# Correspondence

## Picture Thresholding Using an Iterative Selection Method

T. W. RIDLER AND S. CALVARD

*Abstract*—**An object may be extracted from its background in a picture by threshold selection. Ideally, if the object has a different average gray level from that of its surrounding, the effect of thresholding will produce a white object with a black background or vice versa. In practice, it is often difficult, however, to select an appropriate threshold, and a technique is described whereby an optimum threshold may be chosen automatically as a result of an iterative process, successive iterations providing increasingly cleaner extractions of the object region. An application to low contrast images of handwritten text is discussed.**

### I. Introduction

Features of interest in an image may often be extracted from their surroundings using a thresholding technique [1] in which all gray levels below the threshold are mapped into black, those levels above are mapped into white, or vice versa. The success of the technique depends on the object that is desired to be extracted occupying a range of gray levels distinct from that of the background. In practice it is difficult to select the optimum threshold, especially if a range of image scenes with widely differing properties is being considered, and some automatic means of threshold selection is required in each case. Thresholding at too high a level results in a loss of information, while thresholding at low levels can give rise to objectionable background clutter [2].

A common method of automatically deriving a threshold at which to segment a given picture is to examine its gray level histogram [3]. The presence of two peaks in such a histogram demonstrates the existence of two distinct brightness regions in the image, one corresponding to the object and the other to its surroundings. It is reasonable therefore to choose the threshold at the gray level midway between the two peaks. Establishing that the histogram is in fact bimodal and locating the peaks and hence an appropriate splitting level can cause difficulty and represents a considerable disadvantage of the technique.

Object-background discrimination ought to be improved by deriving a threshold from a series of background samples taken close enough to the object to exclude most background clutter but not close enough to include the object. Ideally the background close to the object should then be used to find a mean background gray level, and the object region should be used to provide a mean object gray level. Given these two levels, discrimination could be readily achieved assuming a contrast difference exists. The difficulty lies in choosing optimum background sampling locations.

This correspondence describes a simple automatic method of threshold selection whereby an iterative process yields successively refined threshold levels that converge to an optimum value after about four iterations. An advantage of this technique is that by establishing object and background levels, a processed picture is formed such that the object always appears white on a black background. Should further processing of the image be required, this represents a significant simplification over competing approaches in which thresholding can either result in a black object on a white background or vice versa, depending on the object to background contrast.

Thresholding is of importance in several image analysis problems, for example, in the computerized recognition of text, either machine printed or handwritten, in radiographic image enhancement for the medical field, and in nuclear physics where the enhancement of particle tracks and collisions in a bubble chamber is required. The method discussed is illustrated using the example of text enhancement under low contrast conditions.

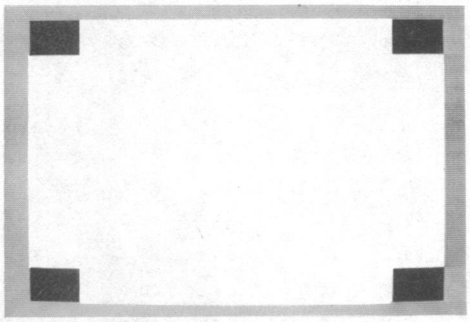


Fig. 1. Initial object-background estimation.

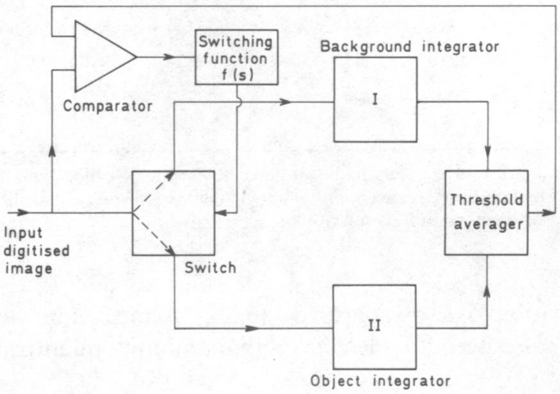



Fig. 2. Schematic image processor for iterative threshold selection.

### II. Iterative Threshold Selection

Suppose an object is located within a square image of picture elements. Without assuming any knowledge of the exact location, consider as a first approximation that the four corners of the scene contain background only and the remainder contains the object. A thresholded version of such a region is illustrated in Fig. 1. This patch may then be used as a switching function $f(s)$ to route a

Manuscript received January 3, 1978; revised March 13, 1978. UK Patent Application (38159/77).
The authors are with the Royal Aircraft Establishment, Farnborough, Hants., UK.

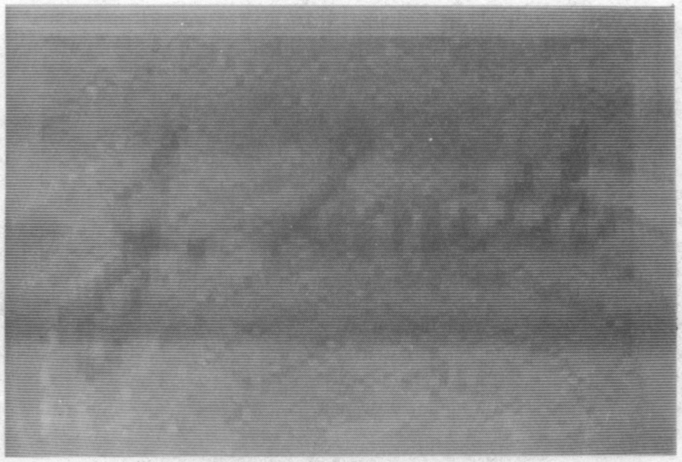


Fig. 3. Handwritten signature under low contrast conditions.

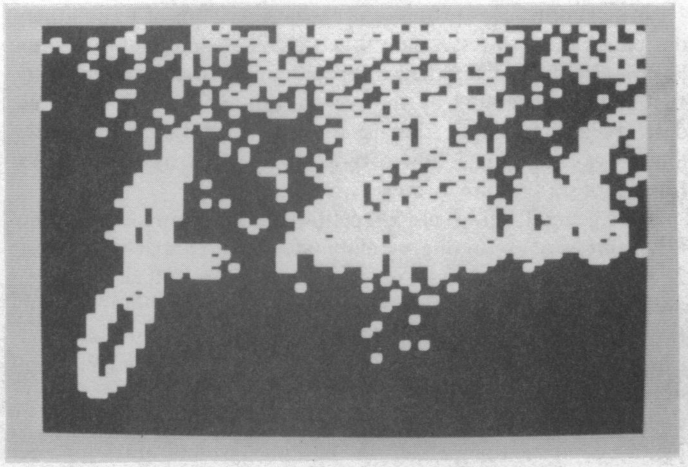
digitized image into one of two integrators. The mechanism is described with reference to the schematic diagram of the processor shown in Fig. 2. The signal controlling the switch is referred to as a switching function $f(s)$ and is in fact a thresholded (i.e., black and white) array of image points. If $f(s) = 0$, the input image signal is fed to integrator I and considered to be background. Whereas if $f(s) = 1$, integrator II receives the signal which now represents the object.

When the final picture element of the input image has been received, the integrator outputs are averaged to find a threshold, $L_0$. The acquired scene is quantized black or white according to $L_0$, and the silhouette produced is used as a new switching function for a second input of the image. This should provide a better estimate than the original switching function of what constitutes object and background and is used in just the same manner to switch the input either into the background or object integrator. A threshold $L_1$ is derived, and the picture split accordingly to produce yet another switching function. The process is repeated on the input image until the thresholded version, i.e., the switching function, remains constant for further iterations.

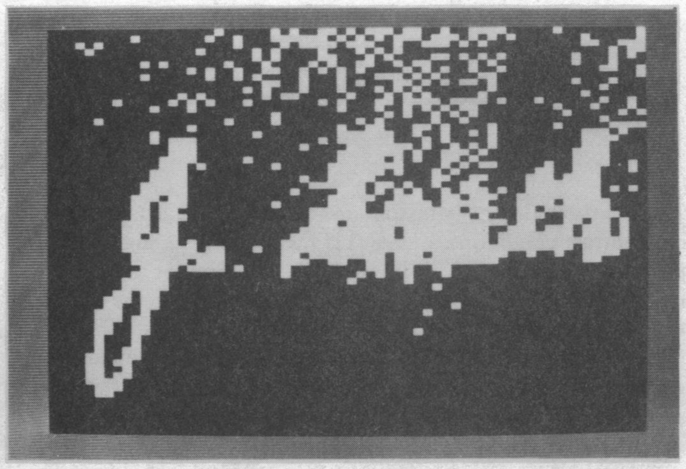
## III. Application to Text Enhancement

Fig. 3 illustrates a handwritten signature, showing how low contrast and shading effects give rise to an enhancement requirement. This image has dimensions of $64 \times 64$ picture elements, quantized to 64 gray levels. Fig. 4(a)–(d) shows thresholded versions derived from four successive iterations of the splitting level.

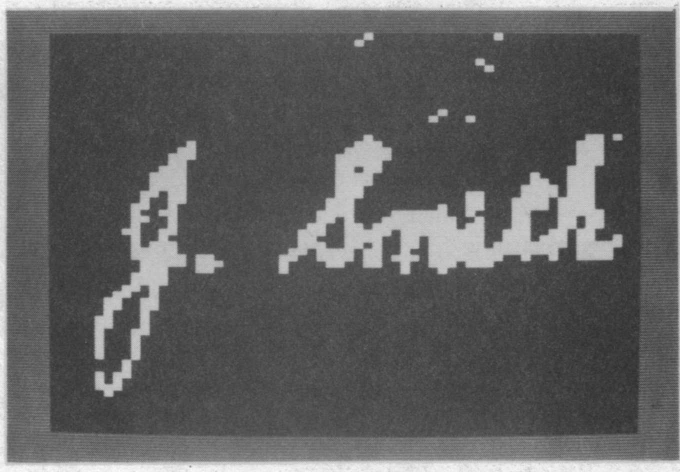
It should be observed from Fig. 4(a) that the first switching function (Fig. 1) is unable to eliminate background clutter. The residual background clutter is removed by the iterations that result in increasingly cleaner extractions of the lettering.



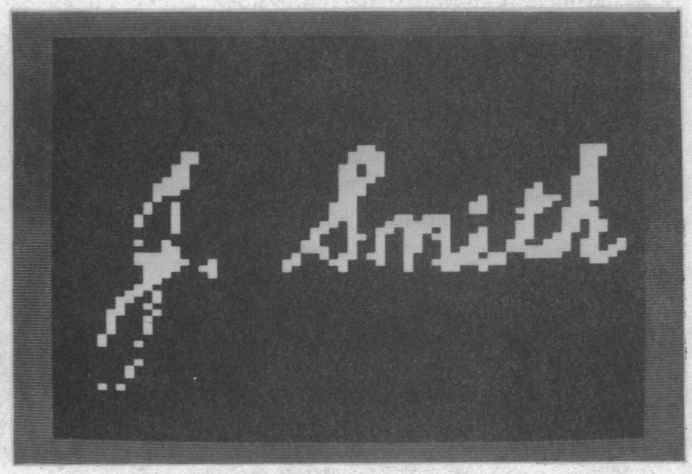
(a)



(b)



(c)



(d)

Fig. 4. Thresholded versions of Fig. 3 for four successive iterations.

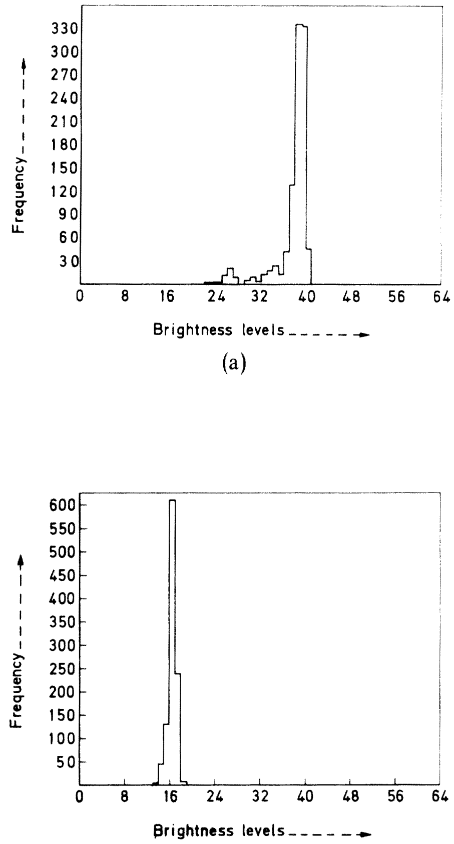



Fig. 5. Intensity distribution. (a) High contrast image. (b) Low contrast image.

Fig. 5 shows intensity histograms of the text image under differing contrast conditions. Fig. 5(a) is the result for a high contrast picture, whereas Fig. 5(b) shows the distribution for the low contrast image of Fig. 3. Although the bimodal nature of the distribution under high contrast conditions readily indicates a threshold level, the low contrast distribution provides a totally inadequate means of threshold selection. The iterative approach, however, is seen to select an appropriate threshold equally well for the two differing contrast situations.

The method is found to provide the required threshold in image scenes having peak-to-peak signal to rms noise ratios as low as 18 dB. Further experiments indicate that useful thresholds are also derived for objects against highly textured backgrounds.

## IV. Conclusions

A technique for determining a threshold at which to segment a picture has been demonstrated. Provided that a picture contains an object and background occupying different average gray levels, the iterative approach discussed provides a simple automatic method for optimum threshold selection.

## References


[1] A. Rosenfeld, *Digital Picture Processing.* New York: Academic, 1976, ch. 8.

[2] L. S. Davis, A. Rosenfeld, and J. S. Weszka, "Region extraction by averaging and thresholding," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-5, pp. 383–388, May 1975.

[3] J. S. Weszaka, R. N. Nagel, and A. Rosenfeld, "A threshold selection technique," *IEEE Trans. Comput.*, vol. C-23, pp. 1322–1326, Dec. 1974.


# A Note on the Use of Local min and max Operations in Digital Picture Processing

YASUO NAKAGAWA AND AZRIEL ROSENFELD, FELLOW, IEEE

*Abstract*—"Shrinking" and "expanding" operations on two-valued digital pictures are useful for noise removal, as well as for detecting dense regions and elongated parts of objects. For gray-scale pictures, the analogs of shrinking and expanding are local min and max operations. These operations commute with thresholding; thus if they are applied to a gray-scale picture, followed by thresholding, the result is the same as if the picture were first thresholded and shrinking and expanding were then performed. Applying min and max operations prior to thresholding thus makes it possible to defer the choice of a threshold, which may be easier to select after these operations have been performed.

## I. Shrinking and Expanding Operations

"Shrinking" and "expanding" operations on two-valued digital pictures are useful for noise removal and segmentation [1, pp. 362–365, 394–397]. Let $\Pi$ be such a picture whose points all have value 0 or 1, and suppose that $\Pi$ suffers from "salt-and-pepper noise." In other words, $\Pi$ contains regions consisting primarily of 0's with a sprinkling of isolated 1's and vice versa. We can remove isolated 1's by performing a "shrinking" operation, in which 1's are changed to 0's if they have 0's as neighbors, followed by an "expanding" operation in which 0's are changed to 1's if they have 1's as neighbors. This process shrinks large regions of 1's and then reexpands them to (essentially) their original state, but it annihilates isolated 1's since they disappear under the shrinking and cannot be regenerated by the reexpansion. In fact, this process destroys any set of 1's that is no more than two pixels wide. Similarly, expanding followed by shrinking destroys isolated 0's (or sets of 0's that are at most two pixels wide).

More generally, if we use $k$ repetitions of shrinking followed by $k$ repetitions of expanding, we eliminate sets of 1's that are at most $2k$ pixels wide. If such a set is connected and has sufficiently large area (e.g., $\geq 10k^2$), it must be elongated since its "length" (= area/width) is much greater than its width. Thus shrinking and expanding operations can be used to detect elongated parts of objects in a segmented digital picture. Another application of such operations is to the detection of clusters or dense regions in a picture composed of isolated 1's on a background of 0's (or vice versa). If we perform $k$ repeated expansions where $k$ is at least half the distance between the 1's in a cluster, the cluster will "fuse" into a solid region; when we subsequently perform $k$ repeated shrinks, this region will remain large (relative to $k^2$). On the other hand, 1's that do not belong to clusters will expand but not fuse with other 1's, so that when they are shrunk, they shrink back to single 1's.

## II. Generalization to Gray-Scale Pictures

Shrinking and expanding operations are defined only for two-valued digital pictures. Suppose that one has a picture $\Pi$ contain-

Manuscript received December 16, 1977; revised April 10, 1978. This work was supported by the U.S. Army Night Vision Laboratory under Contract DAAG53-76C-0138 (ARPA Order 3206).

Y. Nakagawa was with the Computer Science Center, University of Maryland, College Park, MD 20742. He is now with Production Engineering Research Laboratory, Hitachi Ltd., 292 Yoshida-cho, Totsuka-ku, Yokohama, Japan 244.

A. Rosenfeld is with the Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD 20742.

0018-9472/78/0800-0632$00.75 © 1978 IEEE