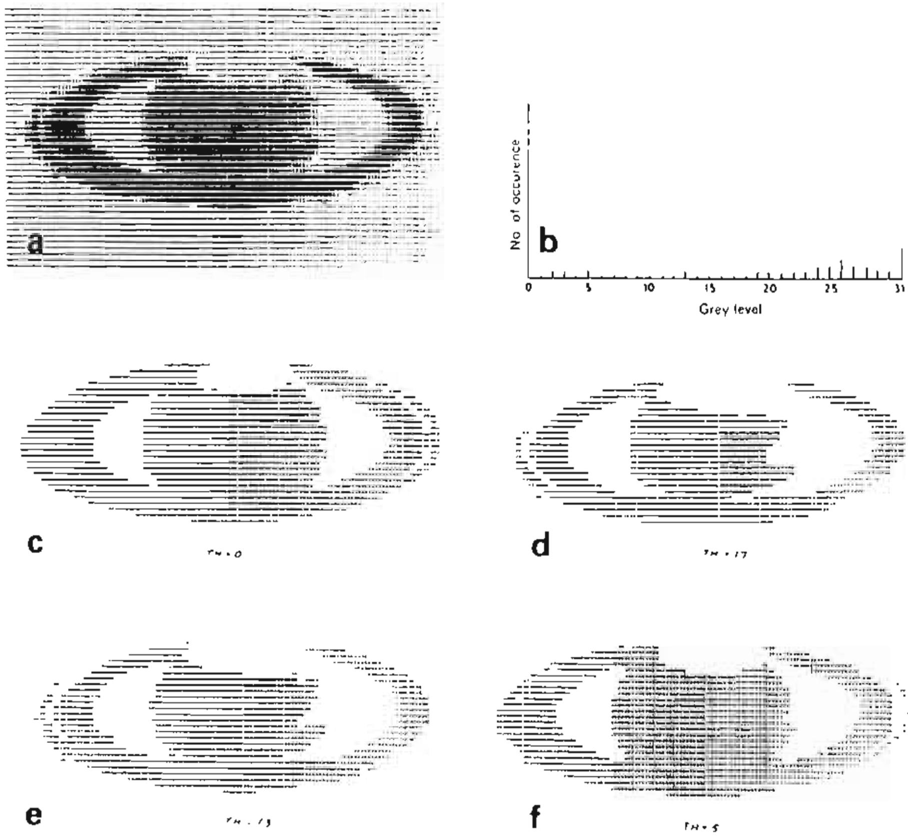


Fig. 6. Saturn image. (a) Input; (b) histogram; (c) proposed Algorithm 1; (d) proposed Algorithm 2; (e) algorithm of Pun; (f) algorithm of Kapur et al.

best classification is done by Algorithm 2. This also conforms well to the recent work of Pal and Rosenfeld [5].

## 6. Conclusion

The concept of different order local entropy and conditional entropy of an image are introduced. The first-order local entropy corresponds to global entropy, as used by Pun [7, 8] and by Kapur et al. [3].

Two algorithms for object-background classification are proposed whereby it is possible to segment/extract the object from the background. The results are compared with those of the existing entropic thresholding methods and are found to be superior for a wide class of images.

The proposed algorithm can be extended to multithresholding problems using the concept of conditional entropy for more than two classes. For example, if $S_R = \{s_1, s_2, \ldots, s_R\}$ denotes $R$ thresholds in an image with $(R+1)$ regions, then $S_R$ partitions the co-occurrence matrix of the image into $(R+1) \times (R+1)$ blocks. Define $p_{ij}^{lm}$, $l, m = 1, 2, \ldots, R+1$, the normalised probabilities for the block $(l, m)$ as

$$p_{ij}^{lm} = \frac{t_{ij}}{\sum_{i=s_{l-1}+1}^{s_l} \sum_{j=s_{m-1}+1}^{s_m} t_{ij}},$$

$$s_{l-1} < i \leq s_l,\ s_{m-1} < j \leq s_m. \tag{31}$$

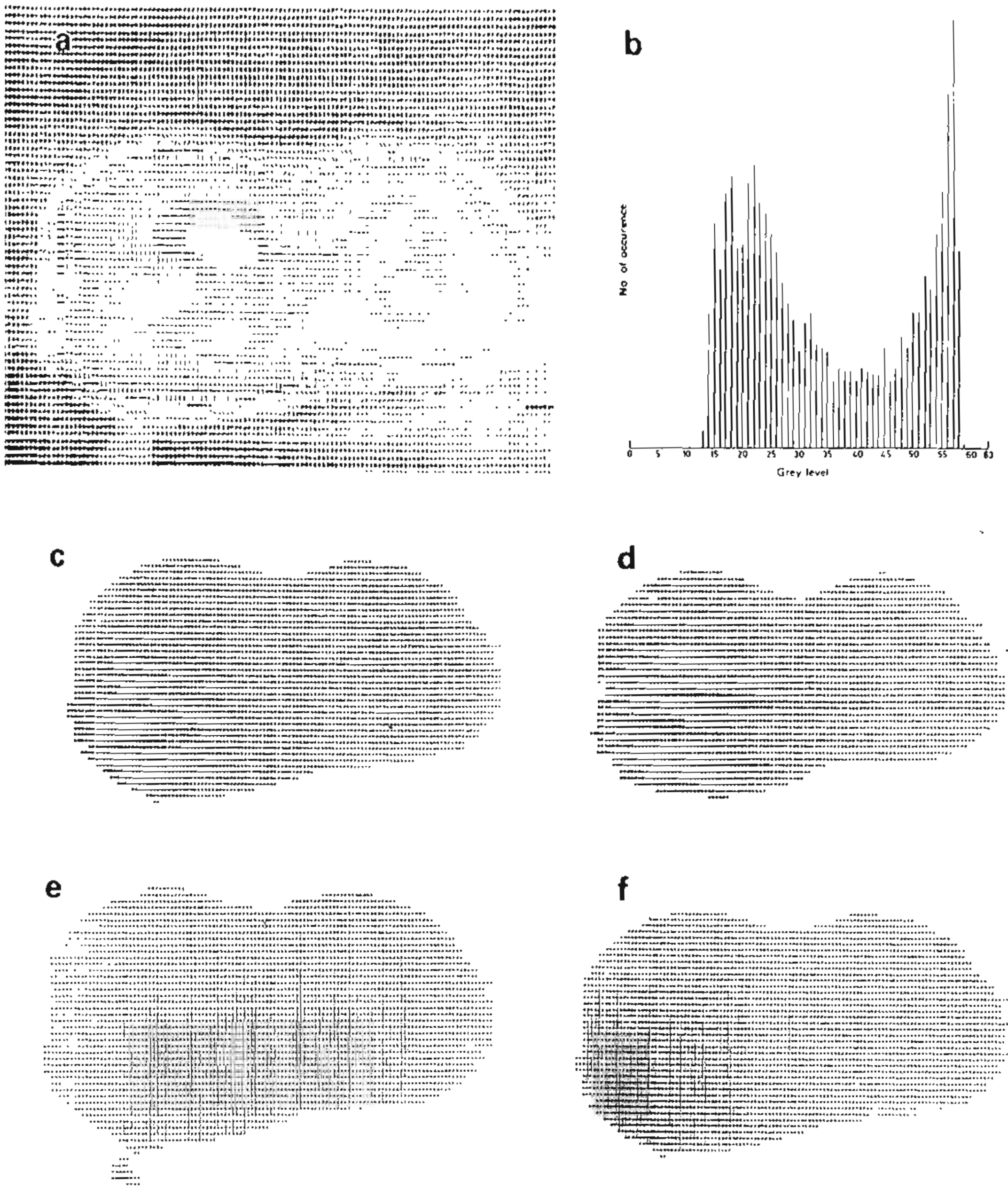


Fig. 7. Blurred chromosome image. (a) Input; (b) histogram; (c) proposed Algorithm 1; (d) proposed Algorithm 2; (e) algorithm of Pun; (f) algorithm of Kapur et al.

with $s_0 = 0$ and $s_{R+1} = L-1$. The total conditional entropy of the partitioned image (into $R+1$ regions) will then be

$$H_T^{(c)}(S_R) = -\sum_{l=1}^{R+1} \sum_{\substack{m=1 \\ m \neq l}}^{R+1} \sum_{i=s_{l-1}+1}^{s_l} \sum_{j=s_{m-1}+1}^{s_m} \times p_{ij}^{lm} \log_2 p_{ij}^{lm}. \tag{32}$$

The vector $S_R = \{s_1, s_2, \ldots, s_R\}$ which maximises $H_T^{(c)}(S_R)$ can be taken as the set of thresholds for multiple segmentation of the image.

Furthermore, the effect of noise on the performance of the algorithms, the quantitative measures for objective evaluation of the resulting thresholded images, and establishment of a link between the entropies presented here and some of the various Markovian models [6], may constitute a part of future investigation.

## Acknowledgements

The authors gratefully acknowledge Mr. J. Dasgupta for typing the manuscript and Prof. D. Dutta Majumder for his interest in this work.

## References

[1] F. Deravi and S.K. Pal, "Grey level thresholding using second-order Statistics", *Pattern Recognition Letts.*, Vol. 1, July 1983, pp. 417-422.

[2] E.L. Hall, *Computer Image Processing and Recognition*, Academic Press, New York, 1979.

[3] J.N. Kapur, P.K. Sahoo and A.K.C. Wong, "A new method for grey-level picture thresholding using the entropy of the histogram", *Comp. Graphics, Vision and Image Proc.* 29, 1985, pp. 273-285.

[4] S.K. Pal and N.R. Pal, "Segmentation using Contrast and homogeneity measures", *Pattern Recognition Letts.*, Vol 5, 1987, pp 293-304.

[5] S K. Pal and A. Rosenfeld, "Image enhancement and thresholding by optimization of fuzzy compactness", *Patern Recognition Letts.*, Vol. 7, 1988, pp. 77-86.

[6] D.K. Pickard, "A curious binary lattice process", *J Appl. Prob.*, Vol. 14, 1977, pp. 717-731.

[7] T. Pun, "A new method for grey-level picture thresholding using the entropy of the histogram", *Signal Processing*, Vol. 2, 1980, pp. 223-237.

[8] T. Pun, "Entropic thresholding, a new approach", *Comp. Graphics and Image Proc.*, Vol. 16, 1981, pp. 210-239.

[9] C.E. Shannon, "A mathematical theory of communication", *Bell Syst. Tech. J.*, Vol. 27, July 1948, pp. 379-423.

[10] C.E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, The University of Illinois Press, Urbana, 1949.