Pattern Recognition Letters 4 (1986) 433–441
North-Holland

December 1986

Thresholding for edge detection using human psychovisual phenomena

M.K. KUNDU and S.K. PAL

Electronics and Communication Sciences Unit, Indian Statistical Institute, Calcutta 700035, India

Received 15 April 1986

Abstract: An algorithm based on the facts of the human visual system is presented here whereby it is possible to select automatically (without human intervention) the thresholds for detecting the significant edges as perceived by human beings. The threshold value adapts with the background intensity according to the criterion governed by a characteristic of one of the De Vries-Rose, Weber's and saturated regions. The algorithm is found to provide a satisfactory improvement in the performance in the conventional edge detection process.

Key words: Human visual system, edge detection, adaptive thresholding.

1. Introduction

Visual information is concentrated at points of large spatial variation of light intensity in a picture. The difference operators such as gradient, Robert gradient, Sobel and Prewitt gradients and Laplacian (Gonzalez and Wintz, 1977) give rise to high values at the places where the change in greylevel occurs. The greater the spatial variation of intensity, the easier is the detection of change and the stronger is the edge intensity. Thresholding of spatial difference of greylevels (gradient image) is one of the popular techniques for sharpening the edges (i.e. selecting valid or significant edge points according to human perception). The problem of edge detection therefore boils down to finding an appropriate threshold which may be global or local or dynamic such that if the spatial difference at a point exceeds that threshold, then and only then it will be considered as a valid edge point.

The present work is an attempt to make this task automatic by providing an adaptive algorithm based on human psychovisual phenomena. The visual increment threshold curve (Buchsbaum, 1980) is piecewise linearly approximated here such that the spatial difference in greylevel at a point is thresholded depending on its background intensity by one of the criteria in De Vries-Rose, Weber and saturated regions. Effectiveness of the algorithm is demonstrated for the existing gradient operators when a set of different images is considered as input.

2. Edge operators and thresholding

Let x_{mn} be the greylevel of the (m,n)th pixel of an $M \times N$ dimensional, L level image, $X=[x_{mn}]$, $m=1,2,\ldots,M$; $n=1,2,\ldots,N$. The gradient x'_{mn} denoting edge intensity at the point (m,n) may be defined as (Hall, 1979)

$$x'_{mn} = (|G_1| + |G_2|)/2 \tag{1}$$

where G_1 and G_2 are given by equations (2), (3) and (4) corresponding to ordinary gradient, Robert's gradient, and Prewitt's and Sobel's gradients:

$$G_1 = x_{mn} - x_{m,n+1}, \tag{2a}$$

$$G_2 = x_{mn} - x_{m+1,n}, \tag{2b}$$

0167-8655/86/$3.50

$$G_1 = x_{mn} - x_{m+1,n+1}, \tag{3a}$$

$$G_2 = x_{m,n+1} - x_{m+1,n}, \tag{3b}$$

$$G_1 = \frac{1}{2+W}\left[(x_{m+1,n+1} + Wx_{m+1,n} + x_{m+1,n-1}) - (x_{m-1,n+1} + Wx_{m-1,n} + x_{m-1,n-1})\right], \tag{4a}$$

$$G_2 = \frac{1}{2+W}\left[(x_{m-1,n+1} + Wx_{m,n+1} + x_{m+1,n+1}) - (x_{m+1,n-1} + Wx_{m,n-1} + x_{m-1,n-1})\right], \tag{4b}$$

with $W=1$ and 2 for Prewitt's and Sobel's gradients respectively. These are called the first difference operators. The second difference operator such as the Laplacian can be expressed as

$$G_1 = \tfrac{1}{2}(x_{m-1,n} + x_{m+1,n} - 2x_{mn}), \tag{5a}$$

$$G_2 = \tfrac{1}{2}(x_{m,n-1} + x_{m,n+1} - 2x_{mn}). \tag{5b}$$

The simple gradient and Robert's gradient (equations (2) and (3)) not only respond strongly to the edges but also to the isolated points. The Laplacian does respond to the edges, but it responds more strongly to the corners, lines, line ends and isolated points. It has zero response to a linear ramp but it responds strongly where there is a change in rate of change of graylevel. Due to some smoothing effect, Prewitt and Sobel operators, on the other hand, possess better noise immunity and good edge response, and added to this the weighted average of the Sobel operator (equation (4) with $W=2$) reduces to a great extent the blurring effect of smoothing.

After applying these operators on an image X, an edge is deemed to be present at the point (m,n) if x'_{mn} exceeds a predefined threshold value T. The general form of T is

$$T = f(x_{mn}, N_{mn}, (m,n)) \tag{6}$$

where N_{mn} denotes some local property at the point (m,n) having pixel intensity x_{mn}.

The threshold is called 'global' if T depends only on x_{mn}. When T depends on both x_{mn} and N_{mn}, it is called 'local'. If T depends on the co-ordinate (m,n) in addition to x_{mn} and N_{mn}, the threshold so computed is defined as 'dynamic'. A good review on different thresholding techniques is given in (Wezka, 1978). Selection of an appropriate threshold therefore poses an important problem in detecting the valid edge points in X.

In the following sections, we are goint to present an automatic threshold selection procedure based on human psychovisual facts. The selection is made on the basis of both local and global information available from X such that for a particular x'_{mn} its threshold value adapts with its average background intensity B_{mn} following either $K_2\sqrt{B_{mn}}$ in the De Vries-Rose region or $K_1 B_{mn}$ in the Weber region or $K_3 B_{mn}^2$ in the saturation region (K_1, K_2 and K_3 being positive constants).

3. Human psychovisual phenomena

Brightness is a psychological sensation associated with the amount of light stimulus. Due to great adaptive ability of the eye, absolute brightness could not be measured by the human eye. The relative brightness is an observer's feeling of difference in greyness between the objects. The term contrast is used to emphasize the difference in illuminance of objects. The perceived greyness of a surface depends on its local background and the perceived contrast remains constant if the ratios of contrasts between object and local background remain constant (Hall, 1979).

In psychology, contrast C refers to the ratio of difference in luminance of an object B_o and its immediate surrounding B_S, i.e.

$$C = |B_o - B_S|/B_S = \frac{\Delta B}{B_S}. \tag{7}$$

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 the same intensity B. It therefore gives a limit for a perceiveable change in luminance or intensity.

The major problems in a low intensity level which any image processing system has to face are (Zuidema et al., 1983):

(i) Detection of changes occuring in a low steady but visible illumination (i.e. minimum detectable change), and

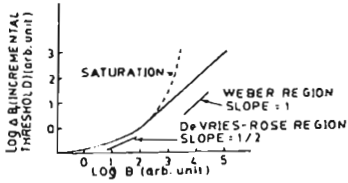


Figure 1. The increment threshold ΔB_T as a function of reference intensity B.

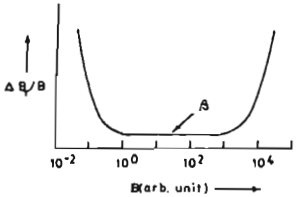


Figure 2. The variation of contrast sensitivity $\Delta B_T/B$ with background intensity B for uniform background.

(ii) Detection of the mere presence or absence of light under a dark adapted condition (i.e. absolute visual threshold).

At low intensity near the absolute visual threshold, the visual increment threshold ΔB_T is constant. With an increase in intensity B, ΔB_T converges assymptotically to the Weber behaviour i.e., $\Delta B_T \propto B$. This type of behaviour is exihibited in a brightness incremental threshold for white broad band spectra and monochromatic narrow band spectra (Buchsbaum, 1980).

Figure 1 presents such a characteristic response 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 1/2 is known as the *De Vries-Rose region*. It has been shown (Buchsbaum, 1980) that if the central visual processor behaves as an optimum probabilistic detector, the incremental visual threshold follows the square root law, i.e. $\Delta B_T \propto \sqrt{B}$. However, in the actual case this rule is followed in a small restricted region (Figure 1). The dashed curve shows the deviation from Weber's law. This deviation (which represents a *saturation region*) is not usually exhibited by retinal cone mechanism even under very high intensities but could happen in very restricted cases.

From Figure 1 it is seen that variation of $\log \Delta B$ against $\log B$ in the De Vries-Rose region is slower than that in the Weber region. In other words, the discrimination ability in the De Vries-Rose region is greater than in the Weber region. The possible reason for this detoriation in discrimination ability can be attributed to inherent visual nonlinearity.

Assuming the response curve (Figure 1) to be piecewise linear, the Weber region and the De Vries-Rose region can be represented as

$$\log \Delta B_T = \log (K_1 B) = \log B + \log K_1 \tag{8}$$

and

$$\log \Delta B_T = \log (K_2 \sqrt{B}) = \tfrac{1}{2} \log B + \log K_2, \tag{9}$$

respectively. K_1 and K_2 are the constants of proportionality.

Furthermore, the saturation which might occur even rarely, only at very high intensities can be approximated as

$$\log \Delta B_T = 2 \log B + K_3, \tag{10}$$

K_3 being a constant.

Therefore, when the brightness value of an object is higher (or lower) than its surrounding or background or a reference intensity B by an amount $\geq \Delta B_T$, it corresponds to a point on or above the curve (Figure 1) and the object will appear either brighter (or darker). In case of variable object and background intensity, the visual system adapts to an average intensity.

Furthermore, Figure 2 shows the variation of $\Delta B_T/B$ with B. It is seen that in the Weber region,

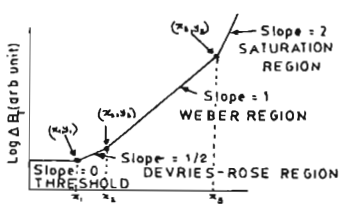


Figure 3. The increment threshold ΔB_T vs. reference intensity B curve (Figure 1) with linear approximation for different regions.

the ratio $\Delta B_T/B$ remains fairly constant at about β % of its maximum value over a very wide range of its B-value (Gonzalez and Wintz, 1977). In the literature the value of β is said to be around 0.02.

4. Adaptive thresholding

The piecewise linear approximation of the visual increment threshold curve (Figure 1) is shown in Figure 3. Here, the threshold values in the De Vries-Rose region, the Weber region and the saturation region are defined by the linear equations (8), (9) and (10) respectively. Although the demarkation between those regions represented by various equations is not very sharp and definite, we assume here for the sake of simplicity and ease of analysis that the De Vries-Rose region extends between x_1 and x_2 (Figure 3) (x_1 corresponds to absolute visual threshold), the Weber region between x_2 and x_3 and the saturation region beyond x_3. The value of B corresponding to $\log B = x_i$ is assumed here to be B_{x_i}, $i = 1, 2, 3$. In other words,

$$x_i = \log B_{x_i}, \quad i = 1, 2, 3. \tag{11}$$

Let x_0 and B_1 be the maximum value of $\log B$ and B respectively. Let us assume further that

$$x_i = \alpha_i x_0, \quad i = 1, 2, 3, \tag{12}$$

and

$$B_{x_i} = \alpha_i' B_1, \quad i = 1, 2, 3, \tag{13}$$

where

$$0 < \alpha_1 < \alpha_2 < \alpha_3 < 1, \qquad 0 < \alpha_1' < \alpha_2' < \alpha_3' < 1.$$

As mentioned in Section 1, the value of $\Delta B_T/B$ remains fairly constant over the entire Weber region with a value approximately equal to β % of $(\Delta B/B)_{max}$, the maximum value of $\Delta B/B$ over the entire dynamic range. Therefore, from equation (8) one can write

$$K_1 = \frac{\Delta B_1}{B} = \frac{1}{100}\,\beta \left(\frac{\Delta B}{B}\right)_{max}. \tag{14}$$

Now, equations (8) and (9) are both satisfied at the point (x_2, y_2) of Figure 3. Therefore, we can write

$$y_2 = x_2 + \log K_1 = \tfrac{1}{2} x_2 + \log K_2$$

or

$$\tfrac{1}{2} x_2 = \log (K_2/K_1). \tag{15}$$

Using equation (11) for $i = 2$ we obtain

$$\tfrac{1}{2} \log B_{x_2} = \log (K_2/K_1)$$

or

$$K_2 = K_1 \sqrt{B_{x_2}}. \tag{16}$$

Similarly, considering equations (8) and (10) at the point (x_3, y_3) we have

$$K_3 = K_1 / B_{x_3}. \tag{17}$$

The minimum values of the increment threshold are therefore

$$\Delta B_T = K_2 \sqrt{B} = \frac{\sqrt{B}}{100}\,\beta \left(\frac{\Delta B}{B}\right)_{max} \sqrt{B_{x_2}}, \quad B_{x_2} \geq B \geq B \tag{18a}$$

$$= K_1 B = \frac{B}{100}\,\beta \left(\frac{\Delta B}{B}\right)_{max}, \qquad B_{x_3} \geq B \geq B \tag{18b}$$

$$= K_3 B^2 = \frac{B^2}{100}\,\beta \left(\frac{\Delta B}{B}\right)_{max} \frac{1}{B_{x_3}}, \quad B \geq B_{x_3}, \tag{18c}$$

corresponding to the De Vries-Rose, Weber and saturation regions.

Equation (12) defines the minimum amount of brightness level, in three different regions, by which the intensity of an object must be greater or less than its background intensity B in order to make the object appear brighter or darker (i.e. detectable). In other words, for a particular point in a picture having intensity B_p if we have

$$\text{either} \quad \frac{\Delta B}{\sqrt{B}} \geq K_2 \quad \text{when } \alpha_2' B_1 \geq B \geq \alpha_1' B_1, \tag{19a}$$

$$\text{or} \quad \frac{\Delta B}{B} \geq K_1 \quad \text{when } \alpha_3' B_1 \geq B \geq \alpha_2' B_1, \tag{19b}$$

$$\text{or} \quad \frac{\Delta B}{B} \geq K_3 \quad \text{when } B \geq \alpha_3' B_1, \tag{19c}$$

with

$$\Delta B = |B_p - B|, \tag{20}$$

then and only then it can be considered as a detectable edge point having edge intensity ΔB.

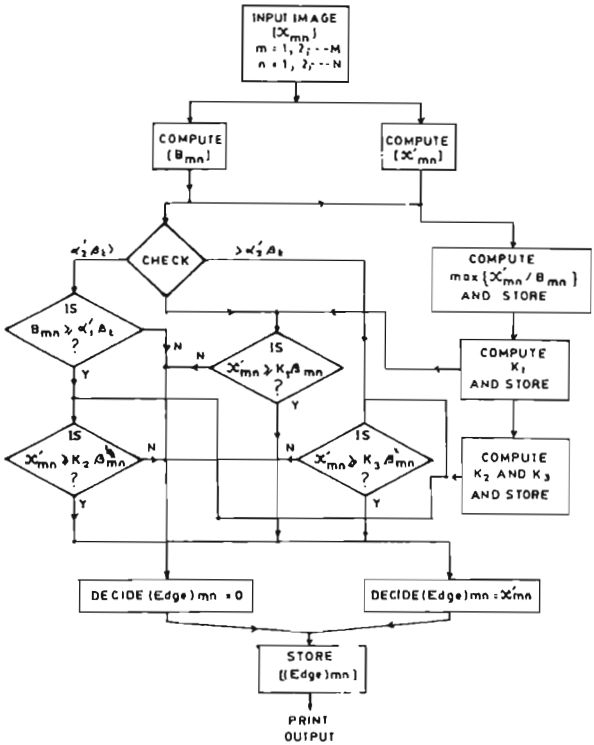


Figure 4. Flow chart for the proposed thresholding algorithm.

5. Algorithm

Figure 4 shows the flow chart for extracting valid edges of an $M \times N$ dimensional, L level image $X = [x_{mn}]$, $m = 1, 2, \ldots, M$; $n = 1, 2, \ldots, N$, using the above mentioned adaptive thresholding scheme.

B_{mn}, denoting the background intensity B at the point (m, n), was considered here to be the weighted average over its 8 neighbours, i.e.

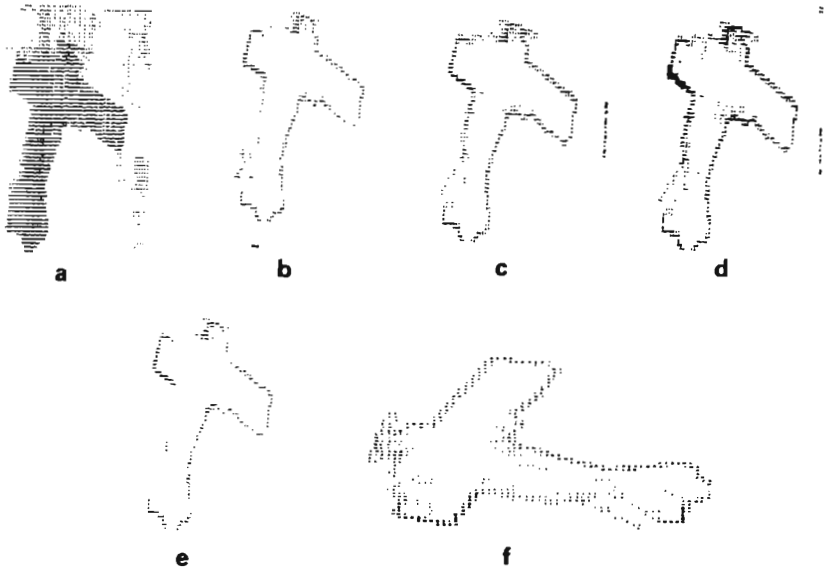
$$B_{mn} = \left\{\frac{1}{2}\left(\frac{1}{4}\sum_{Q} x_{ij} + \frac{1}{4\sqrt{2}}\sum_{Q'} x_{kl}\right) + x_{mn}\right\}\Big/2, \quad (i,j) \in Q,\ (k,l) \in Q', \tag{21}$$

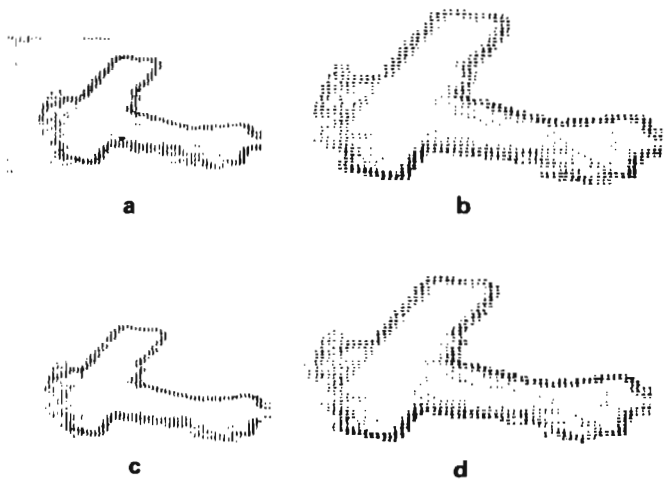
where Q and Q' denote the set the neighbouring pixels x_{ij} and x_{kl} which are at a distance of 1 and $\sqrt{2}$ units respectively from x_{mn}. Similarly, x'_{mn} denoting the ΔB-value at the (m, n)th point (difference in intensity of x_{mn} with respect to its neighbouring pixels) can be obtained by equations (1) to (5) corresponding to different spatial edge detection (gradient) operators.

Finally, for computing the constants K_i, $i = 1, 2, 3$, we consider

$$B_t = \max\{x_{mn}\} - \min\{x_{mn}\} \tag{22}$$

and


a
b
c
d
e
f


a
b
c
d

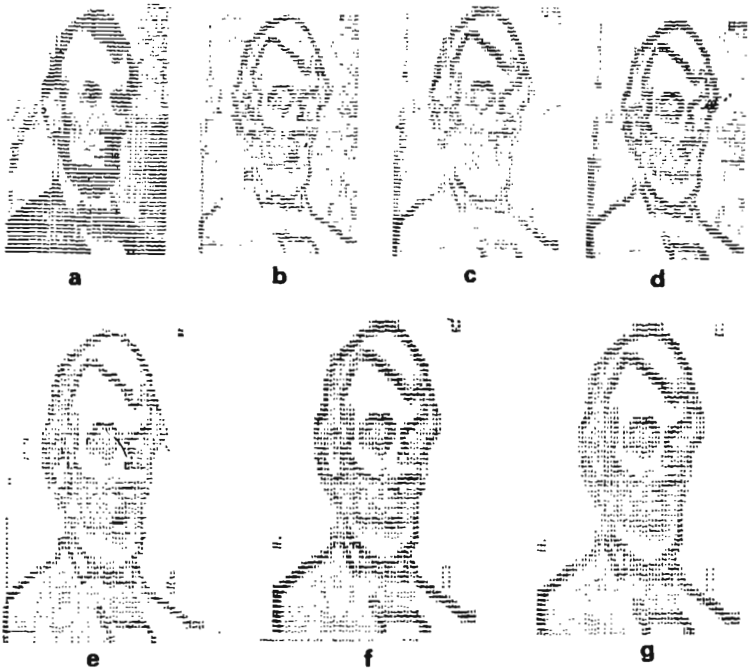


Figure 7. (a) Input image of Lincoln, (b) Edge detected image using Robert's gradient, (c) Edge detected image using the Prewitt operator, (d) Edge detected image using the Sobel operator, (e) When the image 7(b) is thresholded ($\alpha_2' = 0.1$, $\alpha_3' = 0.9$), (f) When the image 7(d) is thresholded ($\alpha_2' = 0.1$, $\alpha_3' = 0.9$), (g) When the image 7(c) is thresholded ($\alpha_2' = 0.1$, $\alpha_3' = 0.9$).

$$\left(\frac{\Delta B}{B}\right)_{\max} = \max\{x'_{mn}/B_{mn}\}, \quad (23)$$

$$m = 1, 2, \ldots, M, \quad n = 1, 2, \ldots, N.$$

Here, B_t has been taken to be equal to the dynamic range of the input image.

As shown in Figure 4, the task of thresholding is two-fold. At the first stage, each pixel intensity x_{mn} is decided, on the basis of its local information (namely, background intensity B_{mn}), to encounter the requirement in one of the three regions. The second stage consists of checking whether x'_{mn} satisfies its respective criterion (equation (19)). If that condition is satisfied, x_{mn} is then

Figure 5. (Previous page, above.) (a) Input iplane image, (b) Edge detected image using Robert's gradient, (c) Edge detected and thresholded image ($\alpha_2' = 0.33$, $\alpha_3' = 0.66$), (d) Edge detected and thresholded image ($\alpha_2' = 0.1$, $\alpha_3' = 0.7$), (e) Edge detected and thresholded image ($\alpha_2' = 0.3$, $\alpha_3' = 0.9$), (f) Edge detected and thresholded image ($\alpha_2' = 0.1$, $\alpha_3' = 0.9$).

Figure 6. (Previous page, below.) (a) Edge detected biplane image using the Prewitt operator. (b) When the image 6(a) is thresholded ($\alpha_2' = 0.1$, $\alpha_3' = 0.9$), (c) Edge detected biplane image using Sobel operator, (d) When the image 6(c) is thresholded ($\alpha_2' = 0.1$, $\alpha_3' = 0.9$).

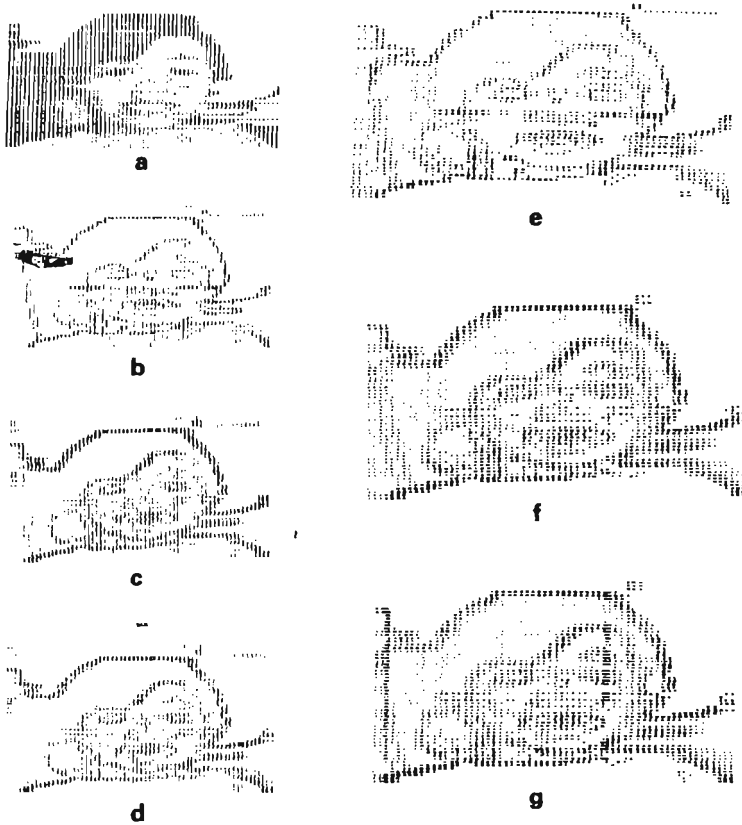


Figure 8. (a) Input image of a coded boy, (b) Edge detected image using Robert's gradient, (c) Edge detected image using the Prewitt operator, (d) Edge detected image using the Sobel operator, (e) When the image 8(b) is thresholded ($\alpha_2^1 = 0.1$, $\alpha_2^1 = 0.9$), (f) When the image 8(c) is thresholded ($\alpha_2^1 = 0.1$, $\alpha_2^1 = 0.9$), (g) When the image 8(d) is thresholded ($\alpha_2^1 = 0.1$, $\alpha_2^1 = 0.9$).

treated as a valid edge point having edge intensity x'_{mn}. Otherwise, the edge intensity at the (m,n)th point is taken to be zero. It is to be mentioned here that the thresholding criterion (equation (18) or (19)) for a particular x'_{mn} is seen to be adaptive (i.e. changing with its B_{mn} value according to $K_2\sqrt{B_{mn}}$ or $K_1 B_{mn}$ or $K_3 B_{mn}^2$ in the respective cases).

6. Implementation and result

The algorithm discussed in the previous section was simulated on an EC 1033 computer. The output images were generated by an overprinting technique as there is no facility for graphic display. A set of three images Figures 5(a), 7(a) and 8(a) were taken as test data of size 64×64 and with 32 gray-

levels. Although it was mentioned in an earlier section that Weber's region covers the major portion of the total dynamic range and both De Vries-Rose and saturated regions cover a very small portion of it (Buchsbaum, 1980), there is no quantitative figure available for the length of span of each region discussed. To estimate the approximate length of span for different regions, four different combinations of α_2' and α_3' were taken. They are ($\alpha_2'=0.33$, $\alpha_3'=0.66$), ($\alpha_2'=0.1$, $\alpha_3'=0.7$), ($\alpha_2'=0.3$, $\alpha_3'=0.9$) and ($\alpha_2'=0.1$, $\alpha_3'=0.9$). These sets of values were used in the present algorithm (Figure 4) and four different results of thresholded edges for the Biplane image (Figure 5a) are shown in Figures 5(c)-5(f).

It is evident that if the value of α_2' increases or decreases, the pixel population for the De Vries-Rose region (having lowest ΔB_T value) will increase or decrease. As a result, a more or less number of edge points will be able to cross the threshold for detection. Similarly, as α_3' increases or decreases, the pixel population for saturation region will decrease or increase. Therefore a more or less number of pixels will appear as significant edge points (due to comparatively high value of ΔB_T in saturated region). These facts are clearly demonstrated in Figures 5(c)-5(f). It is seen from Figure 5(f) that a thin and continuous edge appeared inside the Biplane together with normal outline edges of the Biplane. This is true only for a particular set of α' values ($\alpha_2'=0.1$, $\alpha_3'=0.9$). But for other sets of values for α' the inner edge is not continuous.

It was also mentioned in the literature (Gonzalez and Wintz, 1977) that the value of $\beta \approx 0.02$. If the value of β is increased or decreased, then the values of ΔB_T (equations 18(a)-(e)) for all the regions will simultaneously increase or decrease, which results in overall thinning or thickening of edges. In the present experiment different values of β such as 0.02, 0.03, 0.04, 0.05 and 0.06 were considered. As a typical illustration $\beta \approx 0.04$ were taken for all the results presented here for different images.

The results of the thresholding algorithm are shown in Figure 5(f) when the Robert gradient was used to calculate x'_{mn}. Figures 6(b) and 6(d) correspond to the results when x'_{mn} calculations were based on Prewitt and Sobel operators with $\alpha_2'=0.1$, $\alpha_3'=0.9$ and $\beta=0.04$. This part of the investigation demonstrates the effectiveness of the proposed algorithm for the other edge detection operators as well.

The same sets of experiments were performed on two other images, namely Lincoln (Figure 7(a)) and a coded boy (Figure 8(a)). The corresponding results are shown in Figures 7(e)-7(g) and Figures 8(e)-8(g).

7. Conclusion

The thresholding technique proposed in this paper does not require any human intervention for the choice of threshold values. Although certain portions of the edges are found to be more than one pixel width, those could be made thinned by using any standard edge linking technique. The probable application of the technique will be in the field of robot vision where the scope of human interaction is very limited.

Acknowledgment

The authors wish to thank Dr. D. Dutta Majumder, Head of Electronics and Communication Sciences Unit of Indian Statistical Institute for his constant encouragement for this work. The authors also wish to thank Mr. J. Gupta for typing the manuscript.

References

Buchsbaum, G. (1980). An analytical derivation of visual nonlinearity. *IEEE Trans. Biomed. Engrg.* 27, 237-242.

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

Wezka, J.S. (1978). A survey of threshold selection. *Computer Graphics and Image Processing* 7, 259-265.

Zuidema, P. et al. (1983). A mechanistic approach to threshold behaviour of visual system. *IEEE Trans. System, Man Cybernet.* 13, 923-934.

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