Segmentation of remotely sensed images with fuzzy thresholding, and quantitative evaluation

S. K. PAL†, A. GHOSH‡ and B. UMA SHANKAR§

Machine Intelligence Unit, Indian Statistical Institute, 203 B T Road, Calcutta 700035, India

Abstract. Effectiveness of various fuzzy thresholding techniques (based on entropy of fuzzy sets, fuzzy geometrical properties, and fuzzy correlation) is demonstrated on remotely sensed (IRS and SPOT) images. A new quantitative index for image segmentation using the concept of homogeneity within regions is defined. Results are compared with those of probabilistic thresholding, and fuzzy *c*-means and hard *c*-means clustering algorithms, both in terms of index value (quantitatively) and structural details (qualitatively). Fuzzy set theoretic algorithms are seen to be superior to their respective non-fuzzy counterparts. Among all the techniques, fuzzy correlation, followed by fuzzy entropy, performed better for extracting the structures. Fuzzy geometry based thresholding algorithms produced a single stable threshold for a wide range of membership variation.

1. Introduction

Segmentation is a process of partitioning an image space into some non-overlapping meaningful homogeneous regions. The term 'meaningful' is problem dependent. The success of an image analysis system depends on the quality of segmentation. A large variety of methods for image segmentation is available in the literature (Swain and Davis 1978, Rosenfeld and Kak 1982, Pal *et al.* 1983, Deravi and Pal 1983, Kapur *et al.* 1985, Trivedi and Bezdek 1986, Sahoo *et al.* 1988, Pal and Pal 1993, Richards 1993, Gonzalez and Wood 1993). Segmentation of an image can be done either by grey level thresholding or by clustering of pixels into homogeneous classes. In thresholding one tries to get a set of thresholds $\{T_1, T_2, ..., T_k\}$ such that all pixels with grey values in the range $[T_i, T_{i+1}), i = 0, 1, 2, ..., k$; (irrespective of their positions and neighbourhood information) constitute the ith region type. (T_0 and T_{k+1} are taken as minimum and maximum grey values of the image, respectively.) Thresholds may be detected based on histogram information or spatial information. On the other hand, in clustering homogeneous regions are formed by iterative modification using original or derived features (computed over a neighbourhood). Here the pixels with same grey value may be classified into different region types depending on its position and neighbourhood information.

Remotely sensed images are normally poorly illuminated, highly dependent on

e-mail: †sankar@isical.ac.in, ‡ash@isical.ac.in, §uma@isical.ac.in.

the environmental conditions, and have very low spatial resolution. Most of the times a scene contains too many objects (or regions), and these regions are ill-defined because of both greyness and spatial ambiguities. Moreover, the grey value assigned to a pixel is the average reflectance of different types of ground covers present in the corresponding pixel area. Assigning unique class labels with certainty is thus a problem for remotely sensed images. Fuzzy set theory provides a way of handling this problem by associating certainty factors with class labels.

The problem of segmenting remotely sensed images has been addressed by Cannon *et al.* (1986), Laprade (1988), and Sahasrabudhe and Dasgupta (1992). Laprade presented a split-and-merge technique using F-test and a mean predicate to test the uniformity of regions, and applied it to aerial photographs. A two-stage fuzzy *c*-means algorithm was applied on a Landsat-4 image with six bands to demonstrate the feasibility of the methodology for segmentation by Cannon *et al.* (1986). A method of evaluating the suitability of valleys as threshold has been proposed by Sahasrabudhe and Dasgupta (1992), and applied to satellite image segmentation. Attempts are made to find road like structures (Hu *et al.* 1992, Zlotnick and Carnine (Jr.) 1993, Barzohar and Cooper 1993), and man-made object identification from remotely sensed images (Ton 1988, Mandal *et al.* 1994). Neural network based approaches have also been developed for various land cover type classification (Decatur 1989, Benediktsson *et al.* 1990, Lee *et al.* 1990, Baraldi and Parmiggiani 1995).

Note that most of the approaches for segmentation of remote sensing images are based on pixel classification. Grey level thresholding techniques, although computationally less expensive, did not get much attention. One of the reasons may be that remote sensing images are usually multispectral which makes classification-based methods a natural choice. On the other hand, if thresholding-based techniques are used, then the thresholded output on different bands needs to be again integrated to arrive at a decision. One may note that in certain remote sensing applications the image is available only in one band, for example, panchromatic images, for which segmentation based on thresholding seems to be convenient and appropriate.

Furthermore, the methods available for image segmentation, like other processing techniques, are problem dependent. When an image is segmented for visual interpretation, it is ultimately up to the viewers to judge its quality for a specific application. The process of evaluation of image quality therefore becomes a subjective one.

In the present investigation an attempt is made to demonstrate the effectiveness of various fuzzy thresholding and clustering techniques along with quantitative evaluation for segmentation of remotely sensed images. Comparison of these techniques with the respective non-fuzzy techniques is also made both qualitatively and quantitatively. Five different thresholding techniques based on fuzzy and non-fuzzy (probabilistic) entropy, fuzzy geometry and fuzzy correlation, and two clustering (both fuzzy and non-fuzzy) techniques are considered. The quantitative index is based on the concept of homogeneity within a region. Results are demonstrated on both IRS and SPOT satellite images. Some of the algorithms use only global information of input images and the others use local information.

2. Probabilistic entropy based thresholding
In this section we describe entropy of an image, and a few thresholding algorithms.

2.1. *Global entropy of an image*
Based on the concept of Shannon (1948), *entropy of an image* (or its histogram) can be defined as follows. Let $F = [f(p, q)]_{P \times Q}$ be an image of size $P \times Q$, where

$f(p, q)$ is the grey value at (p, q); $f(p, q) \in G_L = \{0, 1, ..., i, ..., L-1\}$, the set of grey levels. Let n_i be the frequency of occurrence of the grey level i ($i \in G_L$). Then $\Sigma_{i=0}^{L-1} n_i = P \times Q = n$ (say). The global entropy of the image is then expressed as

$$H = - \sum_{i=0}^{L-1} p_i \log_2 p_i; \; p_i = \frac{n_i}{n} \tag{1}$$

H is called global, as it depends only on the histogram of the image.

The concept of global entropy of an image can be viewed from a different angle also. Instead of considering one probability distribution for the entire image, let us consider two probability distributions, one for the object and the other for the background. The sum of the individual entropy of the object and the background gives the total entropy of the image.

If S is an assumed threshold (i.e. S is the boundary grey value between object and background), then the probability distribution of the grey levels over the background portion of the image (assuming lower grey values correspond to background) is

$$\frac{p_0}{P_S}, \frac{p_1}{P_S}, ..., \frac{p_S}{P_S} \tag{2}$$

and that of the object portion of the image is

$$\frac{p_{S+1}}{1 - P_S}, \frac{p_{S+2}}{1 - P_S}, ..., \frac{p_{L-1}}{1 - P_S} \tag{3}$$

where $P_S = \Sigma_{i=0}^{S} p_i$.

The entropy of the background portion of the image

$$H_{Bg}(S) = - \sum_{i=0}^{S} \frac{p_i}{P_S} \log_2 \left(\frac{p_i}{P_S} \right) \tag{4}$$

and that of the object portion is

$$H_{Obj}(S) = - \sum_{i=S+1}^{L-1} \frac{p_i}{1 - P_S} \log_2 \frac{p_i}{1 - P_S} \tag{5}$$

The total entropy of the image is

$$H_{Tot}(S) = H_{Obj}(S) + H_{Bg}(S) \tag{6}$$

In order to segregate the object regions from the background, one needs to maximize (Kapur *et al.* 1985) $H_{Tot}(S)$ which results in equiprobable grey levels in each region; and thus maximizes the sum of homogeneities in grey levels within object and background. Therefore, the value of S which maximizes $H_{Tot}(S)$ gives the threshold for object and background classification.

2.2. *Higher order entropy of an image*

In an image, pixel intensities are not independent of each other. This dependency of pixel intensities can be incorporated by considering sequences of pixels for defining image properties. Entropy of order r ($r = 1, 2, 3 ...$) of an image was defined (Pal and Pal 1989) based on the concept of sequence of pixels as follows.

Let $p(S_i)$ be the probability of occurrence of a sequence S_i of grey levels of length

r, where a sequence S_i of length r is defined as a permutation of r grey levels. Let

$$H^{(r)} = -\frac{1}{r}\sum_i p(S_i)\log_2 p(S_i) \tag{7}$$

where the summation is taken over all grey level sequences of length r. $H^{(r)}$ is a monotonic decreasing function of r and $\lim_{r\to\infty} H^{(r)} = H$, the entropy of the image. For different values of r we get various orders of entropy.

If $r = 1$, we get

$$H^{(1)} = -\sum_{i=0}^{L-1} p_i \log_2 p_i \tag{8}$$

where p_i is the probability of occurrence of the grey level i. Note that equation (8) is the same as equation (1) representing the 'global entropy' of the image.

For $r = 2$

$$H^{(2)} = -\frac{1}{2}\sum_{i=0}^{L-1} p_i(S_i)\log_2 p_i(S_i) \tag{9}$$

$$= -\frac{1}{2}\sum_i\sum_j p_{ij}\log_2 p_{ij} \tag{10}$$

where S_i is the sequence of grey level of length two and p_{ij} is the probability of co-occurrence of grey levels i and j. Thus, $H^{(2)}$ (second-order entropy of an image) can be obtained from the co-occurrence matrix (as shown in figure 1) of an image.

$H^{(2)}$ takes into account the spatial distribution of grey levels. Therefore, two images, with identical histograms but having different spatial distributions will have the same $H^{(1)}$ value, but different $H^{(2)}$ values. Expressions for higher order entropies for $(r > 2)$ can also be deduced in a similar manner. $H^{(r)}$, $r \geqslant 2$, is also called the *local entropy* of order r of an image (Pal and Pal 1989).

2.2.1. *Conditional entropy*

Suppose an image has two distinct portions, the object X and the background Y. Suppose the object consists of the grey levels $\{x_i\}$ and the background contains the grey levels $\{y_i\}$. The conditional entropy of the object X given the background Y, i.e. the average amount of information that may be obtained from X, given that one has viewed the background Y, is defined as

$$H\left(\frac{X}{Y}\right) = -\sum_{x_j \in X}\sum_{y_j \in Y} p\left(\frac{x_i}{y_j}\right)\log_2 p\left(\frac{x_i}{y_j}\right) \tag{11}$$
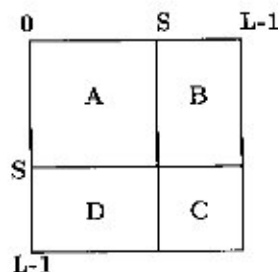

Figure 1.   Four quadrants of the co-occurrence matrix.

Similarly, the conditional entropy of the background Y, given the object X, is defined as

$$H\left(\frac{Y}{X}\right) = - \sum_{y_j \in Y} \sum_{x_i \in X} p\left(\frac{y_j}{x_i}\right) \log_2 p\left(\frac{y_j}{x_i}\right) \tag{12}$$

The pixel having grey level y_j, in general, can be an mth order neighbour of the pixel with grey level x_i. Since the estimation of probability of such an occurrence is very difficult, we assume x_i and y_j to be grey levels of adjacent pixel.

The conditional entropy of an image is then defined as (Pal and Pal 1989)

$$H_{\text{Con}} = \frac{1}{2}\left[H\left(\frac{X}{Y}\right) + H\left(\frac{Y}{X}\right)\right] \tag{13}$$

Computational steps:

Step 1: Compute the probabilities of occurrence of different grey values or sequences (of order r, $r = 1, 2$) of grey values for each of the assumed thresholds (say, S, $S = 1, 2, 3, ..., L - 2$).

Step 2: Using the probability values compute the entropy (e.g. global and local) of the image.

Step 3: Vary S and select those S's for which the entropy values give local maxima. Each of these local maxima corresponds to a threshold. The global optimum corresponds to object background separation.

3. Fuzzy set theory based thresholding

A fuzzy subset A of the universe X is defined as a collection of ordered pairs

$$A = \{(\mu_A(x), x), \forall x \in X\} \tag{14}$$

where $\mu_A(x)$, $(0 \leqslant \mu_A(x) \leqslant 1)$ denotes the degree of belonging of the element x to the fuzzy set A. The support of a fuzzy set A is the crisp set that contains all the elements of X that have a non-zero membership value in A.

Since the theory of fuzzy sets is a generalization of the classical set theory, it has greater flexibility to capture faithfully the various aspects of incompleteness or imperfection in information of a situation (Bezdek and Pal 1992). The flexibility of fuzzy set theory is associated with the elasticity property of the concept of its membership function. The grade of membership is a measure of the compatibility of an object with the imprecise concept represented by a fuzzy set. The higher the value of membership, the lesser will be the amount (or extent) to which the concept represented by a set needs to be stretched to fit an object.

Since the regions in an image are not always crisply defined (imprecision may arise due to both greyness ambiguity and spatial ambiguity), it is natural and appropriate to avoid committing ourselves to a specific hard decision for image segmentation. Thus it is natural to consider the image segments to be fuzzy subsets of the image (first suggested by Prewitt (1970)); the subsets being characterized by the possibility (degree) to which each pixel belongs to them. Moreover (as mentioned in §1), for remotely sensed images the pixel intensities do not reflect the land cover types present in the corresponding area properly. Assigning unique class labels with certainty is thus a problem for remotely sensed images. Hence, fuzzy set theoretic approach will be more appropriate for segmenting regions in remotely sensed images which are usually ill-defined.

3.1. *Fuzzy entropy of an image*

An L level image $F(P \times Q)$ can be considered as an array of fuzzy singletons, each having a membership value denoting its degree of possessing some property (e.g. brightness, darkness, edginess, blurredness, texture, etc.). In the notation of fuzzy sets one may therefore write (Pal and Majumder 1986)

$$F = \{\mu_F(p, q); p = 1, 2, ..., P; q = 1, 2, ..., Q\} \tag{15}$$

where $\mu_F(p, q)$ denotes the grade of possessing such a property μ by the (p, q)th pixel. Membership can be defined based on global information, local information, positional information and a combination thereof, depending on the problem.

Let us construct, say, a fuzzy subset *bright image* characterized by a membership function μ_F using the standard $S(i; a, b, c)$ function of Zadeh (1965) defined as

$$
\begin{aligned}
\mu_F(i) &= 0 & &\text{if } i \leqslant a \\
&= 2\{(i-a)/(c-a)\}^2 & &\text{if } a \leqslant i \leqslant b \\
&= 1 - 2\{(i-c)/(c-a)\}^2 & &\text{if } b \leqslant i \leqslant c \\
&= 1 & &\text{if } c \leqslant i
\end{aligned}
\tag{16}
$$

where $\mu_F(i)$, which is a function only of grey level, represents the degree of belonging of the level i to the fuzzy bright image plane F. The crossover point is $b = (a + c)/2$ (for which the membership value is 0.5) of the membership function μ_F. Similarly, one can use Z-function to construct a *dark image* plane, where

$$Z = (1 - S(i; a, b, c)) \tag{17}$$

Several attempts have been made (Xie and Bedrosian 1988, Murthy and Pal 1992a, b) to determine the appropriate membership function along with its band width and bounds for image processing problems.

Fuzzy entropy of an image (having $n = P \times Q$ pixels) using logarithmic gain function is (Pal 1982)

$$H = \frac{1}{n \ln(2)} \sum_{i=1}^{n} S_n(\mu_F(p, q)) \tag{18}$$

Here,

$$S_n(\mu_F(p, q)) = -\mu_F(p, q) \ln \{\mu_F(p, q)\} - \{1 - \mu_F(p, q)\} \ln \{1 - \mu_F(p, q)\} \tag{19}$$

and $\mu_F(p, q)$ represents the membership for the (p, q)th pixel.

Another definition of fuzzy entropy, given by Pal and Pal (1991) and Ghosh (1995), using exponential gain function is

$$H = \frac{1}{n(\sqrt{e} - 1)} \sum_{i=1}^{n} \{S_n(\mu_F(p, q)) - 1\} \tag{20}$$

with

$$S_n(\mu_F(p, q)) = \mu_F(p, q) e^{1 - \mu_F(p, q)} + \{1 - \mu_F(p, q)\} e^{\mu_F(p, q)} \tag{21}$$

Note that, these entropy measures, first of all, compute the fuzziness related to individual pixel of the image and then make an average over all the pixels to get a quantification of the amount of average ambiguity, the image possesses. Since their computation depends only on the histogram, they may be called global fuzzy entropy.

3.1.1. *Higher order fuzzy entropy*

Pal and Pal (1992) also defined the rth order fuzzy (local) entropy $H^{(r)}$, $r \geqslant 2$ of an image F (using both logarithmic and exponential gain functions) which provides a measure of the average amount of difficulty (ambiguity) in making a decision on any subset of r elements as regards to its possession of an imprecise property. These are as follows:

Out of the n pixels of the image F, consider a combination of r elements. Let S_i^r denote the ith such combination and $\mu(S_i^r)$ denote the degree to which the combination S_i^r, as a whole, possesses the property μ. There are $\binom{n}{r}$ such combinations. The entropy of order r of the image F is defined as (Pal and Pal 1992)

$$H^{(r)} = -\frac{1}{\binom{n}{r}} \sum_{i=1}^{\binom{n}{r}} [\mu(S_i^r) \ln \{\mu(S_i^r)\} + \{1 - \mu(S_i^r)\} \ln \{1 - \mu(S_i^r)\}] \tag{22}$$

with logarithmic gain function, and

$$H^{(r)} = \frac{1}{\binom{n}{r}} \sum_{i=1}^{\binom{n}{r}} [\mu(S_i^r) e_i^{1 - \mu(S_i^r)} + \{1 - \mu(S_i^r)\} e^{\mu(S_i^r)}] \tag{23}$$

with exponential gain function.

$H^{(r)}$ will give a measure of the average amount of difficulty in taking a decision on any subset of size r with respect to the property μ. Note that equations (18)–(21) correspond to a special case of $H^{(r)}$ for $r = 1$. $H^{(r)}$, $r \geqslant 2$ is called higher order fuzzy entropy of the image.

3.1.2. *Membership function and computation of second-order fuzzy entropy*

For computing the higher order fuzzy entropy of an image, represented by a fuzzy set, one needs to choose r pixels at a time and to assign a composite membership value to them. Normally these r pixels are chosen as adjacent pixels. For the present investigation, we have chosen $r = 2$.

Let us consider a two-dimensional S-type membership function (figure 2) representing fuzzy bright image plane (assuming higher grey values correspond to object region). This assigns a composite membership value to a pair of adjacent pixels as follows:

For a particular threshold S,

- (S, S) is the most ambiguous point, i.e. the boundary between object and background. Therefore its membership value for the fuzzy bright image plane is 0.5.
- If one object pixel is followed by another object pixel (i.e. the entries of quadrant C), then its degree of belonging to object region is greater than 0.5. The membership value increases with increase in pixel intensity.
- For quadrants B and D where one object pixel is followed by one background pixel or vice versa, the membership value is less than or equal to 0.5, depending on the deviation from the boundary point (S, S).
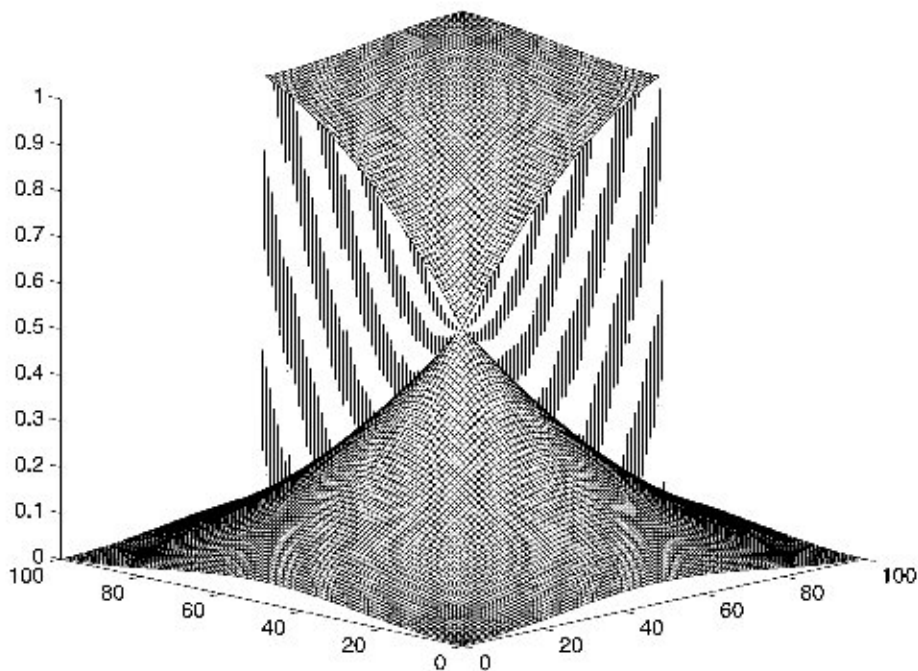- If one background pixel is followed by another background pixel (i.e. for the

Figure 2.   Two-dimensional S-type membership function. (50, 50) is the crossover point.

entries in A), then its degree of belonging to object region is less than 0.5. The membership value decreases with decrease of pixel intensity.

Second-order fuzzy entropy ($r = 2$) computed individually over the entries in quadrants C, A, D and B of the co-occurrence matrix (figure 1) can be termed as object entropy H_O^2, background entropy H_B^2, transitional (object to background) $H_{O/B}^2$ entropy, and transitional (background to object) entropy $H_{B/O}^2$, respectively. Therefore, $H_O^2 + H_B^2$ gives the total within-class second-order local entropy and $H_{(O/B)}^2 + H_{(B/O)}^2$ gives the total second-order transitional (conditional) entropy.

3.2. *Fuzzy geometry of an image subset*

Entropy, as defined in §3.1, may be used in representing greyness ambiguity in an image, i.e. the indefiniteness in making a decision whether an individual pixel is black or white, or a collection of pixels possesses an image property or not. There is another kind of ambiguity in an image called *spatial or geometrical ambiguity* (which refers to indefiniteness in the shape and geometry of regions within the image). These can be represented by fuzzy geometrical properties. Some of them (Pal and Rosenfeld 1988, Pal and Ghosh 1992a) which are used in this investigation are described below.

Fuzzy geometrical properties:

Let μ represent an image fuzzy subset and μ be piecewise constant (for digital image). Then

Area: The area of μ is

$$a(\mu) = \sum \mu \tag{24}$$

Perimeter: The perimeter of μ is

$$p(\mu) = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} |\mu(i) - \mu(j)| \times t_{ij} \tag{25}$$

where $\mu(i)$ and $\mu(j)$ are the membership values of two adjacent pixels having grey value i and j respectively and t_{ij} is the frequency of occurrence of the grey value i followed by j.

Compactness: The compactness of a fuzzy set μ having area $a(\mu)$ and perimeter $p(\mu)$ is defined as

$$\text{comp}(\mu) = \frac{a(\mu)}{p^2(\mu)} \tag{26}$$

Physically, compactness means the fraction of maximum area (that can be encircled by the perimeter) actually occupied by the fuzzy region (concept) represented by μ. Of all possible fuzzy discs, compactness is minimum for its crisp version.

Length: The length of μ is

$$l(\mu) = \max_p \left\{ \sum_q \mu(p, q) \right\} \tag{27}$$

Breadth: Similarly, the breadth of μ is

$$b(\mu) = \max_q \left\{ \sum_p \mu(p, q) \right\} \tag{28}$$

The length (breadth) of an image fuzzy subset gives its longest expansion in the y direction (x direction). If μ is crisp, $\mu(p, q) = 0$ or 1; then length (breadth) is the maximum number of pixels in a column (row).

Index of area coverage (IOAC): The index of area coverage of μ is defined as

$$\text{IOAC}(\mu) = \frac{a(\mu)}{l(\mu) \times b(\mu)} \tag{29}$$

IOAC of a fuzzy image subset represents the fraction (which may be improper also) of the maximum area (that can be covered by the length and breadth of the image) actually occupied by the image. The IOAC is the minimum for the non-fuzzy case of all possible fuzzy versions of a rectangle.

3.3. *Fuzzy correlation*

Let Ω be a closed interval in $\mathbb{R}$. Let $\mu_1: \Omega \rightarrow [0, 1]$ and $\mu_2: \Omega \rightarrow [0, 1]$ be two continuous fuzzy membership functions. The correlation $C(\mu_1, \mu_2)$ between the fuzzy membership functions μ_1 and μ_2 (defined on the same domain) was defined by Murthy *et al.* (1985), and Pal and Ghosh (1992b). $C(\mu_1, \mu_2)$ basically gives a measure of relation between the natures of μ_1 and μ_2, i.e. with change of x what happens to μ_1 and μ_2.

Now if the functions are discrete in nature (as applicable to a digital image), the

expression for correlation takes the form (Pal and Ghosh 1992b),

$$C(\mu_1, \mu_2) = \begin{cases} 1 - \dfrac{4}{X_1 + X_2} \sum_x \{\mu_1(x) - \mu_2(x)\}^2 & \text{if } X_1 + X_2 \neq 0 \\ 1 & \text{if } X_1 + X_2 = 0 \end{cases} \qquad (30)$$

where $X_1 = \sum_x \{2\mu_1(x) - 1\}^2$ and $X_2 = \sum_x \{2\mu_2(x) - 1\}^2$.

3.3.1. Correlation between two fuzzy representations (properties) of an image

Fuzzy correlation between two representations of an image characterized by μ_1 and μ_2 is defined as (Pal and Ghosh 1992b)

$$C(\mu_1, \mu_2) = 1 - \frac{4}{X_1 + X_2} \sum_p \sum_q \{\mu_1(p, q) - \mu_2(p, q)\}^2 \qquad (31)$$

with $X_1 = \sum_p \sum_q \{2\mu_1(p, q) - 1\}^2$ and $X_2 = \sum_p \sum_q \{2\mu_2(p, q) - 1\}^2$ where $\mu_1(p, q)$ and $\mu_2(p, q)$ denote the degree of possessing the property μ_1 and μ_2, respectively, by the (p, q)th pixel.

Let μ_2 be the nearest two-tone version of μ_1 such that

$$\mu_2(x) = \begin{cases} 0 & \text{if } 0 \leqslant \mu_1(x) \leqslant 0.5 \\ 1 & \text{otherwise} \end{cases} \qquad (32)$$

Let μ_1 (or F_r) denote a fuzzy *bright image* plane of F having the crossover point at S, say, and be dependent only on grey level. Then μ_2 (or $\bar{F}$) represents its closest two-tone version thresholded at S. Then the fuzzy correlation between a fuzzy representation of an image and its nearest two-tone version is expressed as

$$C(\mu_1, \mu_2) = 1 - \frac{4}{X_1 + X_2} \left(\sum_{i=0}^{S} \{[\mu_1(i)]^2 h(i)\} + \sum_{i=s+1}^{L-1} \{[1 - \mu_1(i)]^2 h(i)\} \right) \qquad (33)$$

with

$$X_1 = \sum_{i=0}^{L-1} [2\mu_1(i) - 1]^2 h(i)$$

and

$$X_2 = \sum_{i=0}^{L-1} [2\mu_2(i) - 1]^2 h(i) = \sum_{i=0}^{L-1} h(i) = P \times Q = \text{constant}$$

Here, $h(i)$ is the frequency of the ith grey level.

3.3.2. Correlation measure using local information

Higher order correlation using local information (obtained from co-occurrence matrix) can also be defined (Pal and Ghosh 1992b) similarly as in higher order entropy of an image (§3.1.2).

Correlation between any two properties of F computed over the entries in individual quadrants A, C, B and D of the co-occurrence matrix (figure 1) can be termed background correlation $[C(\mu_1, \mu_2)_B]$, object correlation $[C(\mu_1, \mu_2)_O]$, transitional (background to object) correlation $[C(\mu_1, \mu_2)_{B/O}]$, and transitional (object to background) $[C(\mu_1, \mu_2)_{O/B}]$ correlation, respectively. They may be computed by using

similar expressions with different ranges of i and j. For example,

$$C(\mu_1, \mu_2)_B = 1 - \frac{4}{X_1 + X_2} \sum_{i=0}^{S} \sum_{j=0}^{S} \{[\mu_1(i,j) - \mu_2(i,j)]^2 t_{ij}\} \tag{34}$$

with

$$X_1 = \sum_{i=0}^{S} \sum_{j=0}^{S} \{[2\mu_1(i,j) - 1]^2 t_{ij}\}$$

and

$$X_2 = \sum_{i=0}^{S} \sum_{j=0}^{S} \{[2\mu_2(i,j) - 1]^2 t_{ij}\}$$

where t_{ij} is the frequency of occurrences of the grey level i followed by j.

Similarly, for computing $C(\mu_1, \mu_2)_O$, i and j will range from $S+1$ to $L-1$; for $C(\mu_1, \mu_2)_{B/O}$, i will range from 0 to S and j from $S+1$ to $L-1$; and for $C(\mu_1, \mu_2)_{O/B}$, i will range from $S+1$ to $L-1$ and j from 0 to S.

Note that $C(\mu_1, \mu_2)_O + C(\mu_1, \mu_2)_B$ gives the total within-class local correlation and $C(\mu_1, \mu_2)_{O/B} + C(\mu_1, \mu_2)_{B/O}$ gives the total transitional (conditional) correlation.

From the properties of correlation we notice that if two functions μ_1 and μ_2 are very close then $C(\mu_1, \mu_2)$ is very high whereas, $C(\mu_1, \mu_2)$ is least when $\mu_2 = 1 - \mu_1$. Since F is the nearest two-tone version of F_f, $C(F_f, F)$ gives a measure of closeness of the two images F_f and $\bar{F}$. The principle of maximizing fuzzy correlation for image segmentation is described in Pal and Ghosh (1992b).

Computational steps of fuzzy thresholding:

Given an L level image F of dimension $P \times Q$ with minimum and maximum grey values i_{min} and i_{max}, respectively, the algorithm for its fuzzy segmentation (through thresholding) into object and background may be described as follows:

Step 1: Construct the membership plane using the standard $S(i; a, b, c)$ function (equations (16)) as

$$\mu(p, q) = \mu(i) = S(i; a, b, c) \tag{35}$$

(called bright image plane if the object regions possess higher grey values)

$$\text{or } \mu(p, q) = \mu(i) = 1 - S(i; a, b, c) \tag{36}$$

(called dark image plane if the object regions possess lower grey values) with crossover point b, and a band width $\Delta b = b - a = c - b$.

Step 2: Compute the parameter $I(F)$ representing either greyness ambiguity or spatial ambiguity (as designated by $H^{(r)}$, correlation, compactness and IOAC, say) or both (i.e. product of greyness and spatial ambiguities).

Step 3: Vary b between i_{min} and i_{max} and select those b for which $I(F)$ has local minima or maxima depending on $I(F)$. (Maxima correspond for the correlation measure only.) Among the local minima/maxima, let the global one have crossover point at S.

The level S, therefore, denotes the crossover point of the fuzzy image plane μ_F, which has minimum greyness and/or geometrical ambiguity. The μ_F plane then can be viewed as a fuzzy segmented version of the image F. For the purpose of non-fuzzy segmentation, we can take S as the threshold (or boundary) for classifying or segmenting an image into object and background. In case the image has multiple

regions, there will be a set of local optima corresponding to them. Faster methods of computation of the fuzzy parameters have been described by Pal and Ghosh (1992a).

Note that $w = 2\Delta b$ is the length of the window (such that $[0, w] \rightarrow [0, 1]$) which was shifted over the entire dynamic range. As w decreases, the μ_F plane tends to have more intensified contrast around the crossover point, thus resulting in a decrease of ambiguity in F. As a result, the possibility of detecting some undesirable thresholds (spurious minima) increases because of the smaller value of w. On the other hand, an increase in w results in a higher value of fuzziness and thus leads towards the possibility of losing some of the weak minima.

The criteria regarding the selection of membership function and the length of window (i.e. w) have been reported in Murthy and Pal (1992a), assuming continuous functions for both histogram and membership function.

4. Segmentation by clustering

Clustering is a method of partitioning a given set of patterns into a number of homogeneous groups (clusters) depending on the similarity in features. The number of groups is normally pre-specified; but can also be varying. Initial clusters are normally chosen randomly and gradually modified to obtain the final clusters (or optimal cluster centres). A number of clustering algorithms (both non-fuzzy and fuzzy) exists in the literature (Anderberg 1973, Bezdek 1981). In the present investigation we used the hard and fuzzy c-means (HCM and FCM) techniques. Here we briefly describe them.

4.1. *Hard c-means algorithm*

Let $X = \{x_1, x_2, ..., x_k, ..., x_n\}$, $x_k \in \mathbb{R}^p$, be a finite dataset in the p-dimensional space; x_{kj} is the jth feature of the data point x_k. Let c $(2 \leqslant c < n)$ be the number of clusters and $V = (V_1, V_2, ..., V_i, ..., V_c)$, $V_i \in \mathbb{R}^p$ be the set of cluster centres (prototypes). The dataset is then classified and cluster centres are updated iteratively until the classification in two successive stages remain unaltered; which can be measured by the average difference between the partitions (prototypes) computed in two successive stages. If this average difference is less than a pre-defined small positive value ε (> 0) then the process can be terminated and the clusters can be taken as optimal. The classification strategy is as follows:

A data point k is assigned to class i if

$$d_{ik} = \min_{1 \leqslant s \leqslant c} (d_{sk}) \tag{37}$$

where $d_{ik}^2 = \sum_{j=1}^{p} (x_{kj} - v_{ij})^2$;

and $v_{ij} = \dfrac{1}{n_i} \sum_{k=1}^{n_i} (x_{kj})$, n_i is the number of data points assigned to the ith class.

4.2. *Fuzzy c-means algorithm*

Let X, c and V be defined as above and $U = [u_{ik}]_{c \times n}$ be a fuzzy c-partition of X. The membership u_{ik} represents the degree of belonging of the pattern x_k to the ith

class, where

$$0 \leqslant u_{ik} \leqslant 1, \ \sum_{i=1}^{c} u_{ik} = 1, \ \forall k = 1, 2, \ldots, n \ \text{and}$$

$$0 < \sum_{k=1}^{n} u_{ik} < n, \ \forall i = 1, 2, \ldots, c \tag{38}$$

FCM finds U and V iteratively by minimizing

$$J_m(U, V) = \sum_{i=1}^{c} \sum_{k=1}^{n} (u_{ik})^m \| x_k - V_i \|_A^2 \tag{39}$$

where $m > 1$ and the inner product induced norm metric

$$\| \cdot \|_A^2 = (x_k - V_i)^T A (x_k - V_i) = d_{ik}^2 \tag{40}$$

where A is any $p \times p$ positive definite matrix (in the present study we used the Euclidean norm, i.e. A is an identity matrix). Based on the necessary conditions for a local minimum of J_m, cluster prototypes and memberships are computed as follows:

$$V_i = \frac{\sum_{k=1}^{n} (u_{ik})^m x_k}{\sum_{k=1}^{n} (u_{ik})^m}, \quad 1 \leqslant i \leqslant c \tag{41}$$

Let, $I_k = \{ i \mid 1 \leqslant i \leqslant c, d_{ik} = 0 \}$ and

$$\bar{I}_k = \{ 1, 2, \ldots, c \} - I_k.$$

If $I_k \neq \Phi$, $u_{ik} = 0 \ \forall i \in \bar{I}_k$ and arbitrarily assign u_{ik} such that $\sum_{i \in I_i} u_{ik} = 1$.

$$\text{If } I_k = \Phi, \ u_{ik} = \left(\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}} \right)^{2/(m-1)} \right)^{-1}, 1 \leqslant i \leqslant c, 1 \leqslant k \leqslant n \tag{42}$$

To implement the FCM algorithm we initialize either U or V, and then FCM iterates between (41) and (42) until U or V or both stabilize.

4.2.1. *Computational steps*

Using the above-mentioned clustering techniques image pixels are labelled as one of the region types in the image; thereby providing its segmentation. We used a two-dimensional feature space with feature values as average and busyness (Rosenfeld and Kak 1982) of the concerned image pixel. The computational steps are described below.

Step 1: Choose the number of classes and the initial values for the means (for FCM an initial partition can also be chosen instead of means of the classes).

Step 2: Classify the samples by assigning them to the class corresponding to the closest mean (or assign membership value for each class in the case of FCM).

Step 3: Re-compute the means of the classes (weighted means for FCM).

Step 4: If the change in any of the means is more than some pre-assigned small positive quantity (say, $\varepsilon > 0$) then STOP else go to *Step 2*.

Here, we briefly discuss the two features, average and busyness (Rosenfeld and Kak 1982) used with FCM and HCM. Let us consider a 3×3 window centred at (i, j), with grey levels as indicated in figure 3.

$$\begin{array}{|c|c|c|}\hline a_1 & a_2 & a_3 \\ \hline a_4 & a_5 & a_6 \\ \hline a_7 & a_8 & a_9 \\ \hline \end{array}$$

Figure 3.   Grey values over a 3×3 window.

The *average* grey level $\bar{a}$ over the window centred at the (i,j)th position (grey value a_5) of the image is

$$\bar{a} = \frac{1}{9}\left(\sum_{k=1}^{9} a_k\right) \qquad (43)$$

The *busyness* B over the window in figure 3 is

$$A_1 = |a_1 - a_2| + |a_2 - a_3| + |a_4 - a_5| + |a_5 - a_6| + |a_7 - a_8| + |a_8 - a_9|,$$
$$A_2 = |a_1 - a_4| + |a_4 - a_7| + |a_2 - a_5| + |a_5 - a_8| + |a_3 - a_6| + |a_6 - a_9|, \text{ and}$$

$$\text{busyness} = B = \frac{A_1 + A_2}{12} \qquad (44)$$

5.   Quantitative measure

We have used several algorithms with different parameters in this investigation. Each of these gives rise to a partition (segmentation) of the image space. We intend to evaluate this segmentation results quantitatively using an index (say β).

Let n_i be the number of pixels in the ith $(i = 1, 2, ..., c)$ region obtained by a segmentation method. Let X_{ij} be the grey value of jth pixel $(j = 1, ..., n_i)$ in region i, and X_i the mean of n_i grey values of ith region. Then β is defined as

$$\beta = \frac{\frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n_i}(X_{ij} - \bar{X})^2}{\sum_{i=1}^{c}\frac{n_i}{n} \times \frac{1}{n_i}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2} = \frac{\sum_{i=1}^{c}\sum_{j=1}^{n_i}(X_{ij} - \bar{X})^2}{\sum_{i=1}^{c}\sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2} \qquad (45)$$

where, n is the size of the image and X is the mean grey value of the image.

Note that the above measure is nothing but the ratio of the total variation and within-class variation. This is widely used in feature selection and cluster analysis (Fisher and Van Ness 1971, Richards 1993). Since the numerator is constant for an image, the β value is dependent only on the denominator. The denominator decreases with increase in homogeneity in the regions. Therefore, for a given image and c value, the higher the homogeneity within the segmented regions, the higher would be the β value. The value of β also increases with c. In an extreme case when there is no partition (i.e. the entire image space is being considered as one class), we have

$c = 1$ and $\beta = 1$. On the other extreme, when the number of partitions is maximum, i.e. $c = L$, the number of grey levels in the image, we have $\beta = \infty$.

6. Implementation and results

The algorithms, mentioned in the previous sections, with the quantitative measure have been implemented on a number of remotely sensed images. We present here results on two IRS-1A (band 4) images and one SPOT (band 3) image. The IRS-1A image was taken using the scanner LISS-II (Linear Imaging Self Scanner) in the wavelength range 0.77–0.86 μm and it has a spatial resolution of 36.25 m × 36.25 m (Thiruvengadachari *et al.* 1989). One of the images is covering an area around the city of Calcutta (figure 4), whereas the other one is covering the city of Bombay (figure 5). The SPOT images have a spatial resolution of 20.0 m × 20.0 m and the wavelength ranges in 0.79–0.89 μm (Richards 1993), covering a portion of the city of Calcutta (figure 6). All the three images are of size 512 × 512 pixels. Due to poor illumination of these images, the actual object classes present in the input image are not visible clearly. For this reason we have not included the original input images, instead an enhanced version (histogram equalized) of the input images highlighting the different object regions are shown in figures 4, 5 and 6, corresponding to Calcutta (IRS), Bombay (IRS) and Calcutta (SPOT), for the convenience of readers. However, the algorithms were implemented on the actual input images whose histograms are shown in figures 7, 8 and 9, corresponding to these three images. As seen from the

Figure 4. Enhanced (Histogram equalized) Calcutta (IRS) image.

Figure 5. Enhanced (Histogram equalized) Bombay (IRS) image.

histogram (figure 7), a deep valley is present near the grey level 23, closely surrounded by two significant peaks at grey levels 21 and 33. Some weak valleys are also present at levels 55, 62, 65, 70, 72 and 76, but they are not apparent in figure 7. Corresponding to the Bombay (IRS) image of figure 5, the histogram presented in figure 8, depicts a valley at grey level 18, surrounded by two peaks at levels 11 and 25. The other valleys at 57, 59, 64, 66 and 70 are not significant. For the Calcutta (SPOT) image in figure 6 the histogram is shown in figure 9. There are three main peaks (two strong at 18 and 29, and one weak at 22) with valleys, at 20 and 23. The other valleys, which are not visible, are at levels 66, 70, 72, 76, 78, 80, 88, 90, 92, 95, 98 and 103.

For implementing fuzzy c-means (FCM) and hard c-means (HCM) clustering algorithms we have chosen the number of clusters $c = 4$, 5 and 6. Average (equation (43)) and busyness (equation (44)), computed over a 3×3 neighbourhood (Rosenfeld and Kak 1982) incorporating local information, are used as features. For the FCM fuzzifier m (in equation (39)) was taken as 2.

In tables 1–3, we present the thresholds obtained by fuzzy correlation, fuzzy entropy and fuzzy geometry, respectively, for different window sizes ($w = 7$, 9, 11, 13, 15, 17, 19). The results corresponding to probabilistic entropy are shown in table 4. The computed β values are also shown in the tables. As expected, the number of thresholds (where † indicates global optima) decreases as w increases. Note that
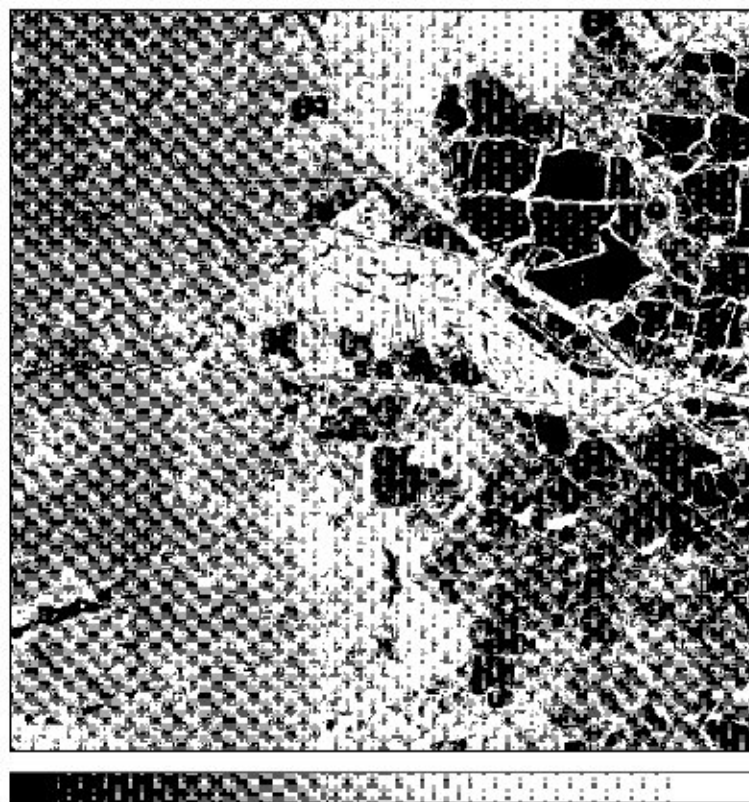
Figure 6.   Enhanced (Histogram equalized) Calcutta (SPOT) image.

membership functions corresponding to $w = 9$ and 11 satisfy the criteria of bounds[1] of Murthy and Pal (1992a).

To demonstrate the significance of the β value we provide segmented results corresponding to the highest and lowest β values, obtained over all these methods, for a fixed number of regions. Figures 10 and 11 depict such example images for $\beta = 9.949$ and $\beta = 4.357$, respectively, when the number of regions, c, is five. For the purpose of comparison we also show the results of FCM (figure 12 with $\beta = 5.880$) and HCM (figure 13 with $\beta = 5.171$) for $c = 5$. From these, the β value is seen to reflect well the quality of segmentation. For example, the details of linear structures like roads, airport runway, water canals are seen to be more prominent in figure 10 (having higher β value) compared with others.

Let us now consider the segmented output (figure 14) corresponding to the best β value ($= 4.790$) when the number of regions, c, is four. Comparing the results of figure 10 (when $c = 5$) we see that some of the roads, canals (linear structures) which are visible in figure 10 are vanished in figure 14. Also it can be seen from figure 14 that the main city area consisting of dense concrete structure is merged with the sparse concrete structure class (habitation class) around the river.

---

[1] Bounds determine the range so that any membership function defined within the range will provide similar segmented results.
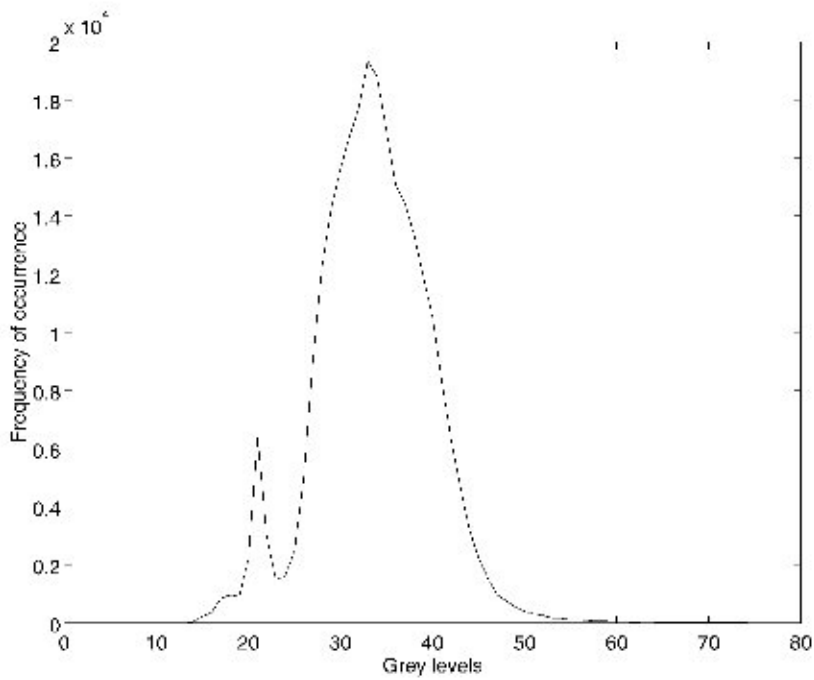
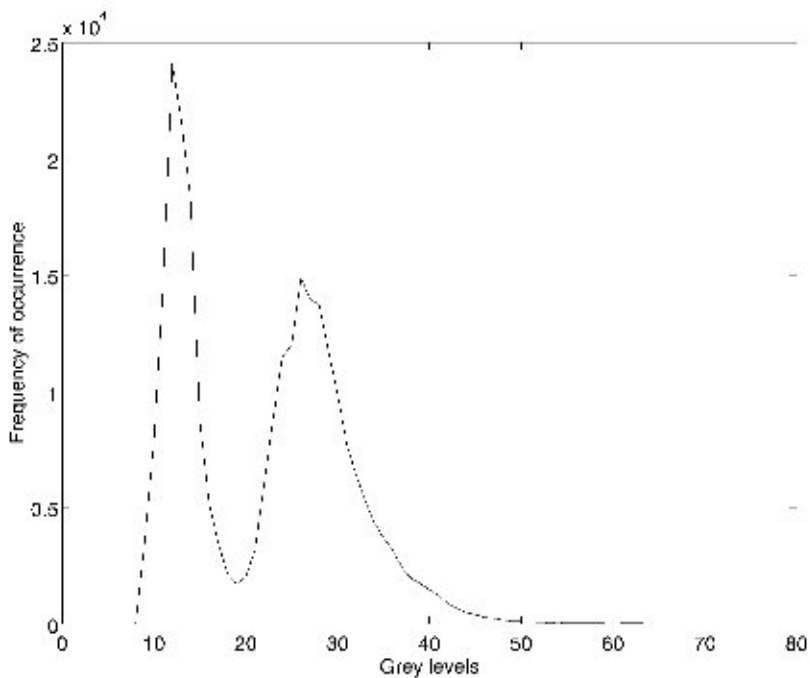Figure 7. Histogram of the original Calcutta (IRS) image.

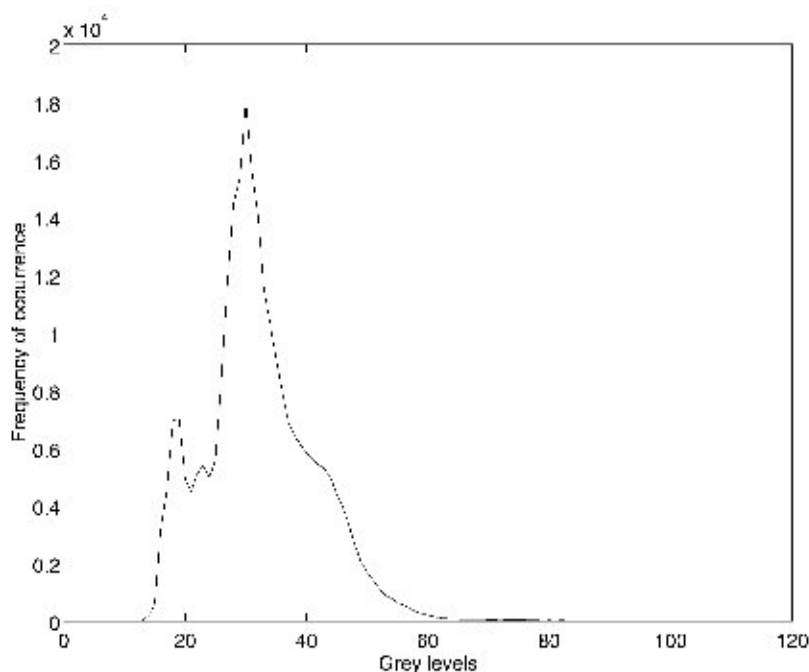Figure 8. Histogram of the original Bombay (IRS) image.

Figure 9. Histogram of the original Calcutta (SPOT) image.

Table 1. Thresholds based on fuzzy correlation and β values.

Serial no.	Kind of information	Window size (w)	Threshold values	β value
1.1	global	7	19†, 23, 31, 36, 39, 47	14.286
1.2		9	24†, 31, 36, 47	9.634
1.3		11	24†, 31, 37, 47	9.949
1.4		13	24†, 32, 38, 48	9.612
1.5		15	24†, 48	1.711
1.6		17	24, 49†	1.681
1.7		19	24†	1.466
1.8	local	7	19, 24†, 32, 36, 47	9.533
1.9	(within-class)	9	19, 24†, 32, 37, 48	9.989
1.10		11	25†, 37, 48	4.486
1.11		13	25†, 37	3.762
1.12		15	25†, 36	4.011
1.13		17	25†, 36	4.011
1.14		19	25†, 36	4.011
1.15	local	7	19, 24, 36, 40, 48†	5.983
1.16	(transitional)	9	19, 24, 36, 48†	4.894
1.17		11	24, 36, 48†	4.790
1.18		13	24, 37, 48†	4.358
1.19		15	24†, 38	3.311
1.20		17	24†, 38	3.311
1.21		19	38†	1.900

†The global maximum.

Table 2. Thresholds based on fuzzy entropy and β values.

Serial no.	Kind of information	Form of gain information	Window size (w)	Threshold values	β value
2.1	global	logarithmic	7	19, 23, 31†, 36, 47	9.793
2.2			9	24, 31†, 37, 47	9.949
2.3			11	24, 31†, 37, 47	9.949
2.4			13	24, 38†, 48	3.788
2.5			15	23†, 49	1.639
2.6			17	23†	1.432
2.7			19	—	—
2.8	global	exponential	7	19, 23, 31†, 36, 47	9.793
2.9			9	24, 31†, 37, 47	9.949
2.10			11	24, 31†, 37, 47	9.949
2.11			13	24, 38†, 48	3.788
2.12			15	24†, 49	1.681
2.13			17	23†	1.432
2.14			19	—	—
2.15	local (within-class)	logarithmic	7	18, 24, 31†, 36, 49	9.473
2.16			9	25, 37†, 48	4.486
2.17			11	25, 37†, 47	4.531
2.18			13	26, 36†	4.146
2.19			15	26, 35†	4.202
2.20			17	26, 36†	4.146
2.21			19	26, 36†	4.146
2.22	local (within-class)	exponential	7	18, 24, 31†, 36, 49	9.473
2.23			9	25, 37†, 48	4.486
2.24			11	25, 37†, 47	4.531
2.25			13	25, 36†	4.011
2.26			15	26, 36†	4.146
2.27			17	25, 36†	4.011
2.28			19	26, 36†	4.146
2.29	local (transitional)	logarithmic	7	19†, 24, 37, 40, 49	4.933
2.30			9	19†, 24, 37, 49	4.357
2.31			11	24, 37†, 49	4.274
2.32			13	24, 38†, 48	3.788
2.33			15	24, 31†	2.981
2.34			17	23, 31†	2.963
2.35			19	30†	2.173
2.36	local (transitional)	exponential	7	19†, 24, 37, 40, 48	4.945
2.37			9	19†, 24, 37, 49	4.357
2.38			11	24, 37†, 47	4.358
2.39			13	24, 38†, 49	3.764
2.40			15	24†	1.466
2.41			17	23, 31†	2.963
2.42			19	31†	2.319

†The global minimum.

Figures 15 and 16 show a comparison of object–background partition (i.e. $c = 2$) as obtained by the thresholding with highest β (= 2.422) value, and HCM/FCM ($\beta = 2.198$). Both visually and by β values the segmented version in figure 15 is seen to be superior to that of figure 16.

Table 3. Thresholds based on fuzzy geometry and β values.

Serial no.	Optimizing property	Window size (w)	Threshold values	β value
3.1	compactness	7	32†	2.422
3.2		9	32†	2.422
3.3		11	32†	2.422
3.4		13	32†	2.422
3.5		15	32†	2.422
3.6		17	32†	2.422
3.7		19	31†	2.319
3.8	IOAC	7	31†	2.319
3.9		9	30†	2.173
3.10		11	30†	2.173
3.11		13	31†	2.319
3.12		15	31†	2.319
3.13		17	31†	2.319
3.14		19	31†	2.319

†The global minimum.

Table 4. Thresholds based on probabilistic entropy and β values.

Serial no.	Kind of information	Form of grain function	Threshold values	β value
4.1	global	logarithmic	20, 26†	1.603
4.2	global	exponential	19, 26, 34†	4.208
4.3	local (within-class)	logarithmic	20†	1.149
4.4	local (transitional)	logarithmic	19, 24†	1.475
4.5	local (within-class)	exponential	19, 35†	2.712
4.6	local (transitional)	exponential	19, 23†, 37	3.579

†The global maximum.

Figure 17 demonstrates the variation of maximum value of β, obtained over different methods, for $c = 2, 3, 4, 5, 6, 7$. As expected (mentioned in §5), the β value increases with increase in c.

Let us now consider figure 15, which has highest β value among those with $c = 2$, for object–background classification. Here the threshold is seen to be 32, distinguishing object region containing water body and city area from the background containing habitation, vegetation and open spaces. Interestingly, it is seen from the correlation-based segmentation results (table 1) that the segmented images with a threshold at or around 32 usually posses high β values. This indicates the significance of the said threshold.

From tables 1 and 2 it is seen that the thresholds corresponding to global correlation and global entropy measures for $w = 9$ and 11 are almost identical. (Note that thresholds 24, 31, 37 and 47 correspond to the boundary between water body, city area, habitation, vegetation and open spaces.) Similar is the case with transitional
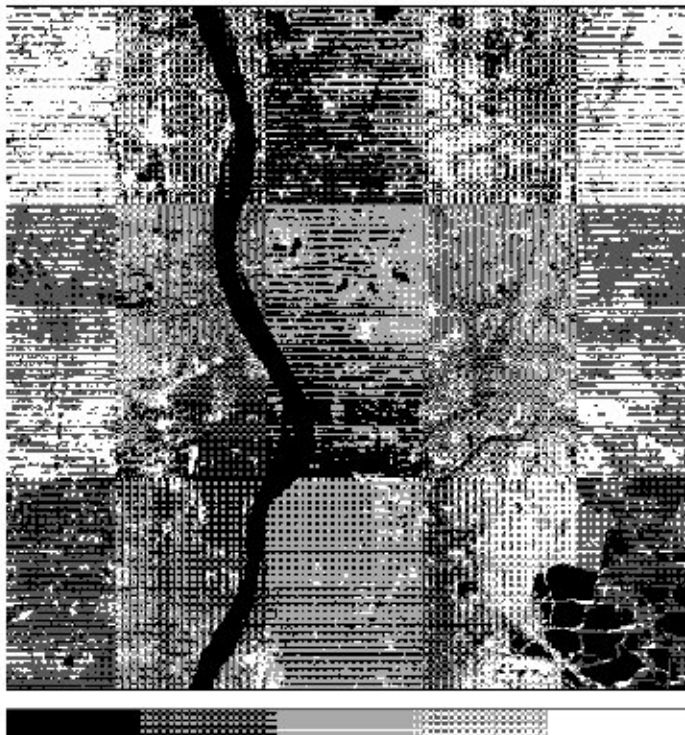
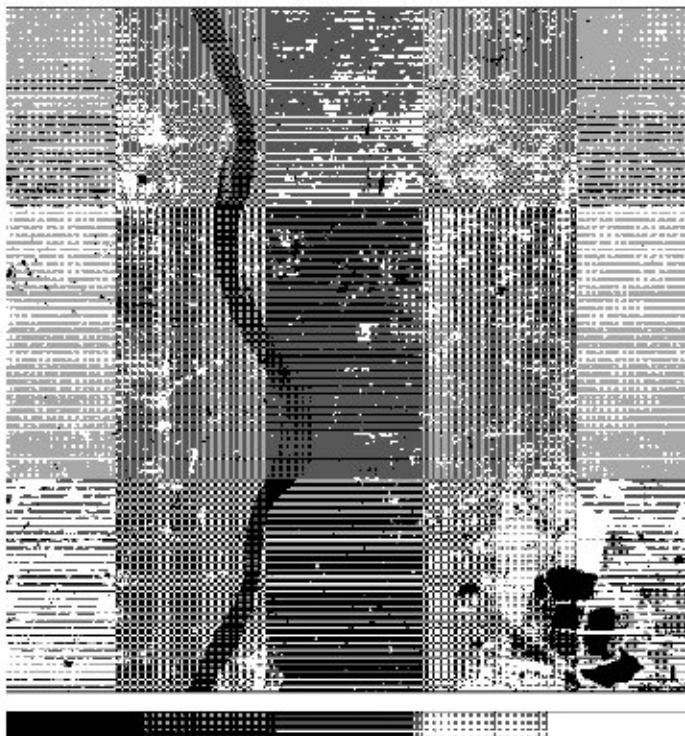Figure 10. Segmented Calcutta (IRS) image with highest β (= 9.949) value and c = 5.

Figure 11. Segmented Calcutta (IRS) image with lowest β (= 4.357) value and c = 5.

Figure 12.   Segmented Calcutta (IRS) image using FCM when $c = 5$, $\beta = 5.880$.

Figure 13.   Segmented Calcutta (IRS) image using HCM when $c = 5$, $\beta = 5.171$.
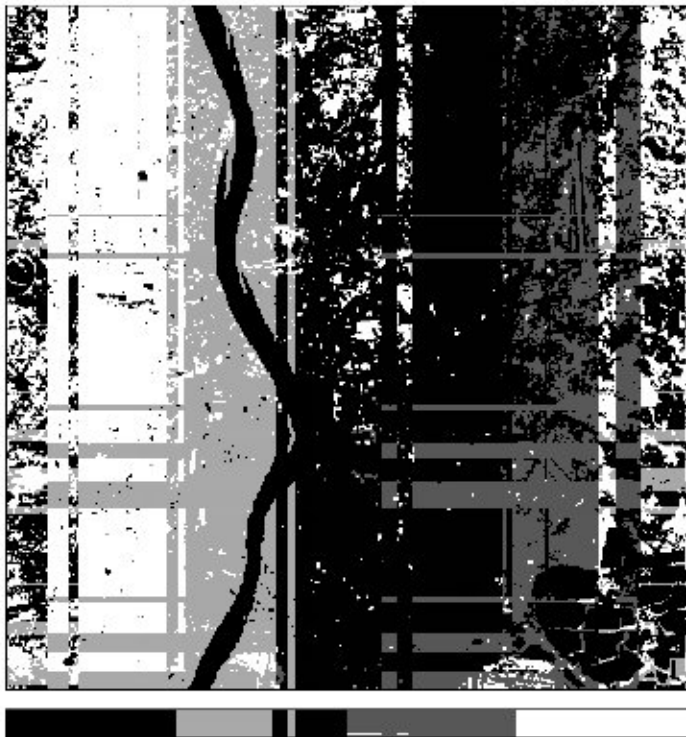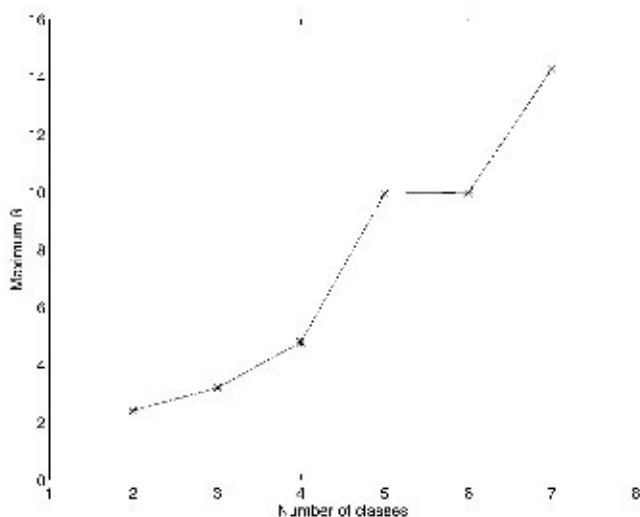
Figure 14. Segmented Calcutta (IRS) image with highest β (= 4.790) value and c = 4.

Figure 15. Segmented Calcutta (IRS) image with highest β (= 2.422) value and c = 2.

Figure 16. Segmented Calcutta (IRS) image using FCM/HCM when $c = 2$, $\beta = 2.198$.

Figure 17. Variation of maximum value of β with the number of classes.

correlation (table 1) and transitional entropy measures (table 2). Further, the thresholds obtained by two gain functions are seen to be almost similar for both global and local entropy.

In the case of IOAC and Compactness (table 3), interestingly, only one threshold in the range 30–32 is obtained irrespective of the window size. Although the method

produced only one threshold, its importance is evident from figure 15 where it is seen to clearly demarcate the silhouettes of the objects present in the input image. This may be due to the incorporation of spatial ambiguities in the optimizing function. Note that this important threshold is missing in all the cases in table 4.

From the previous discussion we therefore see that β values provide a good quantitative index for measuring homogeneity in segmented regions. For a fixed c, its values increases with the quality of segmentation. Fuzzy set theoretic approaches are better than the probabilistic entropic methods. This may be due to the fact that fuzzy approaches described here exploit the ambiguities (both in greyness and in spatial domain) of the image in an effective way. Among the fuzzy techniques, fuzzy geometry based optimization (which basically optimizes the spatial ambiguities) is seen to provide a single threshold, over a wide variation of window size; which is able to segregate the basic structures in the image well. Surprisingly, the global information based fuzzy correlation and fuzzy entropy measures provided better performance (higher β value) for extracting the structures. Among the local information based techniques, within-class fuzzy correlation based algorithms, showed an upper edge. From all the results obtained by thresholding algorithms, it therefore appears that fuzzy correlation and fuzzy entropy using global information (with $w = 9$ and 11) are the best optimizing criteria from the point of possessing β value and detecting the structural details (figure 10).

Let us now consider the Bombay IRS image (figure 5). Figures 18 and 19 depict the segmented images for highest ($\beta = 18.308$) and lowest ($\beta = 11.243$) β values for $c = 5$. For the purpose of comparison we also consider the results of FCM (figure 20
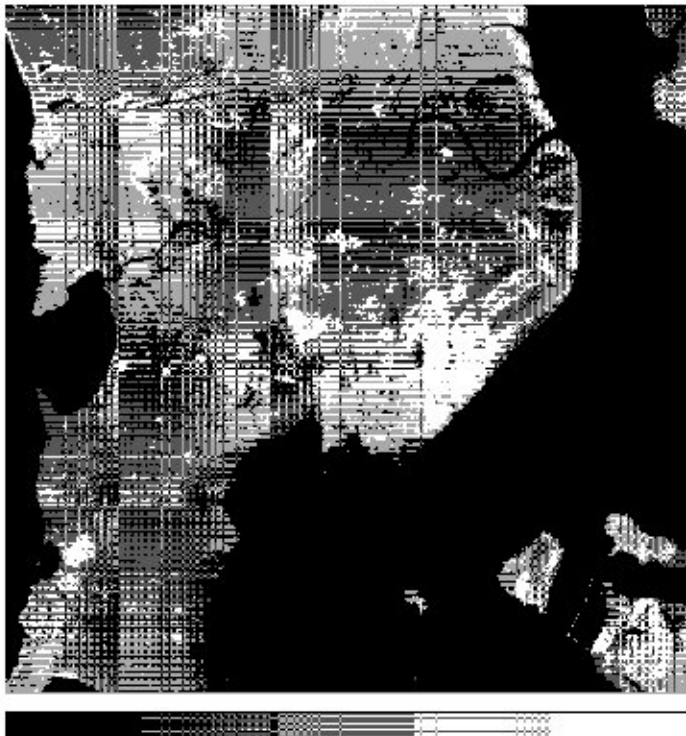

Figure 18.   Segmented Bombay (IRS) image with highest β (= 18.308) value and $c = 5$.

Figure 19. Segmented Bombay (IRS) image with lowest β ($= 11.243$) value and $c = 5$.

Figure 20. Segmented Bombay (IRS) image using FCM when $c = 5$, $\beta = 15.815$.

with $\beta = 15.815$) and HCM (figure 21 with $\beta = 16.056$) for $c = 5$. Like the Calcutta IRS image, the β value reflects the quality of segmentation. The details of linear structures (railroads, airport runway, etc.) are seen to be more prominent in figure 18 (with highest β value) than others.

As in the case of the Calcutta IRS image, here also IOAC and Compactness resulted in a single threshold for all the cases irrespective of w. However, the threshold obtained by IOAC is different from that of Compactness. Among the different segmented regions for $c = 2$, IOAC has highest β value. Here the threshold at 19 is seen to discriminate the water body well from the land.

Finally, we consider the Calcutta SPOT image (figure 6). Figures 22 and 23 depict the segmented version for highest ($\beta = 9.375$) and lowest ($\beta = 6.467$) values of β obtained over all the thresholding methods for $c = 5$. Comparing them with FCM (figure 24, $\beta = 6.388$) and HCM (figure 25, $\beta = 6.676$) we see that the details of structure in figure 22 are more prominent. This also strengthens the fact that β values provide a good quantitative measure of image segmentation.

7. Conclusions and discussion

Fuzzy thresholding provides a useful segmentation technique for remote sensing images. β provides a good quantitative index for image segmentation. For a given c, the higher the value of β, the better is the homogeneity within the segmentated regions. The maximum value of β for a given c, increases with c.

Among the various thresholding techniques, fuzzy correlation provided the best performance, followed by fuzzy entropy as far as β value and the detection of various
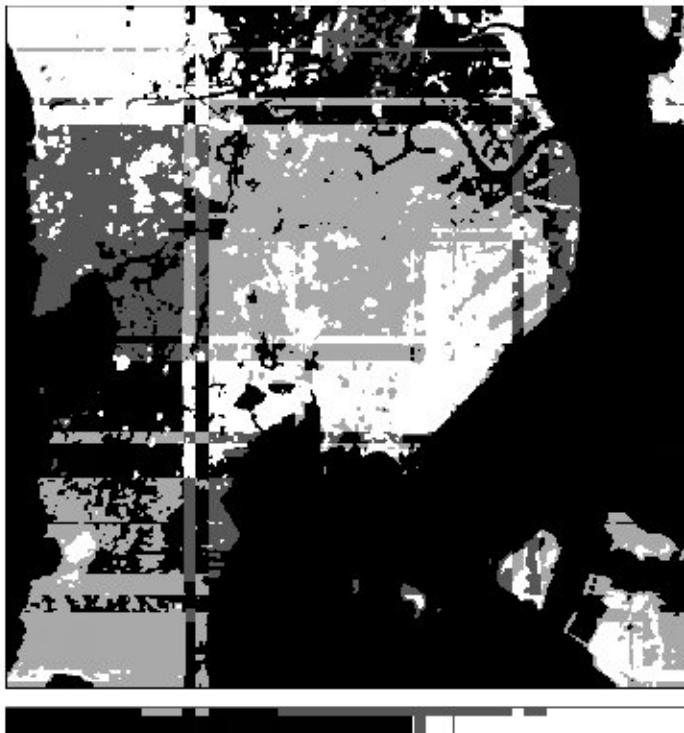

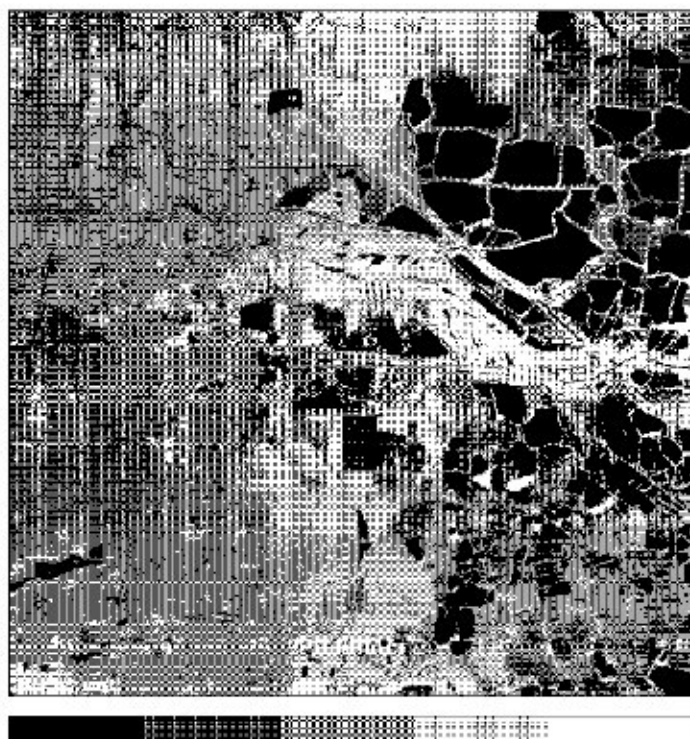Figure 21. Segmented Bombay (IRS) image using HCM when $c = 5$, $\beta = 16.056$.

Figure 22.   Segmented Calcutta (SPOT) image with highest β (= 9.375) value and $c = 5$.
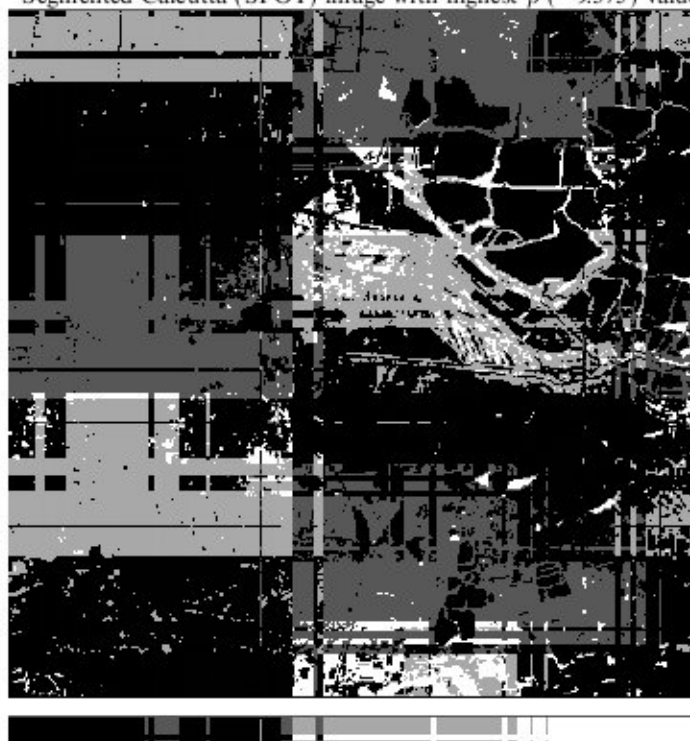

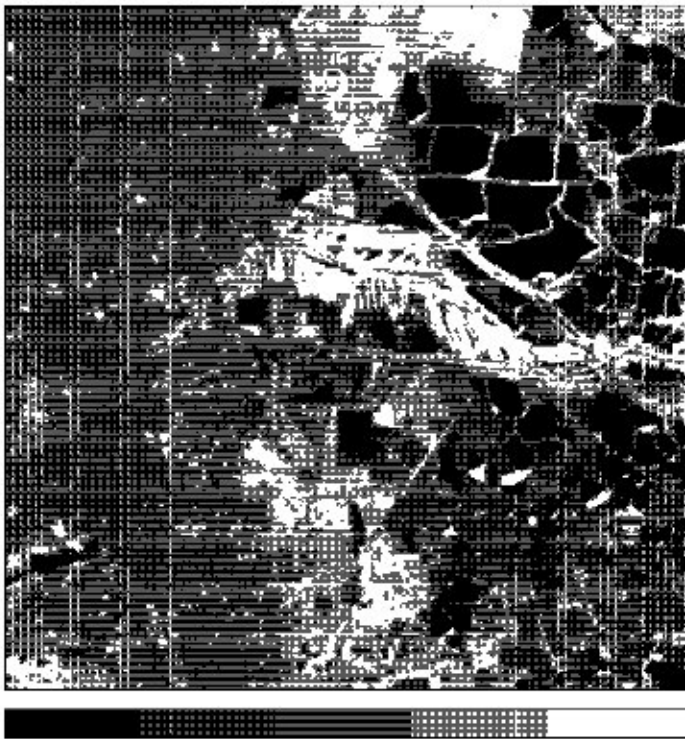Figure 23.   Segmented Calcutta (SPOT) image with lowest β (= 6.467) value and $c = 5$.

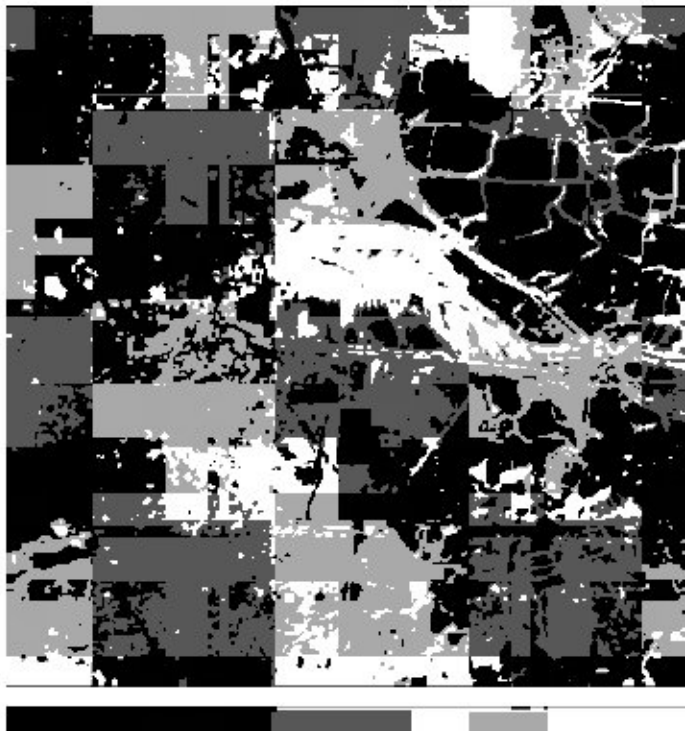Figure 24. Segmented Calcutta (SPOT) image using FCM when $c = 5$, $\beta = 6.388$.

Figure 25. Segmented Calcutta (SPOT) image using HCM when $c = 5$, $\beta = 6.676$.

land cover types are concerned. Note that membership functions corresponding to $w = 9$ and 11 satisfy the criteria of bounds of Murthy and Pal (1992a). Fuzzy geometry based techniques which incorporate spatial ambiguities produced single and almost stable thresholds irrespective of the window size. Overall, the study indicates that fuzzy set theoretic techniques mostly have an upper edge over the non-fuzzy techniques.

FCM provided superior performance to HCM. The time requirement of a thresholding technique, on average, is seen to be of the order of $\frac{1}{100}$th of HCM and $\frac{1}{1000}$th of FCM. The possibility of producing better results by thresholding as compared with clustering methods is not ruled out. For example, the segmentation produced in figure 10 is much better than those in figures 12 and 13 as far as the extraction of both regions and structural details is concerned.

References

ANDERBERG, M. R., 1973, *Cluster Analysis for Applications* (New York: Academic Press).

BARALDI, A., and PARMIGGIANI, F., 1995, A neural network for unsupervised categorization of multivalued input patterns: an application to satellite image clustering. *IEEE Transactions on Geoscience and Remote Sensing*, **33**, 305–316.

BARZOHAR, M., and COOPER, D. B., 1993, Automatic finding of main roads in aerial images by using geometric–stochastic models and estimation. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (Los Alamitos: IEEE CS Press), pp. 459–464.

BENEDIKTSSON, J. A., SWAIN, P. H., and ERSOY, O. K., 1990, Neural network approach versus statistical methods in classification of multisource remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, **28**, 540–552.

BEZDEK, J. C., 1981, *Pattern Recognition with Fuzzy Objective Function Algorithms* (New York: Plenum).

BEZDEK, J. C., and PAL, S. K. (Eds), 1992, *Fuzzy Models for Pattern Recognition: Methods that Search for Structures in Data* (New York: IEEE Press).

CANNON, R. L., DAVE, R., BEZDEK, J. C., and TRIVEDI, M., 1986, Segmentation of a thematic mapper image using the fuzzy c-means clustering algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, **24**, 400–408.

DECATUR, S. E., 1989, Application of neural network to terrain clasification. In *Proceedings of IJCNN'89, Washington, DC*, Vol. I (N.J.: IEEE Press), pp. 283–288.

DERAVI, F., and PAL, S. K., 1983, Gray level thresholding using second order statistics. *Pattern Recognition Letters*, **1**, 417–422.

FISHER, L., and VAN NESS, 1971, Admissible clustering procedure. *Biometrika*, **58**, 91–104.

GHOSH, A., 1995, Use of fuzziness measures in layered networks for object extraction: a generalization. *Fuzzy Sets and Systems*, **72**, 331–348.

GONZALEZ, R. C., and WOOD, R. E., 1993, *Digital Image Processing* (Reading: Addison-Wesley).

HU, J., SAKODA, B., and PAVLIDIS, T., 1992, Interactive rod finding for aerial images. In *Proceedings of IEEE Workshop on Applications of Computer Vision* (Los Alamitos, California: IEEE Computer Society Press), pp. 56–63.

KAPUR, J. N., SAHOO, P. K., and WONG, A. K. C., 1985, A new method for gray level picture thresholding using the entropy of histogram. *Computer Vision, Graphics and Image Processing*, **29**, 273–285.

LAPRADE, R. H., 1988, Split-and-merge segmentation of aerial photographs. *Computer Vision, Graphics and Image Processing*, **48**, 77–86.

LEE, J., WEGER, R. C., SENGUPTA, S. K., and WELCH, R. M., 1990, A neural network approach to cloud classification. *IEEE Transactions on Geoscience and Remote Sensing*, **28**, 846–855.

MANDAL, D. P., MURTHY, C. A., and PAL, S. K., 1994, Utility of multiple choices in detecting ill-defined roadlike structures. *Fuzzy Sets and Systems*, **64**, 213–228.

MURTHY, C. A., and PAL, S. K., 1992a, Bounds for membership functions a correlation based approach. *Information Sciences*, **65**, 143–171.

MURTHY, C. A., and PAL, S. K., 1992b, Histogram thresholding by minimizing gray level fuzziness. *Information Sciences*, **60**, 107–135.

MURTHY, C. A., PAL, S. K., and DUTTA MAJUMDER, D., 1985, Correlation between two fuzzy membership functions. *Fuzzy Sets and Systems*, **17**, 23–38.

PAL, N. R., and PAL, S. K., 1989, Entropic thresholding. *Signal Processing*, **16**, 97–108.

PAL, N. R., and PAL, S. K., 1991, Entropy: a new definition and its applications. *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-21**, 1260–1270.

PAL, N. R., and PAL, S. K., 1992, Higher order fuzzy entropy and hybrid entropy of a set. *Information Sciences*, **61**, 211–231.

PAL, N. R., and PAL, S. K., 1993, A review on image segmentation. *Pattern Recognition*, **24**, 1277–1294.

PAL, S. K., 1982, A note on the quantitative measure of image-enhancement through fuzziness. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **PAMI-4**, 204–208.

PAL, S. K., and DUTTA MAJUMDER, D., 1986, *Fuzzy Mathematical Approach to Pattern Recognition* (New York: John Wiley, Halsted Press).

PAL, S. K., and GHOSH, A., 1992a, Fuzzy geometry in image analysis. *Fuzzy Sets and Systems*, **48**, 23–40.

PAL, S. K., and GHOSH, A., 1992b, Image segmentation using fuzzy correlation. *Information Sciences*, **62**, 223–250.

PAL, S. K., and ROSENFELD, A., 1988, Image enhancement and thresholding by optimization of fuzzy compactness. *Pattern Recognition Letters*, **7**, 77–86.

PAL, S. K., KING, R. A., and HASHIM, A. A., 1983, Automatic grey level thresholding through index of fuzziness and entropy. *Pattern Recognition Letters*, **1**, 141–146.

PREWITT, J. M. S., 1970, Object enhancement and extraction. In *Picture Processing and Psychopictorics*, edited by B. S. Lipkin and A. Rosenfeld (New York: Academic Press), pp. 75–149.

RICHARDS, J. A., 1993, *Remote Sensing Digital Image Analysis: An Introduction*, 2nd edn (New York: Springer Verlag).

ROSENFELD, A., and KAK, A. C., 1982, *Digital Picture Processing*, Vols I and II (New York: Academic Press).

SAHASRABUDHE, S. C., and DASGUPTA, K. S., 1992, A valley-seeking threshold selection technique. In *Computer Vision, Graphics and Image Processing*, edited by L. Shapiro and A. Rosenfeld (Boston: Academic Press), pp. 55–65.

SAHOO, P. K., SOLTANI, S., WONG, A. K. C., and CHEN, Y. C., 1988, A survey of thresholding techniques. *Computer Vision, Graphics and Image Processing*, **41**, 233–260.

SHANNON, C. E., 1948, A mathematical theory of communication. *Bell System Technical Journal*, **27**, 379–423.

SWAIN, P. H., and DAVIS, M., 1978, *Remote Sensing: The Quantitative Approach* (New York: McGraw Hill).

THIRUVENGADACHARI, S., and KALPANA, A. R., and revised by ADIGA, S., and SREENIVASI, M., 1989, *IRS Data Users Handbook (Revision 1)* (India: Department of Space, Government of India, NRSA Data Centre, NRSA).

TON. J., 1988, A Knowledge Based Approach for LANDSAT Image Interpretation. PhD thesis, Michigan State University, Michigan, USA.

TRIVEDI, M., and BEZDEK, J. C., 1986, Low-level segmentation of aerial images with fuzzy clustering. *IEEE Transactions on Systems, Man, and Cybernetics*, **16**, 589–598.

XIE, W. X., and BEDROSIAN, S. D., 1988, Experimentally driven fuzzy membership function for gray level images. *Journal of Franklin Institute*, **325**, 154–164.

ZADEH, L. A., 1965, Fuzzy sets. *Information and Control*, **8**, 338–353.

ZLOTNICK, A., and CARNINE, P. D., JR., 1993, Finding road seeds in aerial images. *Image Understanding*, **57**, 307–330.