Pattern Recognition Letters 1 (1983) 417–422
North-Holland

July 1983

Grey level thresholding using second-order statistics

F. DERAVI and S.K. PAL*

Electrical Engineering Department, Imperial College of Science and Technology, London SW7 2BT, England

Received 28 February 1983
Revised 1 March 1983

Abstract: This letter describes algorithms for global thresholding of grey-tone images which use second-order grey level statistics. Two measures of interaction between classes of intensity levels are defined on simple co-occurance matrices and are used to evaluate and select thresholds. One of these measures is seen to be independent of the grey level histogram and effective in selecting thresholds for images with unimodal grey level distributions. The algorithms are also used for multithresholding without modifications.

Key words: Segmentation, threshold selection, co-occurance matrices, unimodal images.

1. Introduction

Grey level thresholding for the purpose of image segmentation is essentially a classification problem. The intensity (grey) levels are to be subdivided into bands so as to provide classes of intensity levels corresponding to regions of similar attribute.

Various threshold selection techniques have been derived based on the grey level histogram or 'improved' versions of such histograms (using edge strength information) (Weszka (1978), Pal et al. (to appear)). These techniques follow the simple heuristic of threshold selection at the minimum between histogram peaks (valleys). Weszka and Rosenfeld (1978) suggested a cost function based on the joint probability matrices of grey levels which can be used for threshold evaluation and selection. Unlike grey level histograms, such co-occurrence matrices (or grey tone spatial dependency matrices (Haralick et al. (1973))) contain information about the spatial relationship between the intensity levels and can therefore be the basis of more meaningful criteria for grey level classification. However, the 'business' measure of Weszka and Rosenfeld (1978) as a function of the threshold level is essentially an improved histogram and in this respect is similar to other methods aimed at using second-order statistics to define improved histograms.

This paper describes two 'interaction measures' for the selection of thresholds based on similar second-order grey level statistics as those mentioned above. The measures are defined on simple joint frequency matrices for grey levels occurring in horizontal and vertical nearest neighbour relative positions. The relationship between grey levels at these relative displacements is here referred to as 'intensity transition' and the corresponding co-occurance matrix is referred to as a transition matrix. The interaction measures are defined to represent the 'cost' of a threshold in terms of the probabilities of transition between the intensity classes which the threshold defines. Therefore the optimum threshold is chosen so as to minimise the interaction measures.

One of the measures is similar to the 'business'

* On leave from the Electronics and Communication Sciences Unit, Indian Statistical Institute, Calcutta 700 035, India.

measure of Weszka and Rosenfeld (1978) in having the same general shape as the image histogram. It is therefore unable to facilitate the selection of thresholds when different regions are not separated by 'valleys' in the histogram (e.g. unimodal histograms) or when the valleys are long and flat. The second measure, however, is seen to be independent of the shape of the histogram and can be used for threshold selection even when different regions are not separated by valleys. The measures are used for selecting multiple thresholds without further modifications.

The results of application of these measures on a number of images are compared and reported in this paper.

2. Definitions

Given an $M \times N$ dimensional, L-Level grey tone image

$$X = \{x_{mn} : m = 1, \dots, M;\ n = 1, \dots, N\}$$

with grey levels $x_{mn} = k$; $k = 0, 1, \dots, L-1$, an $L \times L$ transition matrix T_h is defined for intensity transitions between adjacent pixels on a horizontal line (row of image X) from left to right such that

$$T_{\mathrm{h}} = [n_{ij}], \quad n_{ij} = \sum_{n=1}^{N-1} \sum_{m=1}^{M} [x_{mn} = i \wedge x_{mn+1} = j], \tag{1}$$

$$i, j = 0, 1, \dots, L-1,$$

where the (i, j)th element of T_{h} specifies how frequently the level i is followed by the jth level in the specified horizontal spatial displacement. Similarly, we can define a matrix T_{v} for vertical (top to bottom) transitions along the columns of the image and a matrix $T_{\mathrm{vh}} = T_{\mathrm{h}} + T_{\mathrm{v}}$ which considers both vertical and horizontal transitions.

Note that unlike the co-occurence matrices used in Weszka and Rosenfeld (1978), Haralick et al. (1973), Auja and Rosenfeld (1978), the transition matrices as defined here are in general not symmetric. This is because only right to left and top to bottom transitions are considered and the opposite senses to these are ignored. The resulting matrices still contain the same amount of information while some redundant computations are avoided.

Choosing a grey level threshold s subdivides the image into two pixel intensity classes $C2(s)$ and $C2(s)$:

$$C1(s) = \{x: x = 0, \dots, s\}, \tag{2a}$$

$$C2(s) = \{x: x = s+1, \dots, L-1\}. \tag{2b}$$

It also leads to defining four regions in the transition matrix as shown in Figure 1. For each region a set of parameters is defined giving the total number of transitions such as

$$a = \sum_{i=0}^{s} \sum_{j=0}^{s} n_{ij}, \quad b = \sum_{i=s+1}^{L-1} \sum_{j=s+1}^{L-1} n_{ij}, \tag{3a, 3b}$$

$$c = \sum_{i=0}^{s} \sum_{j=s+1}^{L-1} n_{ij}, \quad d = \sum_{i=s+1}^{L-1} \sum_{j=0}^{s} n_{ij}, \tag{3c, 3d}$$

where a, b, c and d represent the total number of transitions within $C1$, within $C2$, from $C1$ to $C2$ and from $C2$ to $C1$ respectively.

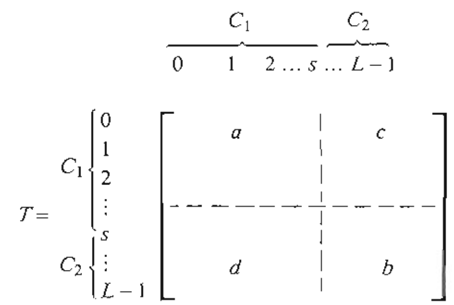

Fig. 1. Regions in the transition matrix.

3. Interaction measures and threshold selection

To evalute the 'goodness' of thresholds we define two measures of interaction between the intensity classes using the above parameters. These are estimates of the joint and conditional probabilities of intensity transition between the intensity classes which are defined by a given threshold:

$$p_{\mathrm{j}}(s) = \frac{c+d}{a+b+c+d}, \tag{4}$$

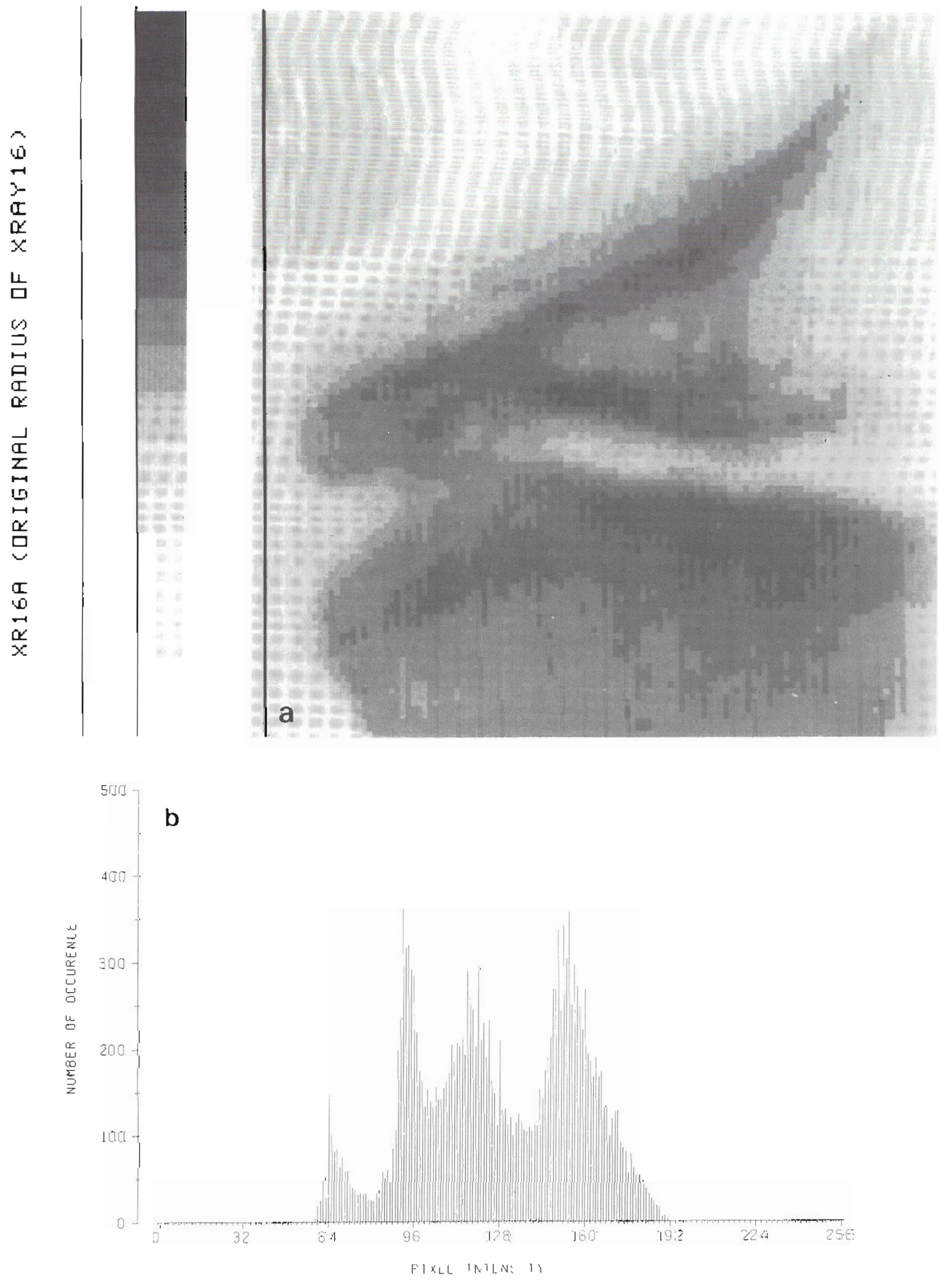

Fig. 2. (a) Multimodal image and (b) its histogram.

$$p_c(s) = \tfrac{1}{2}\left(\frac{c}{a+c} + \frac{d}{b+d}\right); \qquad (5)$$

$p_j(s)$ is a normalised measure of the total number of pixels from one of the classes that are followed by a pixel belonging to the other class. The lower its value the less is the proportion of transitions between the classes. Therefore a minimum of $p_j(s)$ would correspond to a threshold level where most transitions are within the classes and a few across them. The classes then form maximally self-contained regions with minimum transitions across separation boundaries.

An estimate of the conditional probability of transition from $C1$ to $C2$ is $c/(a+c)$ and from $C2$ to $C1$ is $d/(b+d)$. The average of these two values is used to define $p_c(s)$. The lower the value of $p_c(s)$, the lower is the probability that the next transition will be to a different intensity class.

These measures indicate the spatial discontinuity of the segmented regions. Therefore it is conjectured that meaningful sets of thresholds would correspond to the minima of the above measures.

It should be noted that $p_j(s)$ is similar to the business measure in Weszka and Rosenfeld (1978) and should have the same general shape as the grey level histogram. This is because when the image is threshold near a histogram peak, the level of interclass transitions ($c+d$ in (4)) is expected to be high while when the threshold is selected near a valley in the histogram it should be relatively low. Therefore if the histogram is unimodal, $p_j(s)$ curve would also be unimodal. $p_c(s)$ is not directly related to the histogram and it is expected that it will exhibit minima even for unimodal histograms.

4. Implementation

Experiments were conducted on a number of im-

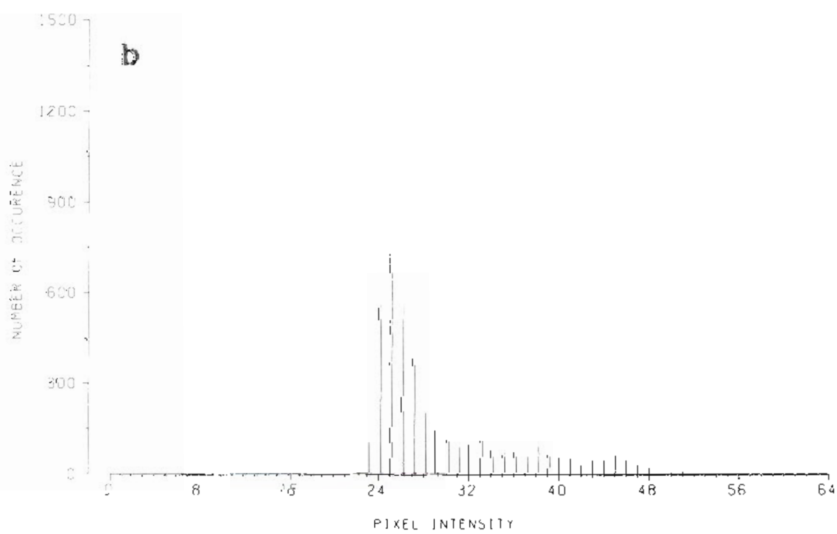

Fig. 3. (a) Unimodal image and (b) its histogram.

ages using T_h, T_v and T_{vh} matrices. Here the results using T_{vh} matrices for two images (Figures 1 and 3) are reported. The results obtained using T_h and T_v were only slightly different.

Figure 2 is a 256-level radiograph of a part of the wrist together with its multimodal histogram. Values of $p_j(s)$ and $p_c(s)$ show minima around 75, 100 and 135 corresponding to boundary levels between flesh and different regions of bones (Table 1 and Figure 2b). This set of minima is found to agree well with that obtained manually from the histogram for extracting different regional boundaries of the X-ray image (Pal and King (1981)). Also p_c exhibits a further minimum at $s=160$ which isolates the hard bone regions (palmar and dorsal surfaces (Pal and King (1981)), although this region is not separated from the rest of the histogram by a valley (Figure 2b). It is to be mentioned here that the recent algorithms based on fuzzy set theory (Pal et al. (1983)) was not able to detect this fourth minimum required for X-ray image identification.

Table 1
Values of $p_j(s)$ and $p_c(s)$ for multimodal image

s	p_j	p_c	s	p_j	p_c
70	0.0055	0.0173	125	0.0294	0.0294
75	0.0045*	0.0110*	130	0.0228	0.0227
80	0.0051	0.0111	135	0.0216*	0.0216*
85	0.0070	0.0184	140	0.0222	0.0225
90	0.0127	0.0346	145	0.0289	0.0304
95	0.0285	0.0515	150	0.0464	0.0534
100	0.0208*	0.0296	155	0.0497	0.0697
105	0.0227	0.0290*	160	0.0359	0.0671*
110	0.0265	0.0314	165	0.0295	0.0789
115	0.0326	0.0357	170	0.0214	0.0869
120	0.0372	0.0382			

* local minima

Table 2
Values of $p_j(s)$ and $p_c(s)$ for unimodal image

s	p_j	p_c	s	p_j	p_c
10	0.0001	1.000	34	0.0683	0.0936
20	0.0001	1.000	36	0.0611	0.0950
22	0.0007	0.5456	38	0.0570	0.1017
24	0.0180	0.3352	40	0.0539	0.1153
26	0.0815	0.0919*	42	0.0467	0.1210
28	0.1062	0.1078	50	0.0188	0.2295
30	0.0802	0.0892*	60	0.0008	0.3252
32	0.0731	0.0894			

* local minima

Figure 3 is a 64-level picture of a boy on a boat together with its unimodal histogram. Such histograms are typical for images with many different (small) objects such as natural outdoor scenes or aerial photographs. These images do not exhibit a clear background–foreground distinction, and a single threshold is not likely to detect most of the interesting boundaries. As expected $p_j(s)$ fails to detect a minimum for this image, while $p_c(s)$ exhibits two minima at 26 and 30. The threshold at 30 gives the lowest value of p_c and is therefore the 'best' threshold. It results in the segmentation of the image into two regions, main object (boy, buildings and bank) and background (water and air). The threshold at 26 further subdivides the background into a darker region (at the right) and a lighter region (at the left) due to reflection from the water (Figure 4).

5. Conclusions

Algorithms based on simple second order grey

Fig. 4. Segmented version of Fig. 3 corresponding to threshold levels 26 and 30.

level statistics are outlined for automatic thresholding of an image. Unlike the business measure in Weszka and Rosenfeld (1978) and the $p_j(s)$ measure, the conditional probability of interclass transitions $p_c(s)$ as defined here is seen to be independent of the grey level histogram. It is an effective criterion for threshold evaluation and selection even when the grey level histogram does not have any valleys.

Acknowledgement

Provision of data by Professor J.M. Tanner and Dr. A.G. Constantinides, and the interest of Dr. R.A. King in this work are gratefully acknowledged by the authors.

References

[1] Weszka, J.S. (1978). Survey of threshold selection techniques. *Comput. Graphics and Image Processing* 7, 259–265.

[2] Pal, S.K., R.A. King and A.A. Hashim (1983). Automatic grey level thresholding through index of fuzziness and entropy. *Pattern Recogn. Lett.* 1, 141–146.

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

[4] Haralick, R.M. et al. (1973) Textural Features for Image Classification, *IEEE Trans. Systems Man Cybernet.* 3, 610–621.

[5] Auja, N. and A. Rosenfeld (1978) A note on the use of second-order gray-level statistics for threshold selection. *IEEE Trans. Systems Man Cybernet.* 8 (12), 895–898.

[6] Pal, S.K. and R.A. King (1981). Application of fuzzy set theory in detecting X-ray edges. *Proc. IEEE ICASSP*, Atlanta, Georgia, U.S.A., Vol. 3, pp. 1125–1128.