Correspondence

A Spatial Thresholding Method for Image Segmentation

K. V. MARDIA AND T. J. HAINSWORTH

Abstract—**There has been recent interest in the segmentation of images by thresholding. We present several model based algorithms for threshold selection. We concentrate on the important two population univariate case when an image contains an object and background. However the methods are applicable to multispectral k-population images. We show how the main ideas behind two important nonspatial thresholding algorithms follow from classical discriminant analysis. We then give various new thresholding algorithms which make use of available local/spatial information. We consider one FLIR image and two artificial examples. A comparative study indicates that a new "alternating mean thresholding and median filtering" algorithm provides an acceptable method when the image is relatively highly contaminated. This method seems to depend less on initial values.**

Index Terms—**Classification, image segmentation, iterated conditional modes, median filtering, shape analysis, spatial cluster analysis, thresholding algorithms.**

I. Introduction

We will consider the problem of image segmentation by thresholding. This problem and its importance were fully described recently in [1], [2], and [3]. Our main interest is with the $k = 2$ population case which is related to object identification. However, we shall also consider the extension to the general k-population problem.

We first show that the threshold value in the segmentation algorithms of [1] and [2] can be deduced from the well-known statistical discriminant rule. Unlike [3], their rule is not spatial, i.e., it does not use contextual information. We give a spatial allocation rule based on the work of [4] and [5]. This is utilized to give a new thresholding algorithm. We also consider the iterated conditional modes (ICM) method [6].

Section II describes the nonspatial allocation rule, and shows how the allocation rules of [1] and [2] are particular cases. We summarize their iterative thresholding method in Section III. In Section IV we give a spatial allocation rule which takes into account the spatial relationship between neighboring pixels and describe the modified iterative algorithms in Section V. In Section VI we describe ICM and its implementation. The methods are compared using synthetic images (following [3]) and one "real" FLIR (forward looking infrared) image. We conclude with a discussion of the methods in Section VIII.

Our method follows naturally from a model introduced in Section II. The method in [3] is not discussed here since it is not model orientated. Also our method applies to multispectral data, i.e., color images.

In all the algorithms considered here, we do not require any prior information of the statistical parameters, so the segmentation method can clearly be described as a "spatial clustering method" rather than a "spatial discrimination method."

Manuscript received August 8, 1986; revised January 15, 1988. Recommended for acceptance by A. K. Jain. This work was supported by the Procurement Executive, Ministry of Defence.

The authors are with the Department of Statistics, University of Leeds, Leeds LS2 9JT, England.

IEEE Log Number 8823847.

II. Nonspatial Bayes' Thresholding

Suppose that an image is known to contain k populations Π_i, $i = 1, \cdots, k$. Let p_i be the prior probability associated with population Π_i, $i = 1, \cdots, k$, respectively. Let Z be any gray-level at a pixel on a lattice. If $L_i(Z)$ is the likelihood under Π_i, $i = 1, \cdots, k$, then the Bayes' allocation rule simply assigns an observed gray-level Z to that population Π_j which maximizes the likelihood

$$p_j L_j(Z).$$

Assuming Π_i: $Z \sim N(\mu_i, \sigma_i^2)$, i.e., Gaussian with mean μ_i and variance σ_i^2, $i = 1, \cdots, k$, this just requires maximization of the score

$$S_i = \log p_i - \tfrac{1}{2}(Z - \mu_i)^2/\sigma_i^2 - \tfrac{1}{2}\log(2\pi\sigma_i^2).$$

This can be reformulated in terms of a set of threshold values. For example, if $\sigma_i = \sigma$, $i = 1, \cdots, k$ and $\mu_1 < \mu_2 < \cdots < \mu_k$, then the thresholding algorithm is:

1) Evaluate the threshold values,

$$t_{ij} = \frac{1}{2}(\mu_i + \mu_j) + \frac{\sigma^2}{(\mu_i - \mu_j)}\log(p_j/p_i), \qquad i, j = 1, \cdots, k;\ i < j. \tag{2.1}$$

2) Set $i = 1$.
For $j = 2$ to k: if $Z > t_{ij}$ then set $i = j$.
3) Assign Z to Π_i.

This algorithm requires $(k - 1)$ comparisons for each pixel of the image; this number can be further reduced by, for example, using a "split and merge" algorithm. The bulk of the computation, i.e., evaluation of the thresholds, is performed once for the entire image.

The case of $k = 2$ populations is important in practice, e.g., Π_1: background and Π_2: targets. The thresholds, which will be used later on, simplify for this case as follows.

If $\sigma_1 = \sigma = \sigma_2$, then the Bayes' allocation rule yields the single threshold value

$$t = \frac{1}{2}(\mu_1 + \mu_2) + \frac{\sigma^2}{(\mu_1 - \mu_2)}\log(p_2/p_1). \tag{2.2}$$

If $\sigma_1 \neq \sigma_2$ then the Bayes' allocation rule requires two threshold values, t_+ and t_- given by

$$t_\pm = \frac{\mu_2\sigma_1^2 - \mu_1\sigma_2^2}{(\sigma_1^2 - \sigma_2^2)} \pm \frac{\sigma_1\sigma_2}{(\sigma_1^2 - \sigma_2^2)}\left\{(\mu_1 - \mu_2)^2 + 2(\sigma_1^2 - \sigma_2^2)\log\left[\sigma_1 p_2/(\sigma_2 p_1)\right]\right\}^{1/2}. \tag{2.3}$$

In this case the allocation rule is: assign Z to Π_1 if $t^+ < Z < t^-$, otherwise assign Z to Π_2.

In practice, the population parameters μ_i, σ_i, p_i, $i = 1, \cdots, k$ are all unknown and must be estimated by some method, e.g., maximum likelihood based on some provisional allocation.

In the $k = 2$ population case, where Π_1 is background and Π_2 is a target, it is common to assume $\sigma_1 = \sigma_2$. However, in practice we expect the background to be more uniform than the target which may contain more variable gray-level information; thus, $\sigma_1 < \sigma_2$ would be more appropriate.

0162-8828/88/1100-0919$01.00 © 1988 IEEE

III. Iterative Threshold Selection Methods

We now consider iterative thresholding methods [1] and [2] for the segmentation of images containing two populations.

A. Nonspatial Thresholding (Ridler and Calvard)

The algorithm is as follows:

1) Select some initial estimate of the threshold t, e.g., we could use the average gray-level for the whole image.

2) Segment the image into two regions Π_1 and Π_2 using t: assign all pixels with gray-level less than t to Π_1, and all other pixels to Π_2.

3) Calculate the mean gray-levels $\bar{z}_1$ and $\bar{z}_2$, and the associated numbers of pixels n_1 and n_2, within the regions Π_1 and Π_2, respectively.

4) Calculate the new threshold value

$$t = \tfrac{1}{2}(\bar{z}_1 + \bar{z}_2). \tag{3.1}$$

5) Repeat steps 2), 3), and 4) until convergence is achieved, i.e., the values remain stable.

It has been proved that the method of Ridler and Calvard always terminates; see [9]. However, different starting values can lead to different solutions for the threshold value for this method. Kittler and Illingworth [10] propose a method for selecting a solution from the multiple solutions. Methods of obtaining a globally optimum solution have been considered; see [11]. It is known that the estimates of the means are inconsistent, but that the threshold is unbiased.

It is important to note that the main step is the use of the allocation rule (2.2) (with $p_1 = p_2$) in step 4).

B. Lloyd's Method

Lloyd [2] considered a modification of the iterative threshold selection method of Ridler and Calvard [1]. This replaced the calculation of the new threshold in step 4) of Section III-A with

$$t = \frac{1}{2}(\bar{z}_1 + \bar{z}_2) + \frac{\sigma^2}{(\bar{z}_1 - \bar{z}_2)} \log (n_2/n_1). \tag{3.2}$$

Here $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

It is important to note that this is simply the Bayes' allocation rule (2.4), estimating the prior probabilities by

$$p_i = n_i \Big/ \sum_j n_j, \qquad i = 1, 2. \tag{3.3}$$

These prior probabilities take into account the relative sizes of background and objects and will tend to allow more accurate thresholding of an image in which the object (or background) is relatively small.

It is possible to estimate σ from the data. However, the estimated value of σ may cause the iteration to assign all pixels to either Π_1 or Π_2. To prevent this, an upper limit

$$\hat{\sigma}^2 = \tfrac{1}{2}(\bar{z}_1 - \bar{z}_2)^2 / \log (n_1 + n_2) \tag{3.4}$$

(see [2] for details) is used to estimate σ in (3.2).

Lloyd [2] found that this modified iterative thresholding method converges to a satisfactory threshold more often than the methods of either Ridler and Calvard [1] or a slightly more complex method (see [17]). However, the output still suffers from a great deal of noise, and it is expected that this might be reduced if we use the spatial information available. For a general discussion of this rationale in image processing see [3] and [12] and the references therein.

IV. Spatial Thresholding

A. Introduction

In the previous sections we ignored the location of each pixel. Let us assume that $\boldsymbol{x}_i$, $i = 1, \cdots, n$ ($\boldsymbol{x}_i \in \mathbb{R}_2$) is the coordinate of a pixel on a regular lattice labelled in some systematic way, e.g., row by row.

Following [5], we assume that for population Π_i, $i = 1, \cdots, k$, $\{Z(\boldsymbol{x})\}$ is an isotropic Gaussian process with

$$E[Z(\boldsymbol{x})] = \mu_i, \ \text{Cov}\,[Z(\boldsymbol{x}), Z(\boldsymbol{y})] = \sigma_i^2 \rho(|\boldsymbol{x} - \boldsymbol{y}|), \tag{4.1}$$

where $\rho(h)$ is the isotropic correlation function with $\rho(0) = 1$.

Further, we assume local spatial continuity [13] in the sense that for a small neighborhood (window), these assumptions will hold at $\boldsymbol{x}$ with high probability. The "dirty pictures" of [6] satisfy such a spatial model.

Suppose that we wish to assign the observation $Z_0 = Z(\boldsymbol{x}_0)$ to one of the populations Π_i, $i = 1, \cdots, k$. Let $\{\boldsymbol{x}_p\}$, $p = 1, \cdots, s$ be some neighborhood of $\boldsymbol{x}_0$, with $Z_p = Z(\boldsymbol{x}_p)$, $p = 1, \cdots, s$, the observed values at these points. Consider some linear combination G, of the elements of this neighborhood,

$$G = \sum_{p=0}^{s} \gamma_p Z_p, \tag{4.2}$$

where γ_p, $p = 0, \cdots, s$, are the respective weights attached to Z_p, $p = 0, \cdots, s$. Write $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \cdots, \gamma_s)'$.

From the well-known properties of linear combinations of Gaussian random variables, we have that under Π_i, $i = 1, \cdots, k$, and spatial continuity G is Gaussian with

$$\mu_i^* = \nu^2 \mu_i, \ \sigma_i^{*2} = \beta^2 \sigma_i^2, \ i = 1, \cdots, k \tag{4.3}$$

where

$$\nu^2 = \mathbf{1}'\boldsymbol{\gamma}, \ \boldsymbol{P} = \left(\rho\{|\boldsymbol{x}_p - \boldsymbol{x}_q|\}\right), \ \beta^2 = \boldsymbol{\gamma}'\boldsymbol{P}\boldsymbol{\gamma}. \tag{4.4}$$

Note that for general $\boldsymbol{\gamma}$, we could have $\mathbf{1}'\boldsymbol{\gamma} < 0$ and so ν^2 would be negative. However, in this paper we shall only consider $\nu^2 > 0$. Thus we have Π_i: $G \sim N(\mu_i^*, \sigma_i^{*2})$, $i = 1, \cdots, k$. If the prior probability of Π_i is p_i, $i = 1, \cdots, k$ respectively, then we may simply write down the allocation rule and associated threshold value(s) for G using the results of Section II. (Note that (4.3) gives: $\sigma_i = \sigma_j \Leftrightarrow \sigma_i^* = \sigma_j^*$.) For example, in the $k = 2$ population case we have the following results.

1) If $\sigma_1 = \sigma_2 = \sigma$ say, then from (2.2) the spatial threshold value is

$$t^* = \frac{1}{2}\nu^2(\mu_1 + \mu_2) + \frac{\beta^2\sigma^2}{\nu^2(\mu_1 - \mu_2)} \log (p_2/p_1) \tag{4.5}$$

which in the case of equal prior probabilities ($p_1 = \frac{1}{2} = p_2$), reduces to

$$t^* = \tfrac{1}{2}\nu^2(\mu_1 + \mu_2). \tag{4.6}$$

If $\mu_1 < \mu_2$ then the spatial allocation rule is: assign Z_0 to Π_1 if $G < t^*$, otherwise assign Z_0 to Π_2.

2) If $\sigma_1 \neq \sigma_2$ then from (2.3), the spatial threshold values are:

$$t_\pm^* = \nu^2 \frac{(\mu_2\sigma_1^2 - \mu_1\sigma_2^2)}{(\sigma_1^2 - \sigma_2^2)} \pm \frac{\sigma_1\sigma_2}{(\sigma_1^2 - \sigma_2^2)} \Big\{\nu^4(\mu_1 - \mu_2)^2 + 2\beta^2(\sigma_1^2 - \sigma_2^2) \log\left[(p_2\sigma_1)/(p_1\sigma_2)\right]\Big\}^{1/2}. \tag{4.7}$$

If $\sigma_1 < \sigma_2$, then the spatial allocation rule is: assign Z_0 to Π_1 if $t_+^* < G < t_-^*$, otherwise assign Z_0 to Π_2.

These results are true for any choice of the weights $\boldsymbol{\gamma}$.

B. Spatial Thresholding

A set of weights $\boldsymbol{\gamma}$, which follow naturally from the model of Section IV-A are given by

$$\boldsymbol{\gamma} = \boldsymbol{P}^{-1}\mathbf{1}. \tag{4.8}$$

For details, see [5].

Thus, we have $\beta^2 = \nu^2$ so that

$$\sigma_i^{*2} = \nu^2\sigma_i^2, \qquad i = 1, 2. \tag{4.9}$$

Thus, we may write down the spatial threshold values by substituting $\beta = \nu$ into the results of Sections II and IV. For example,

in the $k = 2$ population case, if $\sigma_1 = \sigma_2 = \sigma$, then the spatial threshold is

$$t_1^* = \frac{1}{2}\nu^2(\mu_1 + \mu_2) + \frac{\sigma^2}{(\mu_1 - \mu_2)}\log(p_2/p_1) \quad (4.10)$$

which reduces to (4.6) in the case of equal prior probabilities. If $\mu_1 < \mu_2$ then the spatial allocation rule is: assign Z_0 to Π_1 if $G < t_1^*$; otherwise assign Z_0 to Π_2.

1) A Second Order Window: In practice we have used a 3×3 window centred on Z_0, with points arranged as shown in Fig. 1. In this case (4.8) yields (see [5]):

$$\gamma_0 = 1 - 4\beta A - 4\delta B;\quad \gamma_i = A,\ i = 1, \cdots, 4;$$
$$\gamma_j = B,\ j = 5, \cdots, 8; \quad (4.11)$$

where

$$A = (u - vw)/\{4(1 - w^2)^{1/2}(\alpha - \beta^2)^{1/2}\};$$
$$B = (v - uw)/\{4(1 - w^2)^{1/2}(\gamma - \delta^2)^{1/2}\}$$

with

$$u = (1 - \beta)/(\alpha - \beta^2)^{1/2};\quad v = (1 - \delta)/(\gamma - \delta^2)^{1/2};$$
$$w = (\epsilon - \beta\delta)/\{(\alpha - \beta^2)(\gamma - \delta^2)\}^{1/2};$$

and

$$\alpha = \tfrac{1}{4}\{1 + 2\rho(\sqrt{2}) + \rho(2)\};\quad \beta = \rho(1);$$
$$\gamma = \tfrac{1}{4}\{1 + 2\rho(2) + \rho(2\sqrt{2})\};$$
$$\delta = \rho(\sqrt{2});\quad \epsilon = \tfrac{1}{2}\{\rho(1) + \rho(\sqrt{5})\}.$$

Thus, the spatial allocation variable is simply

$$G = (1 - 4\beta A - 4\delta B)Z_0 + 4A\bar{Z}_e + 4B\bar{Z}_f \quad (4.12)$$

where $\bar{Z}_e = \frac{1}{4}(Z_1 + Z_2 + Z_3 + Z_4)$ and $\bar{Z}_f = \frac{1}{4}(Z_5 + Z_6 + Z_7 + Z_8)$, and the size parameter is

$$\nu^2 = 1 + 4A(1 - \beta) + 4B(1 - \delta). \quad (4.13)$$

Note that no account has been taken of the boundary between populations nor the boundary at the edge of the image. However, the method can be modified.

The method requires the estimation of the quantities $\rho(1)$, $\rho(\sqrt{2})$, $\rho(2)$, $\rho(2\sqrt{2})$, and $\rho(\sqrt{5})$. This can be done using the empirical estimate of the semivariogram $\gamma(h)$ in the usual way; see for example [14]. The method is applicable if there is nonstationarity in the means. If the data is otherwise nonstationary, the method is not applicable.

C. Local Mean Thresholding

A computationally simple method is to place equal weight on each element of the neighborhood of x_0, i.e., $\gamma_p = 1/(s + 1)$, $p = 0, 1, \cdots, s$, and $\nu^2 = 1$. In this case G is now the mean gray-level in the neighborhood of x_0. The appropriate threshold values may be obtained using the results of Sections II and IV. In particular, for $k = 2$ populations with $\sigma_1 = \sigma_2 = \sigma$ we have

$$t_2^* = \frac{1}{2}(\mu_1 + \mu_2) + \frac{\bar{\sigma}^2}{(\mu_1 - \mu_2)}\log(p_2/p_1) \quad (4.14)$$

where

$$\bar{\sigma}^2 = (\mathbf{1}'\boldsymbol{P}\mathbf{1})\sigma^2/(s + 1)^2. \quad (4.15)$$

Equations (4.15) and (2.2) are the same when $p_1 = p_2$. Section V presents an adaptation of the iterative method of Ridler and Calvard [1] for use in spatial thresholding when the population parameters are unknown.

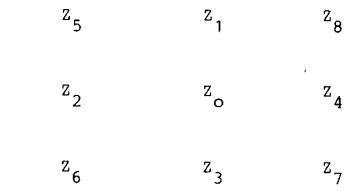


Fig. 1. Arrangement of points in a 3×3 window.

D. Extension

Note that the methods presented in this paper can be readily modified to deal with multivariate data such as multispectral data from color images. For example, we might consider $\{\boldsymbol{Z}(\boldsymbol{x})\}$ where $\boldsymbol{Z} \in \mathbb{R}_3$ and the components of $\boldsymbol{Z}(\boldsymbol{x})$ are the red, green, and blue intensity levels at point $\boldsymbol{x}$. Assuming $k = 2$ populations, with Π_i: $\boldsymbol{Z} \sim N_3(\boldsymbol{\mu}_i, \boldsymbol{\Lambda})$, $i = 1, 2$, the multivariate extension of (4.2) is

$$G = \sum_{p=0}^{s} \gamma_p \boldsymbol{Z}_p. \quad (4.16)$$

For the spatial thresholding method, (4.10) with $p_1 = p_2$ becomes

$$t^* = \tfrac{1}{2}\nu^2(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2), \quad (4.17)$$

where $\boldsymbol{\gamma}$ and ν^2 are defined as before.

The allocation rule is: assign $\boldsymbol{Z}_0$ to Π_1 if $\boldsymbol{w}'(\boldsymbol{G} - t^*) > 0$ otherwise assign $\boldsymbol{Z}_0$ to Π_2, where

$$\boldsymbol{w} = \boldsymbol{\Lambda}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2). \quad (4.18)$$

V. Iterative Spatial Thresholding and Post Smoothing

We now consider an adaptation of the iterative method of Ridler and Calvard [1], described in Section III, which makes use of the spatial threshold values of Section IV.

A. Spatial Thresholding Algorithm

$$(\sigma_i = \sigma,\ i = 1, \cdots, k;\ \mu_1 < \mu_2 < \cdots < \mu_k)$$

i) Segment the image into k regions, Π_i, $i = 1, \cdots, k$ using some initial threshold values, e.g., based on sample quantiles.

ii) Calculate the respective mean gray-level $\bar{z}_i$ and the associated number of pixels n_i for each region Π_i, $i = 1, \cdots, k$.

iii) Obtain estimates of the spatial parameters:

a) Subtract the mean $\bar{z}_i$ from the observed gray-level at each pixel in region Π_i, $i = 1, \cdots, k$, respectively.

b) Evaluate the empirical semivariogram $\hat{\gamma}(h)$ for the modified data set, and then estimate $\hat{\rho}(h)$ using the relation

$$\gamma(h) = \sigma^2\{(1 - \rho(h)\}.$$

(Note that $\hat{\gamma}(h)$ is less sensitive to trend than $\hat{\rho}(h)$.)

c) Hence, evaluate $\boldsymbol{\gamma}$ and ν^2 (from (4.12) and (4.14) for a 3×3 window).

iv) Calculate the spatial threshold values:

$$t_{1;i,j}^* = \frac{1}{2}\nu^2(\bar{z}_i + \bar{z}_j) + \frac{\sigma^2}{(\bar{z}_i - \bar{z}_j)}\log(n_j/n_i),$$
$$i, j = 1, \cdots k,\ i < j.$$

v)

a) Note that the edge pixels do not have 8 neighbors in the 3×3 window. In the present case, the missing pixels are given some suitable gray-level value. Since we might expect the pixels outside the image to be background we will set the missing edge pixels to the current estimate of the background mean, i.e., to $\bar{z}_1$.

b) For each point Z_0 in the image evaluate the quantity

$$G = \sum_{p=0}^{s} \gamma_p Z_p$$

(where $Z_0, Z_1, \cdots, Z_8$ are the observed gray-levels at points in the neighborhood of Z_0). Set $i = 1$. For $j = 2$ to k: if $G < t^*_{1;i,j}$ then set $i = j$. Allocate Z_0 to Π_i.

b) Repeat steps ii)–vi) until the solution is stable.

Note that in step iii) the empirical semivariogram is used directly to evaluate the spatial parameters. An improvement in the method may be achieved if an appropriate model, e.g., spherical scheme is fitted before parameter estimation.

Instead of v)-a) where the boundary pixels are replaced by a suitable value, it will be worthwhile to use in v)-b), the modified weights calculated from (4.9) for the reduced window.

The final thresholded image may still contain noise. Hence, we will smooth the image using median filtering (see Section V-C).

Note that if we omit step iii) and use $\gamma_0 = 1$ and $\gamma_p = 0$, $p \neq 0$, then this algorithm gives the generalization of [1] and [2] to k-populations.

B. Local Mean Thresholding Algorithm

We may use the algorithm of Section V-A with step iii) omitted, and using $\gamma_p = 1/(s+1)$, $p = 0, \cdots, s$ and $\nu^2 = 1$. Note that for $\nu^2 = 1$, (3.4) and (4.10) give $\hat{\sigma}^{*2} = \hat{\sigma}^2$. Estimating the prior probabilities as in (3.3), we replace the thresholds in iv) with

$$t^*_{2;i,j} = \frac{1}{2}(\bar{z}_i + \bar{z}_j) + \frac{\sigma^2}{(\bar{z}_i - \bar{z}_j)} \log (n_j/n_i).$$

A modified window is used at the edges, i.e., missing pixels are not estimated.

C. Post-Smoothing

The iterative thresholding methods described in Sections III-A, III-B, V-A, and V-B have been found to produce final classifications which suffer, to varying degrees, from "blobs" of noise. It is desirable to remove this noise from the thresholded image. This can be achieved to a certain extent by repeated use of some smoothing operation (see for example, [15]).

For general k-population images we shall use median filtering: this reclassifies each pixel as belonging to the median of the classifications within a 3×3 neighborhood of the pixel. For zero-one binary ($k = 2$) images, this has the same effect as either 1) majority voting or 2) mean filtering followed by thresholding at $\frac{1}{2}$. Median filtering has the desirable property that it does not blur "edges."

D. Alternating Mean Thresholding and Median Filtering

It has been found that it is desirable to "clean" a thresholded image by repeated application of median postsmoothing (using a 3×3 window) of the final classification map. Thus, these methods in effect consist of *two* processes.

We might expect an improvement in the mean thresholding method if we were to combine a cycle from each process into a single step of a new process. Thus, at each cycle of the new iterative procedure there are two subcycles:

i) a single subcycle of local mean thresholding (Section IV-C).

ii) a single subcycle of median filtering (Section V-C).

We shall call this new iterative process the Alternating Mean Thresholding and Median Filtering (AMT-MF) algorithm.

We may represent the process pictorially as shown in Fig. 2.

We now give some rationale behind steps i) and ii). In step i), an estimate $\{\hat{Y}\}$ of the classification map $\{Y\}$ is obtained as follows. For each pixel $\boldsymbol{x}_0$, let $\delta\boldsymbol{x}_0$ be some neighborhood of $\boldsymbol{x}_0$ (e.g., a 3×3 window, as in our case). Further, let

$$\bar{Z}_0 = \sum_{\boldsymbol{x} \in \delta_{\boldsymbol{x}_0}} Z(\boldsymbol{x}) \Big/ \sum_{\boldsymbol{x} \in \delta_{\boldsymbol{x}_0}} 1.$$

Then each pixel $\boldsymbol{x}_0$ is allocated to population $\Pi_{\hat{Y}_0}$ such that the conditional probability density function

$$f\left(\bar{Z}_0 \middle| \{Z(\boldsymbol{x}): \boldsymbol{x} \in \delta_{\boldsymbol{x}_0}\} \subset \Pi_{\hat{Y}_0}\right)$$

is maximized; see Section V-B. Note that $\bar{Z}_0$ is replaced by G, defined in (4.12), for the spatial case.

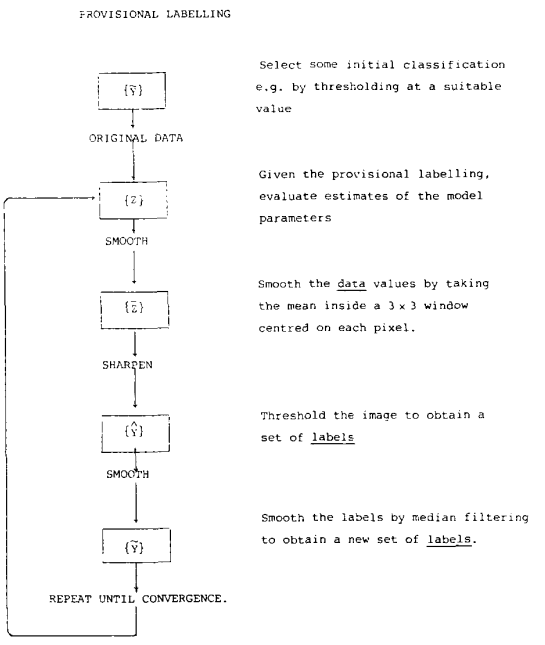


Fig. 2. AMT-MF algorithm.

In step ii), a new classification map $\{\tilde{Y}\}$ is then obtained by assigning pixel $\boldsymbol{x}_0$ to that population $\Pi_{\hat{Y}_0}$ such that the conditional probability density function

$$f\left(\tilde{Y}_0 \middle| \{\hat{Y}_j: \boldsymbol{x}_j \in \delta_{x0}\}\right) \propto \exp\left[-\beta \sum_{\boldsymbol{x}_j \in \delta_{\boldsymbol{x}_0}} |\hat{Y}_j - \tilde{Y}_0|\right], \beta > 0,$$

is maximized. This simply assigns $\boldsymbol{x}_0$ to the population corresponding to the median of the set $\{\hat{Y}_j: \boldsymbol{x}_j \in \delta_{\boldsymbol{x}_0}\}$. In the $k = 2$ population case, this has the same effect as a majority vote rule, cf. the prior of [6]. Note that there are no restrictions on β (other than $\beta > 0$) for this method, since steps i) and ii) are performed in alternation and we are not, therefore, looking for a tradeoff between them (cf. ICM method; see Section VI below).

Note that this method is similar in principle to an iterative presmoothing method used in [15]. Their class probabilities are not based on a spatial model and they use an ad hoc selective mean filter in step ii). Note that our method uses the original observations at each iteration, which is more in the spirit of [6].

For the purpose of comparison we shall also consider the "partially spatial" Alternating Nonspatial Thresholding and Median Filtering (ANST-MF) method, which uses a subcycle of nonspatial thresholding (Section II) and a subcycle of median filtering at each iteration.

1) AMT-MF Algorithm: We may adapt algorithm 5.2 to perform AMF-MF: insert the following between steps v) and vi):

v)′ smooth the thresholded image by assigning each pixel to the median of the classifications within a 3×3 window centred on the pixel.

VI. Iterated Conditional Modes (ICM) Method

A. The Method

In the simplest form, the method iteratively minimizes ([6])

$$S_i = \frac{1}{2\sigma_i^2}(Z_0 - \mu_i)^2 - \beta u_i + \log \sigma_i, \quad i = 1, \cdots, k, \quad (6.1)$$

where u_i is the number of neighbors in a 3×3 window allocated to Π_i at an iteration. Note that this β is not the same as used in Section IV.

For the $k = 2$ population case, we have the following results.

i) If $\sigma_1 = \sigma_2 = \sigma$ say, then the boundary of the decision rule is given by $S_1 = S_2$, which gives

$$Z = \tfrac{1}{2}(\mu_1 + \mu_2) + \beta\sigma^2(u_2 - u_1)/(\mu_1 - \mu_2). \quad (6.2)$$

Thus, we define the decision variable

$$B = Z_0 + \beta\sigma^2(u_1 - u_2)/(\mu_1 - \mu_2), \quad (6.3)$$

and the threshold value

$$t_3 = \tfrac{1}{2}(\mu_1 + \mu_2). \quad (6.4)$$

If $\mu_1 < \mu_2$, then the allocation rule is: assign Z_0 to Π_1 if $B < t_3$, otherwise assign Z_0 to Π_2.

ii) If $\sigma_1 \neq \sigma_2$, then the boundary $S_1 = S_2$ has two solutions for Z:

$$Z = \frac{\mu_2\sigma_1^2 - \mu_1\sigma_2^2}{(\sigma_1^2 - \sigma_2^2)} \pm \frac{\sigma_1\sigma_2}{(\sigma_1^2 - \sigma_2^2)}\left\{(\mu_1 - \mu_2)^2 + 2(\sigma_1^2 - \sigma_2^2)\right.$$
$$\left.\cdot\left[\beta(u_1 - u_2) + \log(\sigma_2/\sigma_1)\right]\right\}^{1/2}. \quad (6.5)$$

Note that the RHS varies with each pixel x_0, since it depends on the counts u_1, u_2. Hence, the RHS is not a threshold in the usual sense. It is not clear whether a thresholding rule based on (6.5) will be computationally more efficient than the comparison of the scores (6.1) at each pixel, especially for $k > 2$ populations. However, it seems likely that for $k = 2$, $\sigma_1 \neq \sigma_2$, the latter will be better than the former.

B. Algorithm ($\sigma_i = \sigma$, $i = 1, \cdots, k$; $\mu_1 < \cdots < \mu_k$)

The thresholding algorithm of Section V-A can be modified to use the local reconstruction method considered in Section VI-A. Delete steps iii)-b), iii)-c), iv), and v)-b), and use in their place the following:

iii)-b)′ Evaluate the variance $\hat{\sigma}^2$, of the modified data set.

iv)′ Evaluate the threshold values $t_{3;i,j} = \frac{1}{2}(\bar{z}_i + \bar{z}_j)$, $i, j = 1, \cdots, k$, $i < j$.

v)-b)′ For each pixel x_0 in the image do the following:

Set $i = 1$.

For $j = 2$ to k:

If $Z_0 + \beta\hat{\sigma}^2(u_i - u_j)/(\bar{z}_i - \bar{z}_j) > t_{3;i,j}$, set $i = j$: next j.

Allocate x_0 to Π_i.

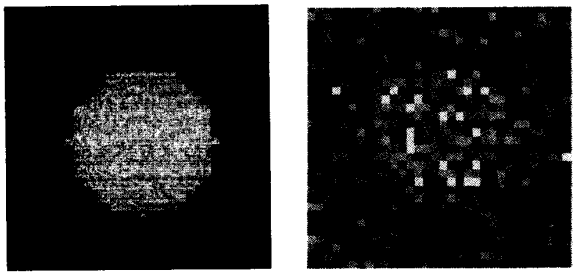
The value of β is chosen by the user. Setting $\beta = 0$ gives the nonspatial method [1], whereas letting $\beta \to \infty$ gives a majority-vote rule. In practice, we choose a value between these extremes. A value of $\beta = 1.5$ was found to work well.

VII. Comparative Study of the Methods

We now present several examples of the application of the thresholding methods considered earlier. In the first two examples, the "naive error" is that observed under the "naive rule," namely, the noniterative discriminant rule based on population parameters. This rule is used to provide an indication of the level of noise in the images.

Example 1

Following [3], we consider a 32 × 32 synthetic set image consisting of two populations, Π_1: background and Π_2: a single circular disk with respective gray-levels $\mu_1 = 100$ and $\mu_2 = 120$. The area of the disk (253 pixels) is approximately one quarter of the total area of the image. Fig. 3(a) shows the "pure" image; the shape of the disc has been distorted to some extent in the process of discretization. We consider 25 simulations obtained by superimposing independent Gaussian $N(0, \sigma^2)$ noise for each value of $\sigma = 2.5$, 5, 10, 15 , 20, and 30 respectively. These values of σ yield average percentage misclassifications of 0.004, 1.8, 14.2, 23.9, 29.7, and 36.0, respectively using the "naive rule." Tables I and II show the mean percentage misclassification errors (and associated standard deviations) for various thresholding methods



(a) (b)

Fig. 3. (a) "Pure" 32 × 32 image containing $k = 2$ populations. (b) Image of Fig. 3(a) corrupted by $N(0, 400)$ noise.

TABLE I

Mean Percentage Misclassifications (with Associated Standard Deviations Shown in Brackets) for Various Thresholding Methods as Applied to 25 Synthetic Images of $k = 2$ Populations with Respective Means $\mu_1 = 100$ and $\mu_2 = 120$, Which Have Been Corrupted by Superimposing Independent Gaussian $N(O, \sigma^2)$ Noise with $\sigma = 2.3$, 5, 10, 15, and 20. In Each Case the Initial Threshold $t_0 = 110$ Is Used.

σ	2.5	5	10	15	20	30
Naive Error	0.004 (0.02)	1.8 (0.4)	14.2 (1.1)	23.9 (1.3)	29.6 (1.2)	36.0 (1.4)
Ridler and Calvard	0 (0)	2.1 (0.4)	19.2 (1.9)	29.2 (1.6)	34.4 (2.0)	39.9 (2.1)
Lloyd	0.02 (0.04)	1.8 (0.5)	14.5 (1.4)	25.48 (2.26)	32.3 (2.7)	38.6 (3.2)
Local mean (equal priors)	0.48 (0.1)	0.74 (0.2)	1.9 (0.3)	4.6 (0.9)	10.9 (2.4)	22.8 (3.3)
Local mean (estimated priors)	1.2 (0.2)	1.6 (0.3)	2.4 (0.5)	4.3 (0.6)	7.9 (1.0)	17.9 (4.3)
Spatial	0.46 (0.1)	0.68 (0.2)	1.8 (0.4)	4.4 (0.9)	9.8 (1.9)	20.7 (3.5)
ANST-MF	0.39 (0)	0.5 (0.2)	2.6 (0.7)	8.4 (1.5)	16.4 (3.2)	26.1 (4.6)
AMT-MF	0.47 (0.1)	0.70 (0.2)	1.6 (0.4)	2.7 (0.5)	4.9 (1.0)	11.0 (2.4)
ICM (β=1.5,fixed)	0 (0)	0.1 (0.1)	1.0 (0.5)	2.2 (0.5)	3.7 (1.2)	7.5 (2.3)

TABLE II

Mean Percentage Misclassifications (with Associated Standard Deviations Shown in Brackets) for Various Thresholding Methods as Applied to 25 Synthetic Images of $k = 2$ Populations with Respective Means $\mu_1 = 100$ and $\mu_2 = 120$, Which Have Been Corrupted by Superimposing Independent Gaussian $N(O, \sigma^2)$ Noise with $\sigma = 2.5$, 5, 10, 15, and 20. In Each Case the Initial Threshold Value Is $t_0 = \bar{Z}$, the Average Gray-Level of the Whole Image.

σ	2.5	5	10	15	20	30
Naive Error	0.004 (0.02)	1.8 (0.4)	14.2 (1.1)	23.9 (1.8)	29.6 (1.2)	36.0 (1.4)
Ridler and Calvard	0 (0)	2.7 (0.7)	22.1 (2.1)	31.2 (1.6)	34.4 (2.0)	40.4 (2.0)
Lloyd	0.02 (0.04)	1.8 (0.4)	19.0 (2.6)	30.8 (1.9)	35.5 (2.0)	40.5 (2.5)
Local mean (equal priors)	0.48 (0.1)	0.8 (0.2)	1.9 (0.3)	4.6 (1.0)	11.2 (2.7)	23.5 (3.6)
Local mean (estimated priors)	1.1 (0.2)	1.6 (0.3)	2.4 (0.5)	4.3 (0.6)	7.9 (1.1)	19.3 (4.9)
Spatial	0.46 (0.1)	0.7 (0.2)	1.8 (0.4)	4.4 (0.9)	9.9 (2.0)	21.3 (4.0)
ANST-MF	0.4 (0)	0.5 (0.15)	2.8 (0.8)	10.4 (2.9)	18.7 (4.2)	28.0 (4.2)
AMT-MF	0.47 (0.1)	0.7 (0.2)	1.6 (0.4)	2.7 (0.5)	4.9 (1.1)	11.2 (2.6)
ICM (β=1.5, fixed)	0 (0)	0.1 (0.1)	1.2 (0.8)	4.8 (2.8)	10.5 (3.3)	20.4 (5.9)

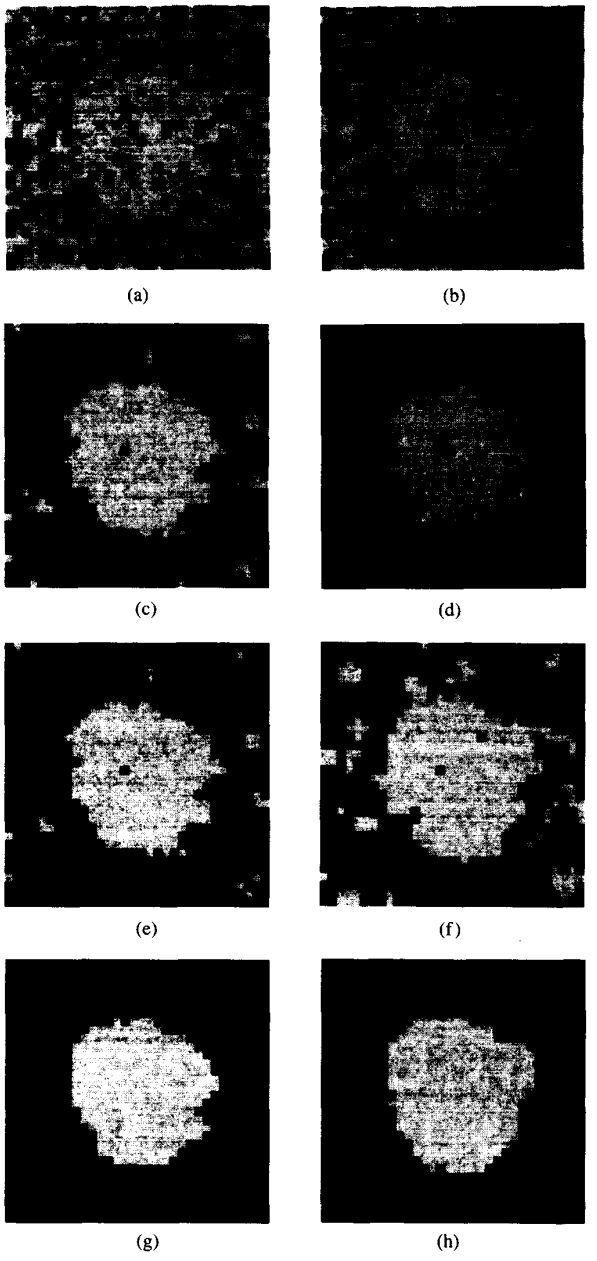


Fig. 4. Image of Fig. 3(b) thresholded using the following methods: (a) nonspatial thresholding (Ridler and Calvard), (b) nonspatial thresholding (Lloyd), (c) local mean (equal priors), (d) local mean (estimated priors), (e) spatial thresholding (Section V-A), (f) ICM ($\beta = 1.5$, fixed), (g) ANST-MF, and (h) AMT-MF.

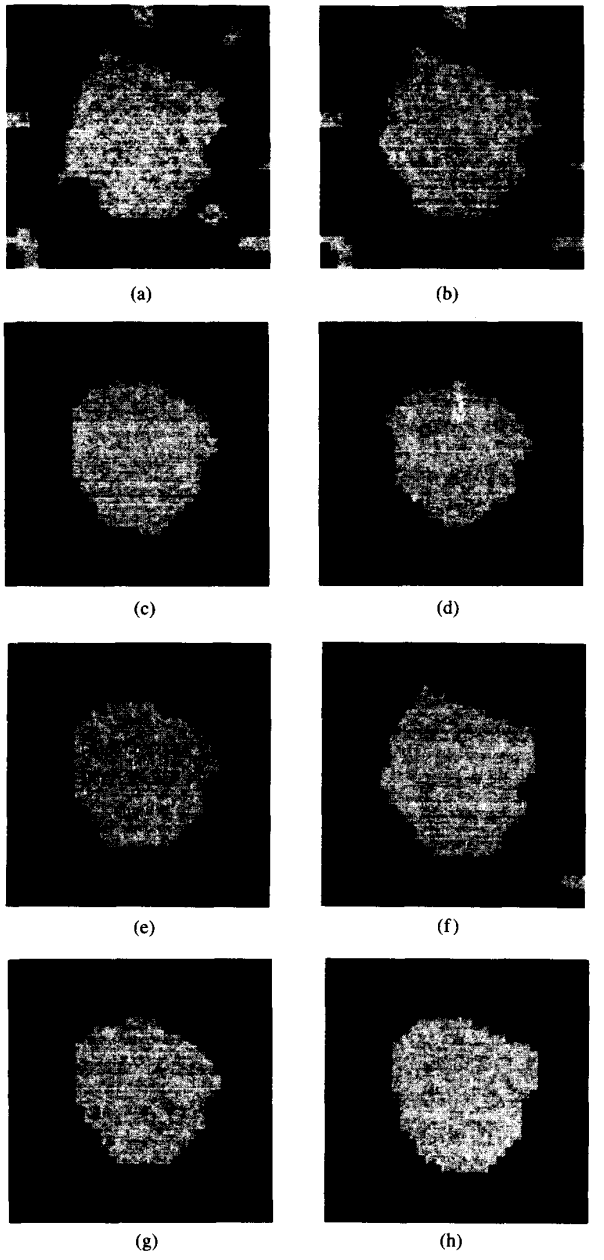
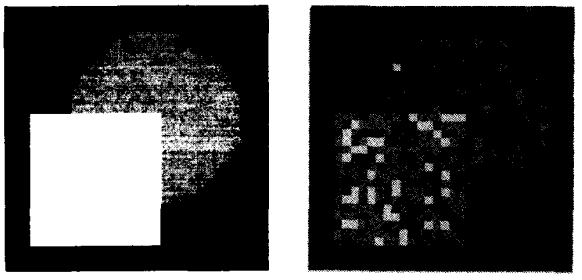


Fig. 5. (a)-(h) show the respective images of Fig. 4 after three iterations of median postsmoothing.

using the initial threshold values $t_0 = \frac{1}{2}(\mu_1 + \mu_2) = 110$ and $t_0 = \overline{Z}$ (the average gray-level of the full image), respectively. For illustrative purposes, we also include results for one further simulated image; the noisy image, with naive error 28.9 percent (generated using $\sigma = 20$), is shown in Fig. 3(b). The results of applying various thresholding methods to the image in Fig. 3(b) are shown in Figs. 4(a)-(h).

All the thresholding methods are, to some extent, dependent on starting values. However, Besag's ICM ($\beta = 1.5$, fixed) seems to be critically dependent on initial values: a "good" initial threshold yields excellent results whereas a "poor" initial threshold yields disappointing results. This effect can be reduced slightly if we make an assumption of the form—"all pixels outside the image are background pixels." However, the error is still greatly affected by the starting values and the latter assumption is not, in general, appropriate. For "poor" initial values the basic spatial methods (i.e., local mean, spatial) are roughly as effective as ICM—particularly for high σ. However, the "best" results on the whole are obtained by AMT-MF; in particular, for $t_0 = \overline{Z}$ and $\sigma > 10$ the results obtained by AMT-MF have less than half the error rate of ICM. AMT-MF seems to be less dependent on starting values than the other methods.

There is very little difference between the performance of methods using equal priors and Lloyd's modification, respectively. It



(a) (b)

Fig. 6. (a) "Pure" 32 × 32 image containing $k = 3$ populations. (b) Image of (a) corrupted by $N(0, 100)$ noise.

remains to be seen whether this is the case with a relatively small object (or background).

Note that both nonspatial methods suffer greatly from noise and the object appears bloated; see Figs. 4(a) and 4(b). These methods, as expected, perform best when the image is relatively uncontaminated, but are wholly unsuitable when the image is noisy. They converge in around 4 to 5 iterations to a solution.

The spatial methods of Section IV produced a thresholded image of much better quality than did the nonspatial methods; see Figs. 4(c), 4(d), and 4(e). The thresholded images still suffer, to a lesser extent, from noise. The spatial method produces "better" results than the local mean method (equal priors) at the expense of increased computational burden. After postsmoothing these methods give very similar results.

ICM ($\beta = 1.5$, fixed) produced the lowest misclassification errors, over the range of values for σ considered, using the initial threshold $t_0 = 110$. However, AMT-MF produced "similar" results. When the starting value, $t_0 = \bar{Z}$, was used, ICM produced disappointing results with high variability for relatively high σ.

ICM may not converge to a *unique* solution. For a majority of the simulations ICM lead to two solutions, but the difference between solutions is small.

The AMT-MF method performs most "reliably" of all the methods considered here. It has a relatively small error variance and is not affected to a significant level by change of starting values. The drawback of the method is that boundaries tend to be smoothed and fine detail lost; see also example 2.

Example 2

Our main interest is with the $k = 2$ population case. However, we now consider the application of the various methods described in this paper to a synthetic test image consisting of $k = 3$ populations, Π_1: background, Π_2: circle (partially obscured) and Π_3: square, with respective gray-levels $\mu_1 = 100$, $\mu_2 = 120$, and $\mu_3 = 140$. Fig. 6(a) shows the "pure" image. We consider 25 simulations obtained by adding independent Gaussian $N(0, \sigma^2)$ noise for each of the following values $\sigma = 2.5$, 5, 10, 15, and 20, respectively. These values of σ give average percentage misclassifications 0.02, 2.7, 18.8, 30.5, and 37.4 respectively under the "naive rule." Fig. 6(b) shows a particular "noisy" image with naive error 20.6 percent (obtained using $\sigma = 10$). Tables III and IV show the mean error rates (and associated standard deviations) for various thresholding methods using initial values based on the population parameters and initial values based on quantiles respectively. For illustrative purposes, Figs. 7(a)–(h) show the results of applying the various thresholding methods to the image shown in Fig. 6(b).

The nonspatial methods [see Figs. 7(a), (b)] produce their best results when σ is "low," but are totally unacceptable for "high" σ. The classified images have errors distributed "randomly" throughout the image, although the use of postsmoothing would confine most errors to interpopulation boundaries. The classification errors for methods utilizing local/spatial information occur mainly at interpopulation boundaries where the assumption of spatial continuity breaks down; see Figs. 7(c)–(e) and (h). This departure is most evident at the Π_1–Π_3 boundary where pixels are almost invariably misclassified as belonging to Π_2. This is because the spatial allocation variable G, being a weighted average of values in the neighborhood of a pixel, has expected value closest to μ_2 along the Π_1–Π_3 boundary. Relatively few errors are due to misclassification of pixels as *belonging* to either Π_1 or Π_3. Further, populations Π_1 and Π_3 having the greatest separation of means, are rarely confused.

The "partially spatial" ANST-MF method seems to produce acceptable results for "low" σ. However, for relatively highly contaminated images the spatial methods produce preferable solutions.

TABLE III

MEAN PERCENTAGE MISCLASSIFICATIONS (WITH ASSOCIATED STANDARD DEVIATIONS SHOWN IN BRACKETS) FOR VARIOUS THRESHOLDING METHODS AS APPLIED TO 25 SYNTHETIC IMAGES OF $k = 3$ POPULATIONS WITH RESPECTIVE MEANS $\mu_1 = 100$ AND $\mu_2 = 120$, AND $\mu = 140$, WHICH HAVE BEEN CORRUPTED BY SUPERIMPOSING GAUSSIAN $N(O, \sigma^2)$ NOISE WITH $\sigma =$ 2.5, 5, 10, 15, AND 20. IN EACH CASE THE INITIAL THRESHOLDS ARE BASED ON THE POPULATION PARAMETERS, i.e., $t_{ij} = \frac{1}{2}(\mu_i + \mu_j)$, $i, j = 1, \cdots, 3$.

σ	2.5	5	10	15	20
Naive Error	0.004 (0.02)	2.6 (0.5)	18.8 (0.9)	30.5 (0.9)	37.4 (1.4)
Ridler and Calvard	0 (0)	2.7 (0.4)	21.4 (2.0)	35.0 (1.8)	42.9 (1.8)
Lloyd	0.004 (0.02)	2.6 (0.3)	20.0 (1.4)	34.6 (2.4)	43.3 (1.9)
Local mean (equal priors)	8.6 (0.1)	8.0 (0.2)	9.1 (2.4)	14.1 (2.4)	25.2 (5.7)
Local mean (estimated priors)	7.9 (0.2)	8.2 (0.4)	8.7 (0.5)	12.2 (2.1)	22.5 (7.0)
Spatial	7.5 (0.1)	7.8 (0.2)	8.9 (0.6)	18.8 (2.2)	26.5 (8.4)
ANST-MF	0.6 (0.00)	0.9 (0.2)	5.1 (0.9)	12.1 (1.9)	20.1 (3.5)
AMT-MF	8.1 (0.1)	8.3 (0.3)	8.6 (0.7)	9.98 (1.2)	14.1 (3.2)
ICM(β=1.5)	0 (0)	0.1 (0.1)	1.2 (0.3)	4.2 (1.5)	9.4 (2.4)

TABLE IV

MEAN PERCENTAGE MISCLASSIFICATIONS (WITH ASSOCIATED STANDARD DEVIATIONS SHOWN IN BRACKETS) FOR VARIOUS THRESHOLDING METHODS AS APPLIED TO 25 SYNTHETIC IMAGES OF $k = 3$ POPULATIONS WITH RESPECTIVE MEANS $\mu_1 = 100$, $\mu_2 = 120$, AND $\mu = 140$, WHICH HAVE BEEN CORRUPTED BY SUPERIMPOSING GAUSSIAN $N(O, \sigma^2)$ NOISE WITH $\sigma =$ 2.5, 5, 10, 15, AND 20. IN EACH CASE THE INITIAL THRESHOLDS ARE BASED ON SAMPLE QUANTILES.

σ	2.5	5	10	15	20
Naive Error	0.004 (0.02)	2.6 (0.5)	18.8 (0.9)	30.5 (1.4)	37.4 (1.4)
Ridler and Calvard	0 (0)	3.0 (0.6)	26.3 (2.0)	37.2 (1.7)	43.9 (1.5)
Lloyd	0.004 (0.02)	2.6 (0.4)	25.0 (2.8)	37.9 (2.0)	45.4 (1.7)
Local mean (equal priors)	7.6 (0.1)	8.0 (0.2)	9.2 (0.5)	16.6 (4.3)	27.5 (6.9)
Local mean (estimated priors)	7.9 (0.2)	8.1 (0.4)	8.8 (0.6)	14.4 (3.6)	28.6 (7.7)
Spatial	7.5 (0.1)	7.8 (0.2)	9.0 (0.6)	16.6 (4.7)	30.4 (8.2)
ANST-MF	0.6 (0)	0.9 (0.2)	5.2 (0.9)	14.3 (3.5)	22.7 (3.9)
AMT-MF	8.1 (0.1)	8.3 (0.2)	8.6 (0.7)	9.98 1.2	19.4 (10.1)
ICM (β=1.5, fixed)	0 (0)	0.09 (0.09)	2.8 (1.5)	9.86 (4.0)	17.3 (5.3)

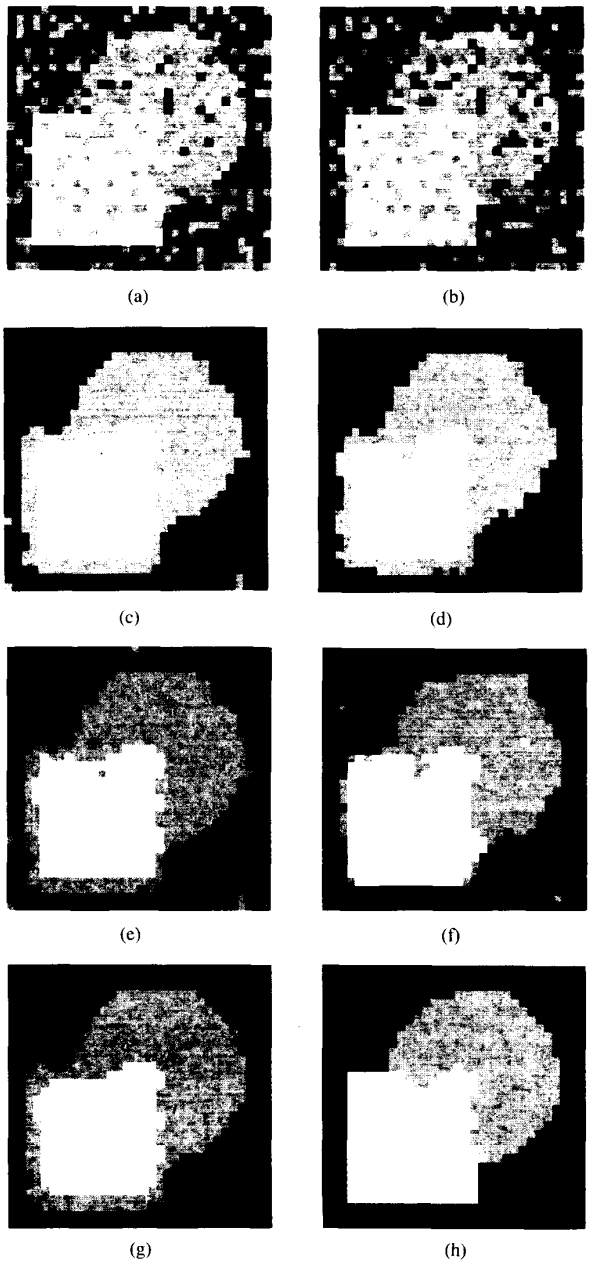


Fig. 7. Results of thresholding Fig. 6(b), with (a)-(h) as in Fig. 4.

In the $k = 3$ population case, ICM ($\beta = 1.5$, fixed) appears to produce the best results. However, the quality of the solution is greatly affected by the initial values; for example, for "large" σ and a "poor" starting value the results for ICM and AMT-MF are similar.

We have kept β fixed at 1.5 for each iteration of ICM to minimise the computational complexity to a level which is similar to that involved in AMT-MF. It should be noted that whilst ICM ($\beta = 1.5$, fixed) is less robust to changes in the initial classification than the other methods (AMT-MF in particular), this may not be the case if β is allowed to vary, e.g., if β is estimated at each iteration.

Most of the errors incurred by the methods using local/spatial information is found at the object/background boundary. In particular, single pixels protruding from the object boundary are misclassified as background because of the smoothing effect. It should be worthwhile incorporating some edge-detecting algorithm (see [16]), but again the computational complexity should be borne in mind.

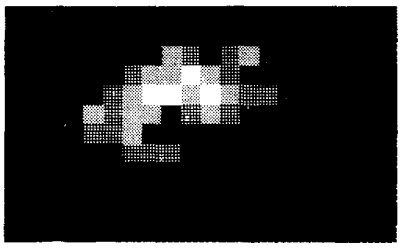


Fig. 8. 20 × 12 subimage containing an object (with "hot-spot") and background (with "clutter") extracted from a 512 × 384 image.

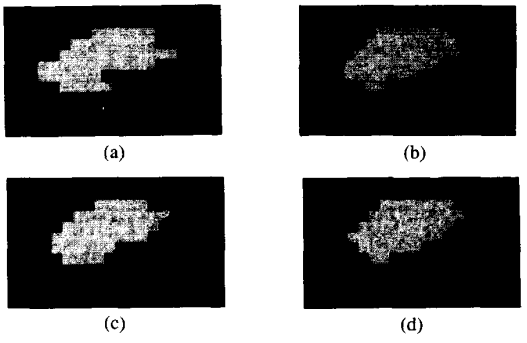


Fig. 9. Image of Fig. 8 thresholded using the following methods: (a) Nonspatial thresholding (Ridler and Calvard), (b) local mean (equal priors), (c) ANST-MF, and (d) AMT-MF. In each case we assume $k = 2$ populations.

Example 3

Fig. 8 shows a 20 × 12 subimage containing an object and background taken from a 512 × 384 FLIR image of a scene. Fig. 9 shows the results of segmenting Fig. 8 into $k = 2$ populations using the following methods: a) Nonspatial thresholding (Section III-A), b) local mean thresholding (Section V-B), c) ANST-MF, and d) AMT-MF.

Note that the observed image (Fig. 8) contains relatively little noise, and so the segmented images do not suffer from "salt and pepper" noise. Consequently, the relative performance of the nonspatial methods is improved. The segmented images are all similar, but the local/spatial methods tend to produce more rounded objects than the nonspatial methods.

It can be argued that the image in Fig. 8 is really composed of $k > 2$ populations: the object containing regions of different brightness ("hot-spots") and the background containing "clutter" (particularly at the top of the image). However, if we use $k = 3$ for example, then the cluster in the background is confused with the object by methods a) and b). Surprisingly, method c) produces the best results on this occasion.

VIII. Discussion

The iterative selection method is a useful technique for determining threshold values. It is relatively simple to implement and requires relatively little computer storage or computer time. This is of great importance for real-time applications. The final output is only a local optimum solution which depends on the guessed initial parameter values. (The argument is similar to that followed in [6].) A global optimum solution could be obtained by evaluating all possible local optimum solutions and then choosing the best of

these according to some criterion, for example that solution with maximum likelihood; or see [10].

Our main concern is with the univariate $k = 2$ population problem when the image contains an object and background. In this case, those methods which utilize local/spatial information present in the image are superior to the nonspatial methods. AMT-MF produces acceptable results which are not greatly affected by starting values. However, ICM does produce better results when a "good" starting value is used. Note that in practice we have no knowledge of the population parameters and so "good" starting values cannot be guaranteed. Thus, in practice we would recommend the use of AMT-MF for $k = 2$.

When the image contains $k \geq 3$ populations then the local/spatial methods are susceptible to boundary confusion. This is most evident at low noise levels when boundaries between populations can be "outlined" with an intermediate population (in the sense that the populations are ordered by magnitude of their respective means). At relatively "low" noise levels it seems that a "partially" spatial method (ANST-MF) produces acceptable results. However, ICM produces the best results over the full range of noise levels considered, although these are greatly affected by starting values.

The "real" image considered in this paper was relatively uncontaminated but did contain $k \geq 3$ populations and background clutter. Therefore, it is not surprising that the spatial methods produced rather disappointing results. Further analysis, of more "highly" contaminated images containing various numbers of populations and background clutter, is required before the practical use of these methods can be assessed. It is also of interest to see how the methods behave when the population variances are unequal. It should be noted that the segmentation of the image is regarded as only a first step in the analysis of the *contents* of the image and any assessment of performance should also take into account the later stages of (unsupervised) image analysis.

We have found that median postsmoothing improves the final classification, with the greatest improvement at the first application. Our results show that "better" results are obtained if the median filtering is incorporated into an alternating process with one of the basic thresholding methods. For two populations and/or high noise levels AMT-MF produces acceptable results.

Note that all the methods are liable to identify more than one population when the image contains only background if we wrongly assume $k \geq 2$ when thresholding. The ICM and AMT-MF methods fare better than the non-spatial methods, producing at worst only "small blobs." However, the nonspatial methods suffer from high levels of random noise; further, if post-smoothing is applied, large "objects" are formed. This demonstrates the need for prior knowledge of k.

There is little to choose between the local mean and the spatial thresholding methods; however, the simplicity of the former may be an advantage. The spatial method may possibly be improved by fitting a suitable model for the semivariogram. However, our work has been limited to the cases where there is no spatial correlation a priori and $\sigma_i^2 = \sigma^2$, for all i.

ICM($\beta = 1.5$, fixed) produces the "best" image if the initial approximation is "good" but can produce disappointing results if a "poor" initial approximation is used. However, the AMT-MF method produces acceptable results and is less dependent on starting values. Note that all the thresholding methods considered here converge to "local" solutions, and so all are to some extent dependent on starting values.

The results for ICM may be improved by allowing β to vary at each iteration, although the computational complexity will be increased; different priors using the shape/size of the object may also be useful. All of the methods may benefit from the inclusion of some edge-detecting routines (see [16]), at the expense of an increased computational burden.

ACKNOWLEDGMENT

The authors are grateful to Dr. T. Ridler and Mr. D. E. Lloyd for their help and to the referees and editor for helpful suggestions.

REFERENCES

[1] T. W. Ridler and S. Calvard, "Picture thresholding using an iterative selection method," *IEEE Trans. Syst., Man., Cybern.*, vol. SMC-8, pp. 630–632, Aug. 1978.
[2] D. E. Lloyd, "Automatic target classification using moment invariants of image shapes," Farnborough, UK, Rep. RAE IDN AW126, Dec. 1985.
[3] S. M. Dunn, D. Harwood, and L. S. Davis, "Local estimation of the uniform error threshold," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 742–747, 1984.
[4] P. Switzer, "Some spatial statistics for the interpretation of satellite data" (with discussion), *Bull. Int. Statist. Inst.*, vol. 50, pp. 962–972 (book 2), pp. 409–426 (book 3), 1983.
[5] K. V. Mardia, "Spatial discrimination and classification maps," *Commun. Statist. Theor. Meth.*, vol. 13, no. 18, pp. 2181–2197, 1984.
[6] J. Besag, "On the statistical analysis of dirty pictures" (with discussion), *J. Roy. Statist. Soc. B*, vol. 48, no. 3, pp. 259–302, 1986.
[7] K. V. Mardia, J. T. Kent, J. M. Bibby, *Multivariate Analysis*. London: Academic, 1979.
[8] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.
[9] F. R. D. Velasco, "Thresholding using the ISODATA clustering algorithm," *IEEE Trans. Syst., Man., Cybern.*, vol. SMC-10, pp. 771–774, Nov. 1980.
[10] J. Kittler and J. Illingworth, "On threshold selection using clustering criteria," *IEEE Trans. Syst., Man., Cybern.*, vol. SMC-15, pp. 652–655. Sept.–Oct. 1985.
[11] S. Dunn, L. Janos, and A. Rosenfeld, "Bimean clustering," *Pattern Recognition Lett.*, vol. 1, pp. 169–173, Mar. 1983.
[12] J. Kittler and J. Föglein, "Contextual classification of multispectral pixel data," *Image and Vision Comput.*, vol. 2, no. 1, pp. 13–29, Feb. 1984.
[13] P. Switzer, "Extension of linear discriminant analysis for statistical classification of remotely sensed satellite imagery," *J. Int. Ass. Math. Geol.*, vol. 12, pp. 367–376, 1980.
[14] M. David, *Geostatistical Ore Reserve Estimation*. New York: Elsevier, 1977.
[15] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, vols. 1 and 2, 2nd ed. New York: Academic, 1982.
[16] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 721–741, 1984.
[17] N. Otsu, "A threshold selection method from grey-level histograms," *IEEE Trans. Syst., Man., Cybern.*, vol. SMC-9, pp. 62–66, Jan. 1979.

Precision Edge Contrast and Orientation Estimation

EDWARD P. LYVERS AND O. ROBERT MITCHELL

***Abstract*—This correspondence examines the contrast and orientation estimation accuracy of several edge operators which have been**

Manuscript received August 19, 1986; revised May 11, 1988. Recommended for acceptance by W. E. L. Grimson. This work was supported by an RCA Zworykin Scholarship, the National Science Foundation, and the industrial members of the Engineering Research Center on Intelligent Manufacturing Systems.

E. P. Lyvers was with the School of Electrical Engineering, Purdue University, West Lafayette, IN 47907. He is now with Lincoln Laboratories, Lexington, MA.

O. R. Mitchell was with the School of Electrical Engineering, Purdue University, West Lafayette, IN 47907. He is now with the Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX.

IEEE Log Number 8823848.

0162-8828/88/1100-0927$01.00 © 1988 IEEE