Pattern Recognition Letters 5 (1987) 293-304
North-Holland

April 1987

Segmentation using contrast and homogeneity measures

S.K. PAL
Electronics & Communication Sciences Unit, Indian Statistical Institute, Calcutta 700035, India

N.R. PAL
Dunlop India Ltd., Sahaganj, Hooghly 712104, India

Received 14 July 1986

Abstract: The present work describes an algorithm for automatic image segmentation using a 'homogeneity' measure and a 'contrast' measure defined on the co-occurrence matrix of the image. The measure of contrast involves the concept of logarithmic response (adaptibility with background intensity) of the human visual system (HVS). A merging algorithm is also introduced in order to remove the undesirable thresholds.

Effectiveness of the algorithm and the comparison of its performance with the existing ones are demonstrated for a set of images.

Key words: Image processing, homogeneity and contrast measures, human visual system, co-occurrence matrix.

1. Introduction

One of the key problems of scene analysis is segmentation of a scene into different regions. Segmentation is essentially a pixel classification problem where one tries to classify the pixels into different classes such that each class is homogeneous and at the same time the union of no two adjacent classes is homogeneous. In other words, given a definition of uniformity, segmentation is a partition of the picture into connected subsets, each of which is uniform, but such that no union of adjacent subsets is uniform [1].

There are several techniques of image segmentation based on global and local information of an image. One of the techniques based on global information is histogram thresholding which selects the valley points as threshold levels. For images where a histogram may not have sharp valleys (i.e., having flat minima or local minima) the histogram is usually sharpened [2] by a suitable transformation so that the task of selecting valley becomes easier. These transformations usually require some parameters whose choices have significant impact in determining the number of thresholds. The co-occurrence matrix, on the other hand, uses local spatial information of an image and provides information regarding the number of transitions between any two gray levels in the image. These information have been used by different authors namely, Weszka and Rosenfeld [3], Deravi and Pal [4] and Chanda et al. [5] for segmentation.

The measures on co-occurrence matrix reported by these authors did not consider the fact of logarithmic response of the human visual system (HVS) [6, 7] in measuring 'contrast' between regions in an image. The present work attempts to bring this factor into consideration while defining a measure of 'contrast' in addition to defining another measure called homogeneity within a region. The combination of these two measures made the algorithm effective in determining threshold levels. Furthermore, a provision is also kept for merging undesirable segments, if generated.

The effectiveness of the algorithm along with its comparison with three other methods [3-5] has been demonstrated on a set of images. A Digital Computer EC-1033 has been used for analysis.

2. Co-occurrence matrix and some measures for segmentation

2.1. Co-occurrence matrix

Let $F=[f(x,y)]$ be an image of size $P\times Q$, where $f(x,y)$ is the gray value at (x,y), $f(x,y)\in G_L=\{0,1,2,\ldots,L-1\}$, the set of gray levels. The co-occurrence matrix of the image F is an $L\times L$ dimensional matrix that gives an idea about the transition of intensity between adjacent pixels. In other words, the (i,j)th entry of the matrix gives the number of times the gray level 'j' follows the gray level 'i' in a specific fashion.

Let 'a' denote the (i,j)th pixel in F and let 'b' be one of the eight neighbouring pixels of 'a', i.e.

$$b\in a_8=\{(i,j-1),(i,j+1),(i+1,j),(i-1,j),(i-1,j-1),(i-1,j+1),(i+1,j-1),(i+1,j+1)\}.$$

Define

$$t_{lk}=\sum_{\substack{a\in F\\ b\in a_8}}\delta$$

where $\delta=1$ if the gray level value of a is l and that of b is k, $\delta=0$ otherwise.

Obviously, t_{lk} gives the number of times the gray level 'k' follows gray level 'l' in any one of the eight directions. The matrix $T=[t_{lk}]_{L\times L}$ is, therefore, the co-occurrence matrix of the image F.

One may get different definitions of the co-occurrence matrix by considering different subsets of a_8, i.e., considering $b\in a_8'$, where $a_8'\subseteq a_8$.

The co-occurrence matrices may again be either non-symmetric or symmetric. One of the non-symmetrical forms can be defined considering

$$t_{lk}=\sum_{i=1}^{P}\sum_{j=1}^{Q}\delta, \tag{1}$$

with

$$\begin{aligned}\delta=1 \quad & \text{if } f(i,j)=l \text{ and } f(i,j+1)=k\\ & \text{or } f(i,j)=l \text{ and } f(i+1,j)=k;\\ \delta=0 \quad & \text{otherwise.}\end{aligned}$$

Here only the horizontally-right and vertically-lower transitions are considered. The following definition of t_{lk} gives a symmetric co-occurrence matrix:

$$t_{lk}=\sum_{i=1}^{P}\sum_{j=1}^{Q}\delta$$

where

$$\begin{aligned}\delta=1 \quad & \text{if } f(i,j)=l \text{ and } f(i,j+1)=k\\ & \text{or } f(i,j)=l \text{ and } f(i,j-1)=k\\ & \text{or } f(i,j)=l \text{ and } f(i+1,j)=k \qquad (2)\\ & \text{or } f(i,j)=l \text{ and } f(i-1,j)=k;\\ \delta=0 \quad & \text{otherwise.}\end{aligned}$$

We consider here both horizontally right and left, and vertically upper and lower transitions.

2.2. Measures for thresholding

Since the co-occurrence matrix contains information regarding the spatial distribution of gray levels in the image, several workers have used them for segmentation. For thresholding at grey level 's', Weszka and Rosenfeld [3] defined the busyness measure as follows:

$$\text{Busy}(s)=\sum_{i=0}^{s}\sum_{j=s+1}^{L-1}t_{ij}+\sum_{i=s+1}^{L-1}\sum_{j=0}^{s}t_{ij}. \tag{3}$$

The co-occurrence matrix used in (3) is symmetric.

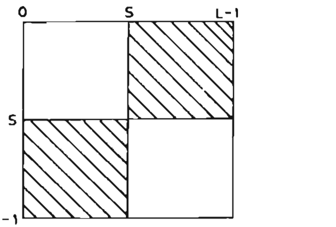

Figure 1. Pictorial representation of the busyness measure.

The sum of the entries of the shaded portion in Figure 1, represents the busyness measure for the level 's'. For an image with only two regions, say, object and background, the value of 's' for which the minimum of Busy(s) occurs, gives the threshold. Similarly, for an image having more than two regions the busyness measure provides a set of minima corresponding to different thresholds.

Deravi and Pal [4] have given a measure for the conditional probability of transition from one region to another as follows.

If the threshold is at 's', the conditional probability of transition from the region $[0,s]$ to $[s+1, L-1]$ is

$$P_1 = \frac{\sum_{i=0}^{s} \sum_{j=s+1}^{L-1} t_{ij}}{\sum_{i=1}^{s} \sum_{j=0}^{s} t_{ij} + \sum_{i=0}^{s} \sum_{j=s+1}^{L-1} t_{ij}} \tag{4}$$

and the conditional probability of transition from the region $[(s+1),(L-1)]$ to $[0,s]$ is

$$P_2 = \frac{\sum_{i=s+1}^{L-1} \sum_{j=0}^{s} t_{ij}}{\sum_{i=s+1}^{L-1} \sum_{j=s+1}^{L-1} t_{ij} + \sum_{i=s+1}^{L-1} \sum_{j=0}^{s} t_{ij}} . \tag{5}$$

t_{ij} in equations (4) and (5) corresponds to the one in equation (1).

$p_c(s)$, the condition probability of transition across the boundary is defined as

$$p_c(s) = (P_1 + P_2)/2. \tag{6}$$

The lower the value of $p_c(s)$, the lower is the probability that the next transition will be to a different class. That means a minimum of $p_c(s)$ will correspond to a threshold such that most of the transitions are within the class and few are across the boundary. Therefore, a set of minima of $p_c(s)$ would be obtained corresponding to different thresholds in F.

Chanda et al. [5] have als used the co-occurrence matrix for thresholding. They defined an average contrast measure as

$$\mathrm{AVC}(s) = \frac{\sum_{i=0}^{s} \sum_{j=s+1}^{L-1} t_{ij} * (i-j)^2}{\sum_{i=0}^{s} \sum_{j=s+1}^{L-1} t_{ij}} + \frac{\sum_{i=s+1}^{L-1} \sum_{j=0}^{s} t_{ij} * (i-j)^2}{\sum_{i=s+1}^{L-1} \sum_{j=0}^{s} t_{ij}} . \tag{7}$$

AVS(s) shows a set of maxima corresponding to the thresholds among various regions in F.

In the computation of t_{ij} they considered only vertical transitions in the downward direction.

It is to be mentioned here that all the three measures described before are basically based on some weighted combinations of the number of entries in the shaded and blank regions of Figure 1.

3. Segmentation based on contrast and homogeneity measure

In this section, we present an algorithm, for segmentation on the basis of the information of contrast and homogeneity (between/within the regions in F) as obtained from the co-occurrence matrix (eqn. (1)). The concept of human visual facts has been incorporated in the contrast measure in order to make the method of segmentation more effective.

Before describing the algorithm let us first of all explain some facts of the human visual system.

3.1. Human psychovisual facts

In psychology, a contrast C refers to the ratio of difference in luminance of an object B_0 and its immediate surrounding B [7] i.e.,

$$C = |B_0 - B|/B = \frac{\Delta B}{B} ,$$

The perceived grayness of a surface depends on its local background and the perceived contrast remains constant if the ratios of *contrasts* between objects and local background remain constant.

The visual increment threshold (or just noticeable difference) is defined as the amount of light ΔB_T necessary to add to a visual field of intensity B such that it can be discriminated from a reference field of same intensity B. It therefore gives a limit for a perceivable change in luminance or intensity.

At low intensity near the absolute visual threshold (mere presence or absence of light intensity detectable under a dark adapted condition), the visual increment threshold ΔB_T is constant. With an increase in B, ΔB_T converges asymptotically to the *Weber behaviour* i.e., $\Delta B_T \propto B$ ($\Delta B_T/B$ being defined as *Weber ratio*).

Figure 2 [6] presents such a characteristic re-

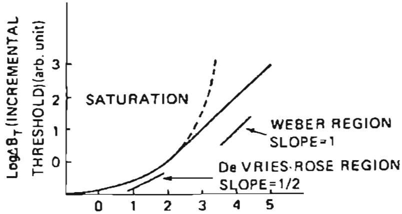

Figure 2. Variation of log ΔB_T with log B (in arbitrary scale) [6].

sponse in the $\log \Delta B_T - \log B$ plane. The Weber behaviour is characterised by the unit slope of the curve. The preceding region with slope ½ is known as *De Vries–Rose region*, characterised by $\Delta B_T \propto \sqrt{B}$. However, in the actual case this rule is followed in a small restricted region. The dashed curve shows the deviation from Weber's law. This deviation is usually not exhibited by the retinal core mechanism even under very high intensities, but can happen in very restricted cases [6].

Therefore, it the brightness value of an object is higher (lower) than its surrounding or background or a reference intensity B by such an amount that it corresponds to a point on or above the curve (Figure 2), the object will only then appear brighter (darker) i.e., discriminable for the human visual system (HVS). Furthermore, an equal amount of ΔB value created at different background intensity (B-values) does not result in an equal perceivable change to the HVS. For example, the discrimination ability in the De Vries–Rose region is greater than in the Weber region and this ability decreases with increase in value of B. The possible reason for this deterioration in discrimination ability can be attributed to inherent visual nonlinearity.

3.2. *Measures of contrast and homogeneity*

It has already been discussed that the problem of segmentation is to partition the set G_L of gray levels into some non-intersecting subsets such that each segment is as homogeneous as possible while the contrast between any segment and its neighbouring segments is as high as possible. Two such measures, namely the contrast of a segment with its neighbouring segments and the homogeneity of a segment are first of all defined. A composite measure of the two is then used to select the threshold levels in F.

Let the gray levels ranging from K to M form one of the segments, say, R_1 of the image F, i.e., $R_1 = [K, M]$, $M \geq K$.

Define $C^B_{K,M}$, the contrast of the segment R_1 with respect to other segments, as follows:

$$C^B_{K,M} = \frac{\sum_{i \in R_1} \sum_{j \notin R_1} t_{ij} \cdot W_{ij}}{(C_1) \cdot \sum_{i \in R_1} \sum_{j \notin R_1} t_{ij}} = \frac{\sum_{i=k}^{M} \sum_{j<k \text{ or } j>M} t_{ij} \cdot W_{ij}}{(C_1) \cdot \sum_{i=k}^{M} \sum_{j<k \text{ or } j>M} t_{ij}}. \tag{9}$$

Since the same amount of change in gray level at different positions on the gray scale does not produce identical impression to the Human Visual System (Section 3.1), the weighting factor W_{ij} is incorporated in the above expression. In other words, W_{ij} provides different phase weights to the changes 'i' to $(i+1)$ and 'i' to $(i+1)$, $i \neq i$, in order to reflect the different impressions created by these changes.

W_{ij} may be defined as

either $W_{ij} = |i-j|/(i+j)$ (10a)

or $W_{ij} = |i-j|/\max\{i,j\}$ (10b)

or $W_{ij} = |i-j|/\min\{i,j\}$. (10c)

It is, therefore, seen that W_{ij} also ensures the equal contribution to equation (9) when the object and background intensities are interchanged.

In the denominator of equation (9), the term $\sum\sum t_{ij}$ is used to make the measure independent of the size of regions while the constant C_1 is introduced to make $0 \leq C^B_{K,M} \leq 1$. C_1 obviously depends on the choice of W_{ij} and is equal to the maximum possible value of W_{ij}.

Therefore

$C_1 = (L-1)/(L+1)$ for equation (10a),

$= (L-1)/L$ for equation (10b),

$= (L-1)$ for equation (10c).

where L is the maximum level in F.

It therefore appears from equation (9) that if $i = j$, then $C^B_{K,M} = 0$ i.e., the contrast in $[K, M]$ is minimum. On the other hand, if $i = 1$ and $j = L$

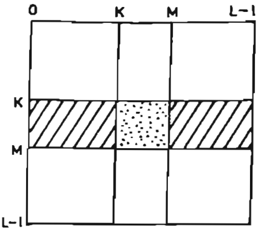

Figure 3. Pictorial representation of contrast and homogeneity measures.

then the contrast is maximum (= 1); since for all t_{ij} those are considered in $C^B_{K,M}$, the values of W_{ij} become equal to C_1 thereby making the numerator and denominator the same.

The t_{ij}'s considered in the equation (9) are shown by the shaded portion in Figure 3, which gives the total number of transitions across the boundary of the segment R_1, i.e., from the region $[K,M]$ to its outside.

Again, define $C^W_{K,M}$, the homogeneity of the region $[K,M]$, as

$$C^W_{K,M} = 1 - \frac{\sum_{i\in R_1}\sum_{j\in R_1} t_{ij} \cdot |i-j|}{(L-1) \cdot \sum_{i\in R_1}\sum_{j\in R_1} t_{ij}}$$

$$= 1 - \frac{\sum_{i=k}^{M}\sum_{j=k}^{M} t_{ij} \cdot |i-j|}{(L-1) \cdot \sum_{i=k}^{M}\sum_{j=k}^{M} t_{ij}}. \tag{11}$$

The t_{ij}'s considered in equation (11) are shown by the dotted portion in Figure 3. $C^W_{K,M}$ lies in the interval [0, 1] i.e., $0 \le C^W_{K,M} \le 1$.

If R_1 is perfectly homogeneous, $K = M$ (i.e. the region contains only one gray level), then $C^W_{K,M} = 1$ as $|i-j| = 0$ for all i and j.

With the decrease of the homogeneity of R_1, $|i-j|$ increases and approaches to $(L-1)$. As a result, the ratio in eqn. (11) approaches to one and $C^W_{K,M}$ tends to zero. Thus $C^B_{K,M}$ and $C^W_{K,M}$ are found to increase with increase in contrast and homogeneity respectively of a region $[K,M]$. On the basis of these two measures we define a composite measure

$$g_{K,M}(C^W_{K,M}, C^B_{K,M}) = C^W_{K,M} \cdot C^B_{K,M}. \tag{12}$$

Since $0 \le C^W_{K,M} \le 1$ and $0 \le C^B_{K,M} \le 1$, g also lies between 0 and 1.

The level at which g attains a maximum value can therefore be considered as a boundary (or threshold) between regions.

3.3. Merging of a single valued region

As mentioned above, the composite measure $g_{K,M}$ is found to be very much sensitive to highly uniform regions. In other words, a region containing only one gray level is very likely to be detected as a separate segment, although one does not desire to have such a segment. To avoid this we have suggested here a merging algorithm whereby such a single valued segment is accepted if the transitions within the segment is higher than those across it. This criterion enables one to retain only those regions which have significant dimension (i.e., informative) in spatial domain.

The algorithm therefore, first of all, finds out the regions of single gray level and determines whether such a region should be merged or not. If it decides to merge a region, then the next task is to determine the adjacent region (left or right) to which it is to be merged.

Let $R_i = [M,M]$ be the region under consideration.

Let T_W = total number of transitions within the region R_i, then

$$T_W = t_{M,M}. \tag{13}$$

If T_0 = total number of transitions from R_i to all other outside regions, then

$$T_0 = \sum_{l=M}^{M} \sum_{j\ne M} t_{lj}. \tag{14}$$

Decision rule

If $T_W > T_0$, then the size of the region can be taken as reasonably big. So accept the region, do not merge it, otherwise merge the region. Now the problem is to select the adjacent region to which it is to be merged.

Case 1. If $M=0$, merge R_l to the right adjacent region.

Case 2. If $M=L-1$, merge R_l to the left adjacent region.

Case 3. If $0<M<L-1$, we proceed as follows.

Let T_L = total number of transitions from R_l to its left adjacent region. If the left adjacent region contains gray levels ranging from L_1 to L_2 then

$$T_L = \sum_{i=M}^{M} \sum_{j=L_1}^{j=L_2} t_{ij}. \qquad (15)$$

Suppose T_R = total number of transitions from R_i to the right adjacent region and the right adjacent region contains gray levels ranging from R_1 and R_2 then

$$T_R = \sum_{i=M}^{M} \sum_{j=R_1}^{R_2} t_{ij} \qquad (16)$$

If $T_L > T_R$ then merge R_i to the left adjacent region, otherwise merge it to the right adjacent region.

The above decision rule can also be formulated in a more general way as follows.

Let T_t = total number of transitions from R_i to all regions including itself, i.e.,

$$T_t = \sum_{i=M}^{M} \sum_{j=0}^{L-1} t_{ij} \qquad (17)$$

and

$$\theta = T_W / T_t, \quad \theta < 1.$$

Now if $\theta > \theta_H$, accept the region, otherwise merge it, where θ_H is some pre-assigned threshold value.

$\theta_H = 0.5$ obviously gives the original decision rule.

The advantage of defining the decision rule in this manner is that in an interactive environment one can change θ_H if necessary and compare the results to pick up the most appropriate value of θ_H for a particular type of image.

In order to extract the thresholds in an image F, we start with $R_1 = [0,0]$ and increase the size of R_1 one by one to the right side of the gray scale until we get a local maximum of $g_{K,M}$, i.e., the process is started with $K = 0$, $M = 0$ and M is incremented one by one until $g_{K,M}$ attains a maximum value. If the maximum occurs at gray level K_1 then K_1 corresponds to a threshold and gray levels ranging from 0 to K_1 represent a region of the image F. Then we start with $K = M = K_1 + 1$ and the process is repeated as described above until we get the next maximum, say at K_2. The gray levels ranging from $K_1 + 1$ to K_2 thus constitute another segment. In this way the process is carried on until the entire gray scale is exhausted.

After obtaining a set of thresholds, the next task is to check if there is any single valued region to be merged.

4. Implementation and results

The segmentation algorithm described in Section 3 is implemented on a set of four different images [8] having dimension 64×64 with 32 gray levels. Figures 4(a), 5(a), 6(a) and 7(a) represent the original input images, while Figures 4(b), 5(b), 6(b) and 7(b) represent the corresponding gray level histograms. These images are produced on a line printer by over printing different character combination for different gray levels.

Figure 4(a) represents an image of Mona Lisa. It is to be noted that the gray level histogram Figure 4(b) is almost unimodal (having some local minima). When the present algorithm (without merging) is applied to it, four thresholds, namely 0, 1, 6 and 17 are produced. The corresponding segmented image is shown in Figure 4(g), where different segments are represented by different textures. When the merging algorithm is applied to it, the segment [1, 1] (Table 1) is merged to its right adjacent segment. The segmented image so obtained after merging is shown in Figure 4(c). Comparing Figures 4(c) and 4(g) we find that there is an undesirable region inside the hair of Mona Lisa, which after being merged results in a more meaningful segmentation (Figure 4(c)). Figure 4(g) is shown, as an illustration, only to demonstrate the effect of the merging algorithm in selecting final thresholds.

Figure 5(a) is an image of Abraham Lincoln, and the corresponding gray level histogram (Figure 5(b)) is found to have a number of deep valleys. The thresholds (before and after merging) generated by the proposed method are shown in Table 1. The output segmented image is shown in Figure 5(c).

In order to demonstrate the validity of the algorithm for images with flat and wide valleys in their histogram, the algorithm is applied to the image of a jet (Figure 6(a)). One can see in Figure 6(a) that

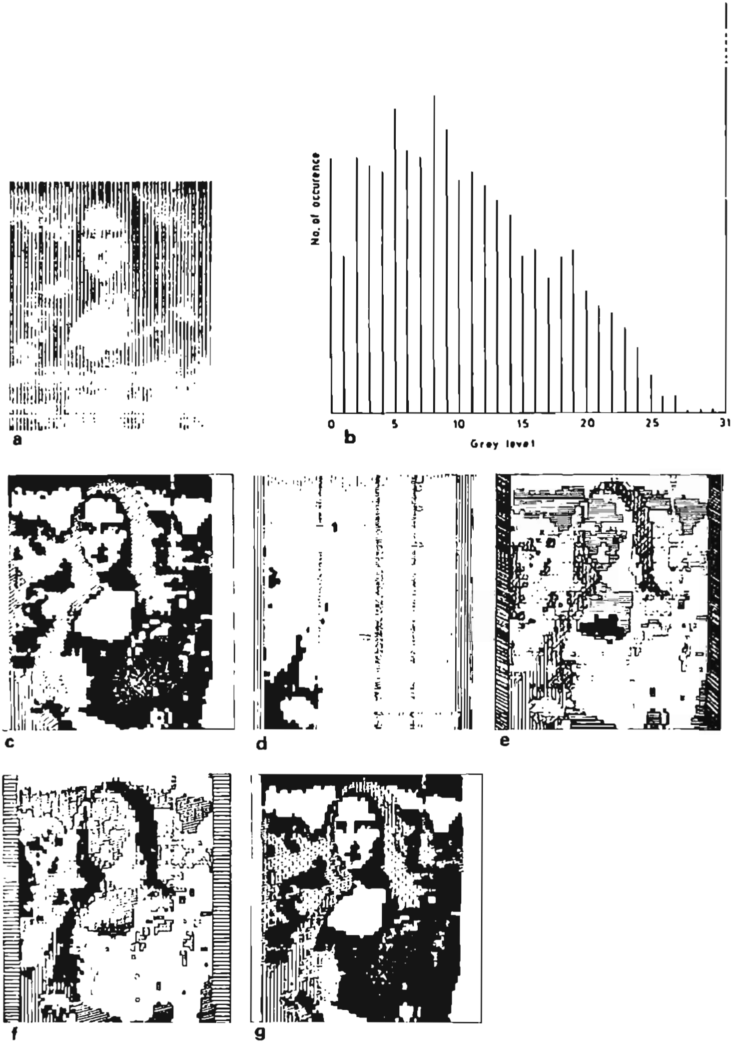

Figure 4. (a) Input image of Mona Lisa; (b) Histogram; (c) Segmented image by the proposed method; (d) Segmented image by [3]; (e) Segmented image by [4]; (f) Segmented image by [5]; (g) Segmented image by the proposed method (before merging).

the right wing of the jet has vanished inside the cloud in such a way that apparently it is very difficult to trace the boundary of the right wing. The present algorithm is found to be successful in separating out that wing from the cloud. Figure 6(c) represents the segmented image, while the thresholds are shown in Table 1.

Figure 7(a) represents the image of a biplane having two dominant modes in its histogram (Figure 7(b)). From Figure 7(c) the object is found

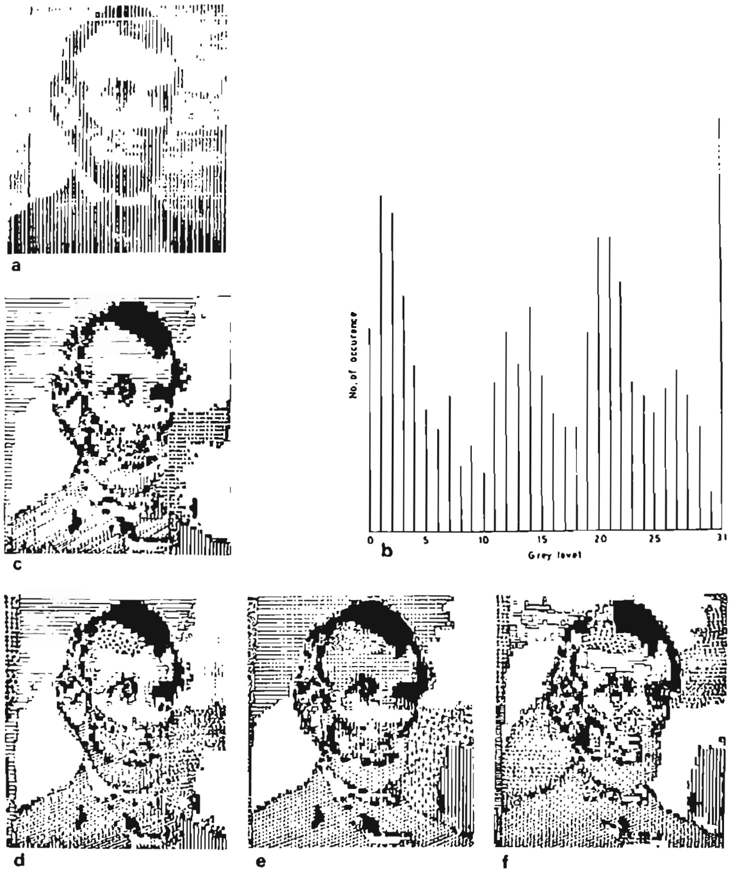

Figure 5. (a) Input image of Lincoln; (b) Histogram; (c) Segmented image by the proposed method; (d) Segmented image by [3]; (e) Segmented image by [4]; (f) Segmented image by [5].

to be clearly separated out from the background.

The above results were obtained by using eqn. (10a) while computing W_{ij} of the contrast measure. Experiments were also carried out with eqn. (10b) and the corresponding performances were found to be almost similar.

Comparison

In order to compare the performance of the algorithm with those of some of the existing algorithms based on co-occurrence matrix, we have considered algorithms of Weszka and Rosenfeld (eqn. (3)) [3], Deravi and Pal (eqn. (4)-(6) [4] and Chanda et al. (eqn. (7)) [5]. The thresholds obtained by these methods are also shown in Table 1.

Equation (3) is found to fail to extract all the meaningful regions of the image of Mona Lisa. It has selected only three segments (Figure 4(d)) with thresholds at 0 and 30, as a result, most of the im-

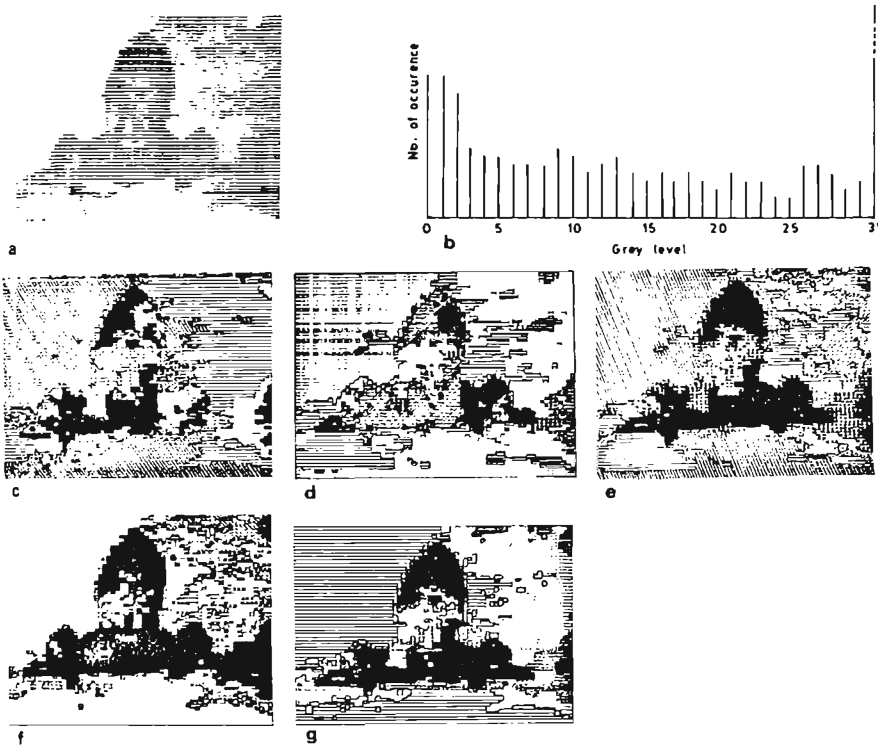

Figure 6. (a) Input image of a jet; (b) Histogram; (c) Segmented image by the proposed method; (d) Segmented image by [3]; (e) Segmented image by [4]; (f) Segmented image by [5]; (g) Segmented image with the proposed method with $\theta = 0.75$.

portant informations are lost.

Equations (4)–(6), on the other hand, have detected two extra segments in the chest and one extra region in the hair (Figure 4(e)), while eqn. (7) has also produced two similar segments (one of them is of smaller size than the corresponding one

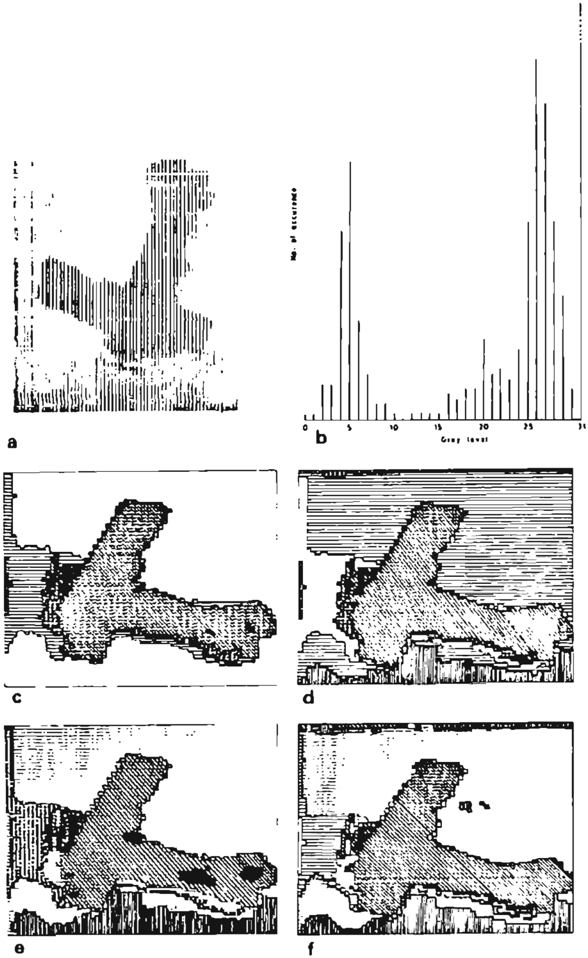

Figure 7. (a) Input image of a biplane; (b) Histogram; (c) Segmented image by the proposed method; (d) Segmented image by [3]; (e) Segmented image by [4]; (f) Segmented image by [5].

Table 1
Thresholds for various methods

Images	Proposed method		Method of Weszka and Rosenfeld [3]	Method of Deravi and Pal [4]	Method of Chanda et al. [5]
	Before merging	After merging			
Mona Lisa (Figure 4)	0, 1, 6, 17	0, 6, 17	0, 30	0, 3, 6, 17, 25, 30	0, 3, 6, 17, 28, 30
Lincoln (Figure 5)	0, 4, 9, 13, 17, 18, 24	0, 4, 9, 13, 17, 24	0, 4, 9, 12, 17, 24, 30	4, 9, 12, 17, 23, 30	5, 10, 12, 17, 22, 27, 30
Jet (Figure 6)	0, 3, 7, 23, 24, 26	0, 3, 7, 24, 26	0, 3, 5, 7, 18, 29	3, 7, 11, 18, 29	7, 11, 13, 18, 21, 23
Biplane (Figure 7)	0, 1, 2, 10, 17, 18, 23	1, 10, 17, 23	0, 2, 10, 17, 22, 30	0, 2, 10, 17, 22, 30	0, 10, 14, 17, 21, 29

produced by eqn. (4)–(6) (Figure 4(f)). From Figure 4(e), it is seen that the regions generated by the proposed method where these additional regions are absent, create a better impression to the eye.

For the image of Lincoln (Figure 5) all the methods except the present one, has divided the fore-head into two regions. This is probably due to the fact that those methods are based on similar concept. Furthermore, the present method and the method by Weszka and Rosenfeld have divided the beard into two regions, which is not the case for the other two methods.

In the case of the jet, eqn. (7) has failed to discriminate between the cloud and the right wing (Figure 6(f)) while the other two methods like ours are successful in doing so (Figure 6(c), 6(d) and 6(e)). The present method and eqn. (3) produced comparable results, while the result produced by eqns. (4)–(6) (Figure 6(e)) seem to create a better impression to the eye. However, if we alter the value of θ_H from 0.5 to 0.75 in the merging algorithm, the first two regions (Table 1) get merged and the resulting image (Figure 6(g)) is found to be much better than Figure 6(c) in that respect (having less noisy background).

In case of the biplane, all the methods have been able to detect its contour (Figure 7). However, eqns. (4)–(6) are found to generate two additional regions inside the tail of the biplane which are absent for other cases. The background is found to be clustered in two parts by all but our method. Furthermore, all the methods except ours have divided the shade of the biplane into three or more regions, which is two in our case.

From Table 1 it appears that the equations (3), (4)–(6) and (7) detected, except for the jet, a threshold at the end of the grey scale (e.g., 30 for Mona Lisa, Lincoln and biplane). These thresholds correspond to some undesirable regions at the frame of the images. The incorporation of the factor W_{ij} in eqn. (9) which accounts for the nonlinear behaviour of the HVS has been found to be able to eliminate such occurrences.

5. Conclusion

A two-stage algorithm for image segmentation is described using the measures of 'homogeneity' and 'contrast' (involving the concept of logarithmic response of the human visual system) within/between regions of the image and incorporating thereafter a provision for merging to eliminate undesirable segments.

The algorithm is found to be much more effective than the existing ones in extracting meaningful regions of images having unimodal, multiple valleyed, flat-wide valleyed and bimodal (with local minima) histograms.

Acknowledgement

The authors gratefully acknowledge the help of Mr. N. Chatterjee in typing the manuscript and Professor D. Dutta Majumder for his interest in this work.

References

[1] Pavlidis, T. (1977). *Structural Pattern Recognition*. Springer, New York.

[2] Rosenfeld, A. and A.C. Kak (1982). *Digital Picture Processing*. Academic Press, New York.

[3] Weszka J.S. and A. Rosenfeld (1978). Threshold evaluation techniques. *IEEE Trans. SMC* 8, 622–629.

[4] Deravi, F. and S.K. Pal (1983) Graylevel thresholding using second-order statistics. *Pattern Recognition Letters* 1(5), 417–422.

[5] Chanda, B., B.B. Chaudhuri and D. Dutta Majumder (1985). On image enhancement and threshold selection using the graylevel co-occurrence matrix. *Pattern Recognition Letters* 3(4), 243–251.

[6] Buchsbaum G. (1980). An analytical derivation of visual nonlinearity. *IEEE Trans. Biomedical Engineering* 27, 237–242.

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

[8] Gonzalez R.C. and P. Wintz (1977). *Digital Image Processing*. Addison-Wesley, London, MA.