Pattern Recognition, Vol. 26, No. 9, pp. 1277–1294, 1993

A REVIEW ON IMAGE SEGMENTATION TECHNIQUES

NIKHIL R. PAL and SANKAR K. PAL
Machine Intelligence Unit, Indian Statistical Institute, 203 B.T. Road, Calcutta 700035, India

Abstract—Many image segmentation techniques are available in the literature. Some of these techniques use only the gray level histogram, some use spatial details while others use fuzzy set theoretic approaches. Most of these techniques are not suitable for noisy environments. Some works have been done using the Markov Random Field (MRF) model which is robust to noise, but is computationally involved. Neural network architectures which help to get the output in real time because of their parallel processing ability, have also been used for segmentation and they work fine even when the noise level is very high. The literature on color image segmentation is not that rich as it is for gray tone images. This paper critically reviews and summarizes some of these techniques. Attempts have been made to cover both fuzzy and non-fuzzy techniques including color image segmentation and neural network based approaches. Adequate attention is paid to segmentation of range images and magnetic resonance images. It also addresses the issue of quantitative evaluation of segmentation results.

Image segmentation Fuzzy sets Thresholding Edge detection Clustering Relaxation Markov Random Field

1. INTRODUCTION

There are several types of images, namely, light intensity (visual) image, range image (depth image), nuclear magnetic resonance image (commonly known as magnetic resonance image (MRI)), thermal image and so on. Light intensity (LI) images, the most common type of images we encounter in our daily experience, represent the variation of light intensity on the scene. Range image (RI), on the other hand, is a map of depth information at different points on the scene. In a digital LI image the intensity is quantized, while in the case of RI the depth value is digitized. Nuclear magnetic resonance images represent the intensity variation of radio waves generated by biological systems when exposed to radio frequency pulses. Biological bodies (humans/animals) are built up of atoms and molecules. Some of the nuclei behave like tiny magnets,[1] commonly known as spins. Therefore, if a patient (or any living being) is placed in a strong magnetic field, the magnetic nuclei tend to align with the applied magnetic field. For MRI the patient is subjected to a radio frequency pulse. As a result of this the magnetic nuclei pass into a high energy state, and then immediately relieve themselves of this stress by emitting radio waves through a process called relaxation. This radio wave is recorded to form the MRI. There are two different types of relaxation: longitudinal relaxation and transverse relaxation resulting in two types of MRIs, namely, T1 and T2, respectively.[1] In digital MRI, the intensity of the radio wave is digitized with respect to both intensity and spatial coordinates. Thus in general, any image can be described by a two-dimensional function $f'(x, y)$, where (x, y) denotes the spatial coordinate and $f'(x, y)$ the feature value at (x, y). Depending on the type of image, the feature value could be light intensity, depth, intensity of radio wave or temperature. A digital image, on the other hand, is a two-dimensional discrete function $f(x, y)$ which has been digitized both in spatial coordinates and magnitude of feature value. We shall view a digital image as a two-dimensional matrix whose row and column indices identify a point, called a pixel, in the image and the corresponding matrix element value identifies the feature intensity level. Throughout this paper a digital image will be represented as

$$F_{P \times Q} = [f(x, y)]_{P \times Q} \tag{1}$$

where $P \times Q$ is the size of the image and $f(x, y) \in G_L = \{0, 1, \ldots, L-1\}$, the set of discrete levels of the feature value. Since the majority of the techniques we are going to discuss in this paper are developed primarily for ordinary intensity images, in our subsequent discussion, we shall usually refer to $f(x, y)$ as gray value (although it could be depth or temperature or intensity of radio wave).

Segmentation is the first essential and important step of low level vision.[2–5] There are many applications of segmentation. For example, in a vision guided car assembly system, the robot needs to pick up the appropriate components from the bin. For this, segmentation followed by recognition is required. Its application area varies from the detection of cancerous cells to the identification of an airport from remote sensing data, etc. In all these areas, the quality of the final output depends largely on the quality of the segmented output. Segmentation is a process of partitioning the image into some non-intersecting regions such that each region is homogeneous and the union of no two adjacent regions is homogeneous. Formally, it can be defined[6] as follows: if F is the set of all pixels and $P(\ )$ is a uniformity (homogeneity) predicate defined on groups of connected pixels, then segmentation is a partitioning of the set F into a set of connected subsets or regions $(S_1, S_2, \ldots, S_n)$ such that

$$\bigcup_{i=1}^{n} S_i = F \quad \text{with} \quad S_i \bigcap S_j = \varnothing, \quad i \neq j. \tag{2}$$

The uniformity predicate $P(S_i) =$ true for all regions (S_i) and $P(S_i \cup S_j) =$ false, when S_i is adjacent to S_j. Note that this definition is applicable to all types of images we have described. For LI images the uniformity predicate measures the uniformity of light intensity, while for range images it could be the uniformity of surfaces.

Hundreds of segmentation techniques are present in the literature, but there is no single method which can be considered good for all images, nor are all methods equally good for a particular type of image. Moreover, algorithms developed for one class of image (say ordinary intensity image) may not always be applied to other classes of images (MRI/RI). This is particularly true when the algorithm uses a specific image formation model. For example, some visual image segmentation algorithms are based on the assumption that the gray level function $f(x, y)$ can be modeled as a product of an illumination component and a reflectance component.[7] On the other hand, in Pal and Pal[8] the gray level distributions have been modeled as Poisson distributions, based on the theory of formation of visual images. Such methods[7,8] should not be applied to MRI/RIs. However, most of the segmentation methods developed for one class of images can be easily applied/extended to another class of images. For example, the variable order surface fitting method,[9] although developed for range images can be applied for other images that can be modeled as a noisy version of piece-wise smooth surfaces.

There are many challenging issues like, the development of a unified approach to image segmentation which can (probably) be applied to all kinds of images. Even the selection of an appropriate technique for a specific type of image is a difficult problem. Up to now, to the knowledge of the authors there is no universally accepted method of quantification of segmented output. Authentication of edges is also a very important task. Different edge operators[3-5,10] like Sobel, Prewitt, Marr–Hildreth, etc. produce an edginess value at every pixel location. However, all of them are not valid(!) candidates for edges. Normally, edges are required to be thresholded. The selection of the threshold is very crucial as for some part of the image low intensity variation may correspond to edges of interest, while the other part may require high intensity variation. Adaptive thresholding[11-13] often is taken as a solution to this. Obviously it cannot eliminate the problem of threshold selection. A good strategy to produce meaningful segments would be to fuse region segmentation results and edge outputs.[14,15] Incorporation of psychovisual phenomena[16,17] may be good for light intensity images but not applicable for range images. Actually semantics and a prior information about the type of images are critical to the solution of the segmentation problem.[18] According to Pavlidis[18,19] (visual) image segmentation is a problem of psycho-physical perception, and therefore, not susceptible to purely analytical solution. Any mathematical algorithm usually should be supplemented by heuristics which involve semantic information about the class of images under consideration.

One may attempt to extract the segments in a variety of ways. Broadly, there are two approaches namely, classical approach and fuzzy mathematical approach. Under the classical approach we have segmentation techniques based on histogram thresholding, edge detection, relaxation, and semantic and syntactic approaches.[11-113] In addition to these, there are certain other methods which do not fall clearly in any one of the above classes.[114-121] Similarly, the fuzzy mathematical approach[122-154] also has methods based on edge detection, thresholding and relaxation. Some of these methods, particularly the histogram based methods are not at all suitable for noisy images. Several attempts have also been made to develop image processing algorithms using neural network (NN) models,[155-170] particularly the Hopfield and Kohonen networks. These algorithms work well even in a highly noisy environment and they are capable of producing outputs in real time. Though many algorithms are available for color image segmentation,[171-177] the literature is not that rich as it is for the gray level images. In this context it may be mentioned that the literature is very rich on the methods of segmentation, but not many attempts have been made for the objective evaluation of segmented outputs.

This paper attempts to critically review and summarize some of the existing methods of segmentation. Before we proceed further, we summarize some of the earlier surveys on image segmentation. Fu and Mui[18] categorized segmentation techniques into three classes: (1) characteristic feature thresholding or clustering, (2) edge detection, and (3) region extraction. This survey was done from the viewpoint of cytology image processing. A critical appreciation of several methods of thresholding, edge detection and region extraction[20-33] has been done. This includes some graph theoretic approaches[30] also. For color image thresholding, only a brief mention about it has been given.[22] The section on edge detection makes a good summarization of several edge detection approaches including some adaptive local operators.[31,32] Hueckel's[33] approach of viewing edge detection as a functional approximation problem has been discussed.

Haralick and Shapiro[34] classified image segmentation techniques as: (1) measurement space guided spatial clustering, (2) single linkage region growing schemes, (3) hybrid linkage region growing schemes, (4) centroid linkage region growing schemes, (5) spatial clustering schemes, and (6) split and merge schemes. According to them, the difference between clustering and segmentation is that in clustering, the grouping is done in measurement space; while in image segmentation, the grouping is done in the spatial domain of the image. We like to emphasize that segmentation tries to do the groupings in the spatial domain but it can be achieved through groupings in the measurement space, particularly for multispectral images.[137,177] For multispectral data, instead of clustering in the full measurement space, Haralick and Shapiro[34] suggested to work in multiple lower order projection spaces, and then reflect these clusters back to the full measurement space as follows: suppose, for example, that the clustering is done on a four band image. If the clustering done in bands 1 and 2 yields clusters c_1, c_2, c_3 and the clustering done in bands 3 and 4 yields clusters c_4 and c_5 then each possible 4-tuple from a pixel can be given a cluster label from the set "$\{(c_1, c_4), (c_1, c_5), (c_2, c_4), (c_2, c_5), (c_3, c_4)(c_3, c_5)\}$". A 4-tuple (x_1, x_2, x_3, x_4) gets the cluster labels (c_2, c_4) if (x_1, x_2) is in cluster c_2 and (x_3, x_4) is in cluster c_4. However, this does not seem to be of any use to us as this virtually assigns a point (a 4-tuple) in two different classes. Note that it is neither a probabilistic assignment nor a fuzzy assignment. A good summary of different types of linkage region growing algorithm[35-41] has also been presented.

Sahoo *et al.*[42] surveyed only segmentation algorithms based on thresholding and attempted to evaluate the performance of some thresholding algorithms using some uniformity and shape measures. They categorized[42] global thresholding techniques into two classes: point dependent techniques (gray level histogram based) and region dependent techniques (modified histogram or co-occurrence based). A fairly detailed discussion on probabilistic relaxation[42] is available. They also reviewed several methods of multi-thresholding techniques.[43-45] We offer the following comments about the previous reviews on image segmentation:

(1) None of these surveys[18,34,42] considers fuzzy set theoretic segmentation techniques.

(2) Neural networks based techniques are also not included.

(3) The problem of objective evaluation of segmentation results has not been adequately dealt with except in Sahoo *et al.*[42]

(4) Color image segmentation has not been paid proper attention.

(5) Segmentation of range images/magnetic resonance images has not been considered at all.

This review paper attempts to incorporate all these points to a limited but reasonable extent. However, by no means is it an exhaustive survey.

2. GRAY LEVEL THRESHOLDING

Thresholding is one of the old, simple and popular techniques for image segmentation. Thresholding can be done based on global information (e.g. gray level histogram of the entire image) or it can be done using local information (e.g. co-occurrence matrix) of the image. Taxt *et al.*[46] refer to the local and global information based techniques as contextual and non-contextual methods, respectively. Under each of these schemes (contextual/non-contextual) if only one threshold is used for the entire image then it is called global thresholding. On the other hand, when the image is partitioned into several subregions and a threshold is determined for each of the subregions, it is referred to as local thresholding.[46] Some authors[11-13] call these local thresholding methods adaptive thresholding schemes. Thresholding techniques can also be classified as bilevel thresholding and multithresholding. In bilevel thresholding the image is partitioned into two regions—object (black) and background (white). When the image is composed of several objects with different surface characteristics (for a light intensity image, objects with different coefficient of reflection, for a range image there can be objects with different depths and so on) one needs several thresholds for segmentation. This is known as multithresholding. In such a situation we try to get a set of thresholds $(t_1, t_2, \ldots, t_k)$ such that all pixels with $f(x, y) \in [t_i, t_{i+1})$, $i = 0, 1, \ldots, k$; constitute the ith region type (t_0 and t_{k+1} are taken as 0 and $L - 1$, respectively). Note that thresholding can also be viewed as a classification problem. For example, bilevel segmentation is equivalent to classifying the pixels into two classes: object and background. Mardia and Hainsworth[58] have shown that the main idea behind the iterative thresholding schemes of Ridler and Calvard[63] and Lloyd[81] can be defined as special cases of the classical Bayes' discrimination rule. Under the assumption that object and background pixels are normally distributed with the same variance, Bayes' allocation rule yields the formula used for threshold computation in reference (81). With an additional assumption that the prior probabilities for object and background pixels are the same, Bayes' formula reduces to the computation formula for threshold in Ridler and Calvard.[63]

If the image is composed of regions with different gray level ranges, i.e. the regions are distinct, the histogram of the image usually shows different peaks, each corresponding to one region and adjacent peaks are likely to be separated by a valley. For example, if the image has a distinct object on a background, the gray level histogram is likely to be bimodal with a deep valley. In this case, the bottom of the valley (T) is taken as the threshold for object background separation. Therefore, when the histogram has a (or a set of) deep valley(s), selection of threshold(s) becomes easy because it becomes a problem of detecting valleys. However, normally the situation is not like this and threshold selection is not a trivial job. There are various methods[3-8,42-81] available for this. For example, Otsu[52] maximized a measure of class separability. He maximized the ratio of the between class variance to the local variance to obtain thresholds. Nakagawa and Rosenfeld[12] assumed that the object and background populations are distributed normally with distinct means and standard deviations. Under this assumption they selected the threshold by minimizing the total misclassification error. This method is computationally involved. Kittler and Illingworth,[62] under the same assumption of normal mixture, suggested a computationally less involved method. They proposed a method which optimizes a criterion function related to average pixel classification error rate that finds out an approximate minimum error threshold. Pal and Bhandari[71] optimized the same criterion function but assumed Poisson distributions to model the gray level histogram.

Pun[65] assumed that an image is the outcome of an L symbol source. He maximized an upper bound of the total a posteriori entropy of the partitioned image for the purpose of selecting the threshold. Kapur *et al.*,[49] on the other hand, assumed two probability distributions, one for the object area and the other for the background area. They then maximized the total entropy of the partitioned image in order to arrive at the threshold level. Wong and Sahoo[69] maximized the a posterior entropy of a partitioned image subject to a constraint on the uniformity measure of Levine and Nazif[95] and a shape measure. They maximized the a posterior entropy over min (s_1, s_2) and max (s_1, s_2) to get the threshold for segmentation; where s_1 and s_2 are the threshold levels at which the uniformity and the shape measure attain the maximum values, respectively. Pal and Pal[8] modeled the image as a mixture of two Poisson distributions and developed several parametric methods for segmentation. The assumption of the Poisson distribution has been justified based on the theory of image formation. These algorithms maximize either entropy or minimize the χ^2 statistic. Though these methods use only the histogram, they produce good results due to the incorporation of the image formation model.

All these methods have a common drawback, they take into account only the histogram information (ignoring the spatial details). As a result, such an algorithm may fail to detect thresholds if these are not properly reflected as valleys in the histogram, which is normally the case. There are many thresholding schemes that use spatial information, instead of histogram information. For example, the busyness measure of Weszka and Rosenfeld[75] is dependent on the co-occurrence of adjacent pixels in an image. They minimized the busyness measure in order to arrive at the threshold for segmentation. Deravi and Pal[76] minimized the conditional probability of transition across the boundary between two regions. This method also uses the local information contained in the co-occurrence matrix of the image. However, finally all these methods threshold the histogram, but since they make use of the spatial details, they result in a more meaningful segmentation than the methods which use only the histogram information. Based on the co-occurrence matrix, Chanda *et al.*[77] have given an average contrast measure for segmentation. Pal and Pal[16,67] proposed measures of contrast between regions and homogeneity of regions using the brightness perception aspect of the human psycho-visual system, and applied them to segmentation. They also defined[48] the higher order entropy and conditional entropy of an image giving measures of homogeneity and contrast, respectively. These measures are finally applied to develop object extraction algorithms. A concept similar to the second-order local entropy of Pal and Pal[48] has been used by Abutaleb[90] for segmentation. The gray value of a pixel and the average of its neighboring pixels have been used there for the computation of the co-occurrence matrix. As a result the boundary of the segmented object usually becomes blurred.

The philosophy behind gray level thresholding, "pixels with gray level $\leq T$ fall into one region and the remaining pixels belong to another region", may not be true on many occasions, particularly, when the image is noisy or the background is uneven and illumination is poor. In such cases the objects will still be lighter or darker than the background, but any fixed threshold level for the entire image will usually fail to separate the objects from the background. This leads one to the methods of adaptive thresholding. In adaptive thresholding[11–13] normally the image is partitioned into several non-overlapping blocks of equal area and a threshold for each block is computed independently. Chow and Kaneko[15] used the (sub) histogram of each block to determine local threshold values for the corresponding cell centers. These local thresholds are then interpolated over the entire image to yield a threshold surface. They[11] used only gray level information. Yanowitz and Bruckstein[13] extended this idea to use combined edge and gray level information. They computed the gray level gradient magnitude from a smooth version of the image. The gradient values have then been thresholded and thinned using a local maxima directed thinning process. Locations of these local gradient maxima are taken as boundary pixels between object and background. The corresponding gray levels in the image are taken as local thresholds. The sampled gray levels are then interpolated over the entire image to obtain an adaptive threshold surface. Several approaches to the two-dimensional interpolation problem have been discussed. The performance of the algorithm is likely to depend on the choice of the threshold levels for the gradients and no guideline has been provided for this.

3. ITERATIVE PIXEL CLASSIFICATION

3.1. *Relaxation*

Relaxation[3,94,96] is an iterative approach to segmentation in which the classification decision about each pixel can be taken in parallel. Decisions made at neighboring points in the current iteration are then combined to make a decision in the next iteration. There are two types of relaxation: probabilistic and fuzzy. We discuss here the probabilistic relaxation. Suppose a set of pixels $\{f_1, f_2, \ldots, f_n\}$ is to be classified into m classes $\{C_1, C_2, \ldots, C_m\}$. For the probabilistic relaxation we assume that for each pair of class assignments $f_i \in C_j$ and $f_h \in C_k$, there exists a quantitative measure of compatibility $C(i, j; h, k)$ of this pair, i.e. the class assignment of pixels is interdependent. It is reasonable to assume that a positive value of $C(i, j; h, k)$

indicates the compatibility of $f_i \in C_j$ and $f_h \in C_k$, while a negative value represents incompatibility and a zero don't care situation. The function C need not be symmetric.

Let p_{ij} represent the probability that $f_i \in C_j$, $1 \le i \le n$ and $1 \le j \le m$, with $0 \le p_{ij} \le 1$, $\sum_j p_{ij} = 1$. Intuitively, if p_{hk} is high and $C(i, j; h, k)$ is positive, we increase p_{ij} since it is compatible with the high probability event $f_h \in C_k$. Similarly, if p_{hk} is high and $C(i, j; h, k)$ is negative, we reduce p_{ij} as it is incompatible with $f_h \in C_k$. On the other hand, if p_{hk} is low or $C(i, j; h, k)$ is nearly zero, p_{ij} is not changed as either $f_h \in C_k$ has a low probability or is irrelevant to $f_i \in C_j$. The fuzzy relaxation is similar.

3.2. *MRF based approaches*

There are many image segmentation methods[114–121] which use the spatial interaction models like Markov Random Field (MRF) or Gibbs Random Field (GRF) to model digital images. Geman and Geman[118] have proposed a hierarchical stochastic model for the original image and developed a restoration algorithm, based on stochastic relaxation (SR) and annealing, for computing the maximum a posterior estimate of the original scene given a degraded realization. Due to the use of annealing, the restoration algorithm does not stop at a local maxima but finds the global maximum of the a posterior probability. We mention here that the probabilistic relaxation[94] (also known as relaxation labeling (RL)) and stochastic relaxation, although they share some common features like parallelism and locality, are quite distinct. RL is essentially a non-stochastic (deterministic) process which allows jumps to states (configurations) of lower energy. On the other hand, SR transition to a configuration which increases the energy (decreases the probability) is also allowed. In fact, if the new configuration decreases the energy, the system transits to that state, while if the new configuration increases the energy the system accepts that state with a probability. This helps the system to avoid the local minima. RL usually gets stuck in a local minima. Moreover, in RL there is nothing corresponding to an equilibrium state or even a joint probability law over the configurations. Derin *et al.*[117] extended the one-dimensional Bayes smoothing algorithm of Askar and Derin[120] to two dimensions to get the optimum Bayes estimate for the scene value at every pixel. In order to reduce the computational complexity of the algorithm, the scene is modeled as a special class of MRF models, called Markov mesh random fields which are characterized by causal transition distributions. The processing is done over relatively narrow strips and estimates are obtained at the middle section of the strips. These pieces together with overlapping strips yield a suboptimal estimate of the scene. Without parallel implementation these algorithms become computationally prohibitive. Derin and Elliott[121] used a doubly stochastic hierarchical model for image data. At the top level a Gibbs distribution (GD) is used to characterize the clusters of the image pixels into regions with similar features. At the bottom level, the feature or textural properties of region types are modeled by a second set of GD, one for each type of class. The segmentation algorithms are derived by using the maximum a posterior probability (MAP) criterion. To reduce the computational overhead of the exact MAP estimate, they derived suboptimal solutions through simplifying assumptions in the model. They formulated it as a dynamic programming problem. These algorithms require only one raster scan over the image.

3.3. *Neural network based approaches*

For any artificial vision application, one desires to achieve robustness of the system with respect to random noise and failure of processors. Moreover, a system can (probably) be made artificially intelligent if it is able to emulate some aspects of the human information processing system. Another important requirement is to have the output in real time. Neural network based approaches are attempts to achieve these goals. Neural networks are massively connected networks of elementary processors.[155–158] Architecture and dynamics of some networks are claimed to resemble information processing in biological neurons.[157] The massive connectionist architecture usually makes the system robust while the parallel processing enables the system to produce output in real time. Several authors[159–170] have attempted to segment an image using neural networks. Blanz and Gish[159] used a three-layer feed forward network for image segmentation, where the number of neurons in the input layer depends on the number of input features for each pixel and the number of neurons in the output layer is equal to the number of classes. Babaguchi *et al.*[160] used a multilayer network trained with backpropagation, for thresholding an image. The input to the network is the histogram while the output is the desirable threshold. In this method at the time of learning a large set of sample images with known thresholds which produce visually suitable outputs are required. But for practical applications it is very difficult to get many sample images.

Recently Ghosh *et al.*[162,166] used a massively connected network for extraction of objects in a noisy environment. The maximum a posterior probability estimate of a scene modeled as a GRF and corrupted by additive Gaussian noise has been done using a neural network.[162] The hardware realization of neurons to be used for such a network has also been suggested. This NN based method takes into account the contextual information, because the GRF model considers the spatial interactions among neighboring pixels. Another robust algorithm for the extraction of objects from highly noise corrupted scenes using a Hopfield type neural network has been developed in references (167, 169). The energy function of the network has been constructed in such a manner that in a stable state of the net it extracts compact regions from a noisy scene. A multilayer neural network[164] where each neuron in layer i ($i > 1$) is connected to the corresponding neuron in layer $(i-1)$ and some of its neighboring neurons (in layer $i-1$), has been used to segment noisy images. The output status of the neurons in the output layer has been viewed as a fuzzy set (to be defined in Section 7). The weight updating rules have been derived to minimize the fuzziness in the system. For this algorithm the architecture of the network enforces the system to consider the contextual information. Moreover, this algorithm integrates the advantages of both fuzzy sets (decision from imprecise/incomplete knowledge) and neural networks (robustness). Shah[165] formulated the problem of edge detection in the context of an energy minimizing model. The method is capable of eliminating weak boundaries and small regions. Cortes and Hertz[165] proposed a NN to detect potential edges in different orientations. The performance of the system has been investigated through simulation studies using simulated annealing and mean field annealing. In reference (168) the image segmentation problem has been formulated as a constraint satisfaction problem (CSP) and a class of constraint satisfaction neural network (CSNN) is proposed. A CSNN consists of a set of objects, a set of labels, a collection of constraint relations and a topological constraint describing the neighborhood relationships among various objects. The CSNN is viewed as a collection of interconnected neurons. The architecture is chosen in such a way that it represents constraints in the CSP. The proposed method is found to be successful on CT (computed tomography) images and MRIs. However, robustness of the algorithm with noisy data has not been investigated. Moreover, for references (164, 168) a large number of neurons are required even for an image of moderate size.

4. SURFACE BASED SEGMENTATION

This section mainly discusses a few selected techniques for range image segmentation.[9,14,15,82–84] Besl and Jain[9] have developed an image segmentation algorithm based on the assumption that the image data exhibits surface coherence, i.e. image data may be interpreted as noisy samples from a piece-wise smooth surface function. Though, this method is probably most useful for range images, it can be used to segment any type of image that can be modeled as a noisy sampled version of a piece-wise smooth graph surface. This method is based on the fact that the signs of Gaussian and mean curvatures yield a set of eight surface primitives: peak, pit, ridge, saddle ridge, valley, saddle valley, flat (planar) and minimal. These primitives possess some desirable invariant properties and can be used to decompose any arbitrary smooth surfaces. In other words, any arbitrary smooth surface can be decomposed into one of those eight possible surface types. These simple surfaces can be well approximated, for the purpose of segmentation, by bivariate polynomials of order ≤ 4. The first stage of the algorithm creates a surface type label image based on the local information (using mean curvature and Gaussian curvature images). The second stage takes the original image and the surface type image as input and performs an iterative region growing using the variable order surface fitting. In the variable order surface fitting, first it has been tried to represent the points in a seed region by a planar surface. If this simple hypothesis of planar surface is found to be true then the seed region is grown on the planar surface fit. If this simple hypothesis fails, then the next more complicated hypothesis of biquadratic surface fit is tried. If this is satisfied, the region is grown based on that form otherwise, the next complicated form is tried. The process is terminated when either the region growing has converged (same region obtained twice) or when all preselected hypotheses fail. In the later case, possibly a higher order surface should be tried.

Hoffman and Jain[82] have developed a method for segmentation and classification of range images. They have used a clustering algorithm to segment the image into surface patches. Different types of clustering algorithms including methods based on minimal spanning tree, mutual nearest neighbor, hierarchical clustering and square error clustering have been attempted. The square error clustering has been found to be the most successful method for range images. The feature set used contains the coordinate position (x, y), the depth value $f(x, y)$ and the estimated unit surface normal vector. The unit surface normal vector is normal to the tangent plane at a point which is obtained by finding the best (in the least square sense) fitting plane over a neighborhood. In the second phase of the method these patches are classified as planar, convex or concave. In order to make the method of classification more effective they have combined three different methods, namely, "non-parametric trend test for planarity", "curvature planarity test", and the "eigenvalue planarity test". In the final stage, boundaries between adjacent surface patches are classified as crease or non-crease edge, and this information is then used to merge adjacent compatible patches to result in reasonable faces of the object. For this type of method, the choice of the neighborhood to compute the local parameters is an important issue and no theoretical guideline has been provided for this.

Yokoya and Levine[14] also used a differential geometric technique like Besl and Jain[9] for range image segmentation. Yokoya and Levine[14] combined both region and edge based considerations. They approximated object surfaces using biquadratic polynomials. As in reference (9) signs of Gaussian and mean curvatures (curvature sign map) have been used to get the initial region based segmentation. Two edge maps are formed: one for the jump edge and the other for the roof edge. The jump edge magnitude is obtained by computing the maximum difference in depth between a point and its eight neighbors; while the roof edge magnitude is computed as the maximum angular difference between adjacent unit surface normals. These two edge maps and the curvature sign map are then fused to form the final segmentation. This method too requires selection of threshold levels for the maps and the curvature sign map. Improper choice of these parameter values is likely to deteriorate the quality of segmentation output. At this point we note that for range images, detection of jump edges can be done with ordinary gradient operators, but detection of crease edges with ordinary gradient operators becomes difficult. For an inclined plane depth value changes slowly and hence any difference operator is likely to respond resulting in false edges. Often magnitude of the crease edge is computed as the maximum angular difference between adjacent unit surface normals. Note that the maximum angular difference method may (usually will) fail to detect jump edges.

Thus for edge detection in range images one needs to account for both crease and jump edges separately. Rimey and Cohen[84] formulated the problem as a maximum likelihood (ML) segmentation problem. Here also the objective is to divide the range image into windows, classify each window as a particular surface primitive, and group like windows into surface regions. Homogeneous windows are classified according to a generalized likelihood ratio test. This test uses information from adjacent windows and is computationally simple. Once each window has been classified, similar windows are merged using ML clustering analysis.

5. SEGMENTATION OF COLOR IMAGES

Color is a very important perceptual phenomenon related to human response to different wavelengths in the visible electromagnetic spectrum.[22,171] The image is usually described by the distribution of three color components R (red), G (green), B (blue). Color image is often also represented by three psychological qualities—hue, saturation and intensity. These color features and many others can be calculated from the tristimuli R, G and B by either a linear or a non-linear transformation. Ohta *et al.*[174] attempted to find a set of effective color features by systematic experiments in region segmentation. They applied an Ohlander type segmentation algorithm for the experiment.[175] At every step of segmenting a region, calculation of the new color features is done for the pixels in that region by the Karhunen–Loave (KL) transform of R, B and G data. Based on extensive experiments, it has been found that the following three color features $I1 = (R + B + G)/3$, $I2 = (R - B)/2$ or $(B - R)/2$ and $I3 = (2G - R - B)/4$ constitute an effective set of features for segmentation.

Spectrum analysis is another technique of color image segmentation in which prior knowledge about object colors is used to classify pixels. However, in many real life applications prior knowledge about the colors of the object may be difficult to gather. Under this situation clustering techniques can be used. Ohta *et al.*[174] instead of using the R–B–G color coordinate directly, used I1, I2 and I3. Lim and Lee[172] developed a two-stage color image segmentation technique based on thresholding and fuzzy c-means (FCM) methods.[172] The FCM method will be discussed in Section 7. This method[172] can be viewed as a coarse to fine technique which tries to reduce the computational overhead of FCM. The method is similar to the iterative algorithm proposed by Huntsberger *et al.*[173] except it uses the scale space filter for finding the number of clusters. The coarse segmentation attempts to segment using thresholding and then the FCM algorithm is used to classify pixels which have not yet been assigned to any class in the coarse segmentation phase. Though the method is claimed to find the number of classes automatically, it does have some subjective choices. For example, in the coarse segmentation phase if the number of pixels in a class exceeds a prespecified threshold, then only it is taken as a valid class. We mention here that a color image is a special case of multispectral images and algorithms developed for multispectral images[137] usually can be used for color image segmentation.

6. EDGE DETECTION

Segmentation can also be obtained through detection of edges of various regions, which normally tries to locate points of abrupt changes in gray level intensity values. As discussed in the previous section, for range images edges are declared at points of significant changes in depth values. Since edges are local features, they are determined based on local information. A large variety of methods are available in the literature[3–5,97–113] for edge finding. Davis[18,109] classified edge detection techniques into two categories: sequential and parallel. In the sequential technique the decision whether a pixel is an edge pixel or not is dependent on the result of the detector at some previously examined pixels. On the other hand, in the parallel method the decision whether a point is an edge or not is made based on the point under consideration and some of its neighboring points. As a result of this the operator can be applied to every point in the image simultaneously. The performance of a sequential edge detection method is dependent on the choice of an appropriate starting point and how the results of previous points influence the selection and result of the next point. Kelly[110] and Chien and Fu[111] used guided search techniques for this. Chien and Fu[111] detected cardiac and lung boundaries in chest X-ray images using a sequential search technique with an evaluation function.

There are different types of parallel differential operators such as Roberts gradient, Sobel gradient, Prewitt gradient and the Laplacian operator.[3–5] These difference operators respond to changes in gray level or average gray level. The gradient operators, not only respond to edges but also to isolated points. For Prewitt's operator the response to the diagonal edge is weak, while for Sobel's operator it is not that weak as it gives greater weights to points lying close to the point (x, y) under consideration. However, both Prewitt's and Sobel's operators possess greater noise immunity. The preceding operators are called the first difference operator. Laplacian, on the other hand, is a second difference operator. The Laplacian operator is given by

$$\nabla^2 = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}. \tag{3}$$

The digital Laplacian being a second difference operator, has a zero response to linear ramps. It responds strongly to corners, lines, and isolated points. Thus for a noisy picture, unless it has a low contrast, the noise will produce higher Laplacian values than the edges. Moreover, the digital Laplacian is not orientation invariant. A good edge detector, should be a filter with the following two features. First, it should be a differential operator, taking either a first or second spatial derivative of the image. Second, it should be capable of being tuned to act at any desired scale, so that large filters can be used to detect blurry shadow edges, and small ones to detect sharply focused fine details. The second requirement is very useful as intensity changes occur at different scales in an image. According to Marr and Hildreth[10] the most satisfactory operator fulfilling these conditions is the Laplacian of Gaussian (LG) operator. It is normally denoted by $\nabla^2 G$; where the Laplacian is as given by equation (3) and

$$G = e^{(x^2+y^2)/(2\pi\sigma^2)} \tag{4}$$

is a two-dimensional Gaussian distribution, with standard deviation σ. The Gaussian part of the LG operator blurs the image, wiping out all structures at scales much smaller than the σ of the Gaussian.[10] The Gaussian blurring function is preferred over others because it has the desirable property of being smooth and localized in both spatial and frequency domains. In order to find the intensity change at a given scale, Marr and Hildreth, first filtered the image with the $\nabla^2 G$ filter and then found the zero-crossings in the filtered image. The space described by the scale parameter σ and the zero-crossing curves is called the scale space. The behavior of edges in the scale space produced by the LG operator has been studied by Lu and Jain.[106] In order to formulate rules for reasoning in the scale space they studied dislocation of edges, false edges, and merging of edges with nice mathematical frames.

According to Canny[105] a good edge detector should have the following three properties: (1) low probability of wrongly marking non-edge points and low probability of failing to mark real edge points (i.e. good detection); (2) points marked as edges should be as close as possible to the center of true edges (i.e. good localization); and (3) one and only one response to a single edge point (single response). Good detection can be achieved by maximizing signal to noise ratio (SNR), while for good localization Canny used the reciprocal of an estimate of the r.m.s. distance of the marked edge from the center of the true edge. To maximize simultaneously both good detection and localization criteria Canny[105] maximized the product of SNR and the reciprocal of standard deviation (approximate) of the displacement of edge points. The maximization of the product is done subject to a constraint which eliminates multiple responses to single edge points.

In the case of a noise free image, the edge angle can be measured accurately, but in real life images, noise cannot be avoided and it makes it difficult to estimate the true edge angles. Kittler *et al.*[98] suggested three methods to improve the edge angle estimate obtained from Sobel's operator. All the three methods involve averaging of the outputs of the Sobel operator over a 3×3 window. One of the methods, which ignores the effect of the central pixel, at which the angle estimate is wished, is found to produce the best result. They have justified this counterintuitive view also. Haralick[102] attacked the problem of edge and region detection from a new angle. He assumed that the observed image is an ideal image with noise added. Each region in the image is a sloped plane. In order to determine the edge between two pixels, best fitted sloped planes over a neighborhood of each pixel are found. Edges are declared at locations having significantly different planes on either side of them. The least square error procedure has been used to estimate the parameters of a sloped surface for a given neighborhood. An appropriate F statistic has been used to test the significance of the difference of the estimated slope from a zero slope or the significance of the difference of estimated slopes of adjacent neighbors.

An iterative algorithm has been developed by Gokmen and Li[112] using the regularization theory. The energy functional in the standard segmentation has been modified to spatially control the smoothness over the image in order to obtain the accurate location of edges. An algorithm for defining a small, optimal kernel conditioned on some important aspects of the imaging process has been suggested by Reichenbach *et al.*[113] for edge detection. This algorithm takes into account the nature of the scene, the point spread function of the image gathering device, the effect of noise, etc.; and generates the kernel values which minimize the expected mean square error of the estimate of the scene characteristics. We have discussed various operators to get edge values. All the edges produced by these operators are, normally, not significant (relevant) edges when viewed by human beings. Therefore, one needs to find out prominent (valid) edges from the output of the edge operators. Kundu and Pal[17] have suggested a method of thresholding to extract the prominent edges based psycho-visual phenomena. Haddon[108] developed a technique to derive a threshold for any edge operator, based on the noise statistics of the image.

7. METHODS BASED ON FUZZY SET THEORY

Zadeh introduced the concept of fuzzy sets in which imprecise knowledge can be used to define an event. A fuzzy set A is represented as

$$A = \{\mu_A(x_i)/x_i, \quad i = 1, 2, \ldots, n\} \tag{5}$$

where $\mu_A(x_i)$ gives the degree of belonging of the element x_i to the set A.

The relevance of fuzzy sets theory in pattern recognition problems has adequately been addressed in the literature.[122–125] It is seen that the concept of fuzzy sets can be used at the feature level in representing an input pattern as an array of membership values denoting the degree of possession of certain properties and in representing linguistically phrased input features; at the classification level in representing multi-class membership of an ambiguous pattern, and in providing an estimate (or a representation) of missing information in terms of membership values.[125] In other words, fuzzy set theory may be incorporated in handling uncertainties (arising from deficiencies of information: the deficiencies may result from incomplete, imprecise, ill-defined, not fully reliable, vague, contradictory information) in various stages of a pattern recognition system. While the application of fuzzy sets in cluster analysis and classifier design was in the process of development, an important and related effort in fuzzy image processing and recognition[126,133,137,141,177] was evolving more or less in parallel with the aforesaid general developments. This evolution was based on the realization that many of the basic concepts in image analysis, e.g. the concept of an edge or a corner or a boundary or a relation between regions, do not lend themselves well to precise definition. A gray tone image possesses ambiguity within pixels due to the possible multi-valued levels of brightness in the image. This indeterminacy is due to inherent vagueness rather than randomness. Incertitude in an image pattern may be explained in terms of grayness ambiguity or spatial (geometrical) ambiguity or both. Grayness ambiguity means "indefiniteness" in deciding whether a pixel is white or black. Spatial ambiguity refers to "indefiniteness" in the shape and geometry of a region within the image.

Conventional approaches to image analysis and recognition[2–5] consist of segmenting the image into meaningful regions, extracting their edges and skeletons, computing various features/properties (e.g. area, perimeter, centroid, etc.) and primitives (e.g. line, corner, curve, etc.) of and relationships among the regions, and finally, developing decision rules/grammars for describing, interpreting and/or classifying the image and its subregions. In a conventional system each of these operations involves crisp decisions (i.e. yes or no, black or white, 0 or 1) about regions, features, primitives, properties, relations and interpretations.

Since the regions in an image are not always crisply defined, uncertainty can arise within every phase of the

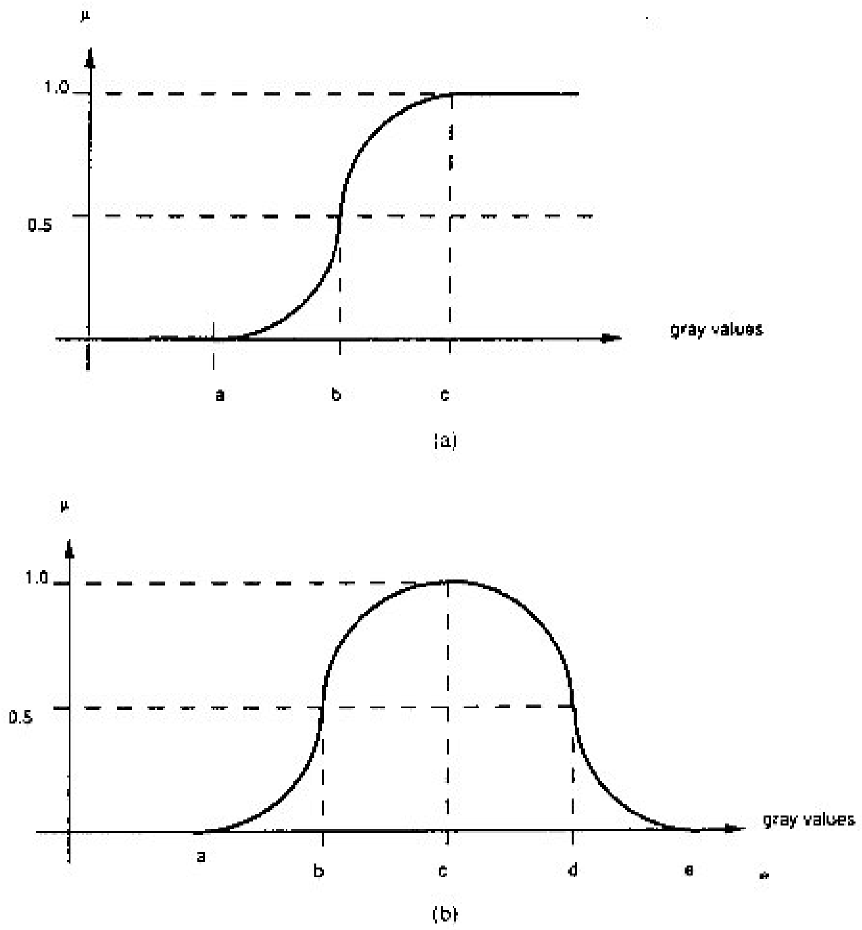

Fig. 1. Different types of membership function: (a) S-type membership function; (b) π-type membership function.

aforesaid tasks. Any decision made at a particular level will have an impact on all higher level activities. A recognition (or vision) system should have sufficient provision for representing and manipulating the uncertainties involved at every processing stage; i.e. in defining image regions, features, matching, and relations among them, so that the system retains as much of the "information content" of the data as possible. If this is done, the ultimate output (result) of the system will possess minimal uncertainty (and unlike conventional systems, it may not be biased or affected as much by lower level decision components).

For example, consider the problem of object extraction from a scene. Now, the question is "How can one define exactly the target or object region in a scene when its boundary is ill-defined?" Any hard thresholding made for the extraction of the object will propagate the associated uncertainty to subsequent stages (e.g. thinning, skeleton extraction, primitive selection) and this might, in turn, affect feature analysis and recognition. Consider, for example, the case of skeleton extraction of a region through medial axis transformation (MAT). The MAT of a region in a binary picture is determined with respect to its boundary. In a gray tone image, the boundaries are not well defined. Therefore, errors are more likely, if we compute the MAT from the hard-segmented version of the image.

Thus, it is convenient, natural and appropriate to avoid committing ourselves to a specific (hard) decision (e.g. segmentation/thresholding, edge detection and skeletonization), by allowing the segments or skeletons or contours to be fuzzy subsets of the image, the subsets being characterized by the possibility (degree) to which each pixel belongs to them. Similarly, for describing and interpreting ill-defined structural information in a pattern, it is natural to define primitives (line, corner, curve, etc.) and relations among them using labels of fuzzy sets. For example, primitives which do not lend themselves to precise definition may be defined in terms of arcs with varying grades of membership from 0 to 1 representing their degree of belonging to more than one class. The production rules of a grammar may similarly be fuzzified to account for the fuzziness (impreciseness) in physical relation among the primitives; thereby increasing the generative power of a grammar for syntactic recognition of a pattern.

We shall describe here a few methods of fuzzy segmentation (based on both gray level thresholding and pixel classification) and edge detection using global and/or local information of an image space. We mention here that the result of segmentation should be fuzzy subsets rather than ordinary subsets was first suggested by Prewitt.[150]

7.1. *Fuzzy thresholding*

Different histogram thresholding techniques in providing both fuzzy and non-fuzzy segmented versions by minimizing the grayness ambiguity (global entropy, index of fuzziness, index of crispness) and geometrical ambiguity (fuzzy compactness) of an image have been described in references (126, 131). These algorithms use different S-type membership functions (Fig. 1(a)) to define fuzzy "object regions" and then select the one which is associated with the minimum (optimum) value of the aforesaid ambiguity measures. The optimum membership function thus obtained enhances the object from background and denotes the membership values of the pixels for the fuzzy object region. Note that the cross-over point (the point with membership value of 0.5; in Fig. 1(a) b is the cross-over point) of the optimum membership function may be considered a threshold for crisp segmentation. Its extension to multithresholding has also been made. An S-type membership function can be asymmetric also. The mathematical framework of the algorithm including the selection of S functions, its bandwidth and bounds has been established by Murthy and Pal.[132] Many other measures of image ambiguity, e.g. fuzzy correlation,[146] index of area coverage,[144] adjacency[144,147] may similarly be used. Pal and Pal[133,134,151] introduced a measure called higher order entropy of a fuzzy set and applied it in a similar way to the object extraction problem using an adaptive membership function.

The problem of determining the appropriate membership function in image processing drew the attention of many researchers. Reconsider the problem of gray level thresholding using S functions. If there is a difference in opinion in defining an S function (i.e. instead of a single membership function, we have a set of monotonically non-decreasing functions), the concept of spectral fuzzy sets[148] can be used to provide soft decisions (a set of thresholds along with their certainty values) by giving due respect to all opinions. In making such a decision, the algorithm minimizes differences in opinions in addition to the ambiguity measures mentioned earlier; thereby managing the uncertainty. The bounds for S-type functions have been defined based on the properties of fuzzy correlation[132] so that any function lying in the bounds would give satisfactory segmentation results. It, therefore, demonstrates the flexibility of fuzzy algorithms. Xie and Bedrosian[145] have also made attempts in determining membership functions for gray level images.

7.2. *Fuzzy clustering*

The fuzzy c-means (FCM) clustering algorithm[122] has has also been used in image segmentation.[137,138,177] The fuzzy c-means algorithm uses an iterative optimization of an objective function based on a weighted similarity measure between the pixels in the image and each of the c-cluster centers. A local extremum of this objective function indicates an optimal clustering of the input data. The objective function that is minimized is given by

$$W_m(U, V) = \sum_{k=1}^{n} \sum_{i=1}^{c} (\mu_{ik})^m (d_{ik})^2 \qquad (6)$$

where μ_{ik} is the fuzzy membership value of the kth pixel in the ith cluster, d_{ik} is any inner product induced norm metric, m controls the nature of clustering with hard clustering at $m = 1$ and increasingly fuzzier clustering at higher values of m, V is the set of c-cluster centers and U is the fuzzy c-partition of the image. Trivedi and Bezdek[137] proposed a fuzzy set theoretic image segmentation algorithm for aerial images. The method is based upon region growing principles using a pyramid data structure. The algorithm is hierarchical in nature. Segmentation of the image at a particular processing level is done by the FCM algorithm. In a multilevel segmentation experiment, level i regions are considered homogeneous when image elements have largest cluster membership values of greater than a prescribed threshold. If the homogeneity test fails, regions are split to form the next level regions which are again subjected to the FCM algorithm. This algorithm is a region splitting algorithm, where the acceptance of a region is determined by fuzzy membership values to different regions. Hall *et al.*[138] segmented magnetic resonance brain images using the unsupervised fuzzy c-means and also by a supervised computational network—a dynamic multilayered perceptron trained with the cascade correlation learning algorithm. The different aspects of both approaches and their utility for the diagnostic process have been discussed. However, computational complexity of fuzzy c-mean is too high to apply it for real time application of MRI segmentation. Cannon *et al.*[139] suggested an approximate version of the algorithm that reduces the computational overhead. One of the advantages in using fuzzy clustering algorithms is that one can dynamically select the appropriate number of clusters depending on the strength of memberships across clusters.[135] Keller and Carpenter[140] used a modified version of FCM for image segmentation. The cluster centers are updated using the FCM formula but new membership values for each point are calculated using an S-type function based on the feature value of each point and the fuzzy means. They[140] also proposed region growing and relaxation algorithms based on membership values.

Backer[135] developed a very general clustering strategy which has been applied to different types of data including images. The set of samples $\mathscr{A}$ is first partitioned into c (number of classes) disjoint sets as an initial guess of the desired partition. Then a membership function is assigned to those initialized clusters according to some "point to point subset affinity" mechanism for all points in $\mathscr{A}$. In fact, he suggested a number of affinity mechanisms based on the distance concept, the neighborhood concept, and the probabilistic concept. Updating of the partitions (repartition, reclassification) is then done under the guidance of some criterion function which characterizes the partition. Three different types of criterion functions, based on measures of fuzziness, inter fuzzy set distance, and measure of fuzzy similarity have been considered there. If changes occur in the earlier step, the process of assigning membership function and updating is repeated; otherwise, the algorithm terminates. Fuzzy measures and fuzzy integral have also been used for image segmentation including multispectral images.[141–143]

We emphasize here that these developments are mainly based on the applications of fuzzy operators, properties and mathematics. Segmentation based on the theory of approximate reasoning (i.e. based on "if-then" rules) should constitute a field of research in the near future.

7.3. *Fuzzy edge detection*

Pal and King[126] used a non-symmetrical membership function **G** to get the fuzzy property plane from the intensity plane. The **G** is defined as

$$G(f(x,y)) = (1 + |f^* - f(x,y)|/F_d)^{-F_e} \qquad (7)$$

where f^* is a reference level, F_e and F_d are the exponential and denominational fuzzifiers, respectively. If $f^* = f_{max}$, the maximum gray level, then **G** approximates the standard S function[130] of Zadeh and when f^* is equal to some other level, $0 < f^* < f_{max}$, it approximates the standard π function[130] of Zadeh shown in Fig. 1(b). The **G** functions under the above cases are denoted by $\mathbf{G}_S$ and $\mathbf{G}_\pi$, respectively. They used these $\mathbf{G}_S$ and $\mathbf{G}_\pi$ functions in conjunction with an intensification operator INT to intensify the contrast in the image. (Note that the non-fuzzy thresholds obtained automatically from fuzzy segmentation techniques can be used in defining $\mathbf{G}_S$ and $\mathbf{G}_\pi$.) Finally, an inverse transformation is applied to get the enhanced spatial domain image. Edges of this enhanced image can then be easily found with any spatial domain technique. Edge detection operators based on max and min operations are available in references (152–154). In references (133, 151) the entropy of a fuzzy set defined by an adaptive membership function, over a neighborhood of a pixel (x, y) is used as a measure of edginess at (x, y). The use of an adaptive membership function makes the detection algorithm robust. The framework of the algorithm is quite general and works with any measure of ambiguity (fuzziness). In the next section we compare a few of the segmentation techniques.

8. COMPARISON OF SOME METHODS

We have discussed several methods of segmentation but so far not shown any results. In this section, for the sake of completeness and illustration, we consider segmentation results produced by a few techniques. We implemented six histogram based methods (methods of Otsu,[52] Pun,[65] Kapur *et al.*,[49] Kittler and Illingworth,[63] Pal and Bhandari,[71] and Pal and Pal[8]), and two iterative pixel classification methods (relaxation[94] and MAP estimate of a scene using NN[162]). Reference (8) has several algorithms, we have implemented only the maximum entropy algorithm (that uses Poisson distributions). Since the first six algorithms are not suitable for highly noisy images,

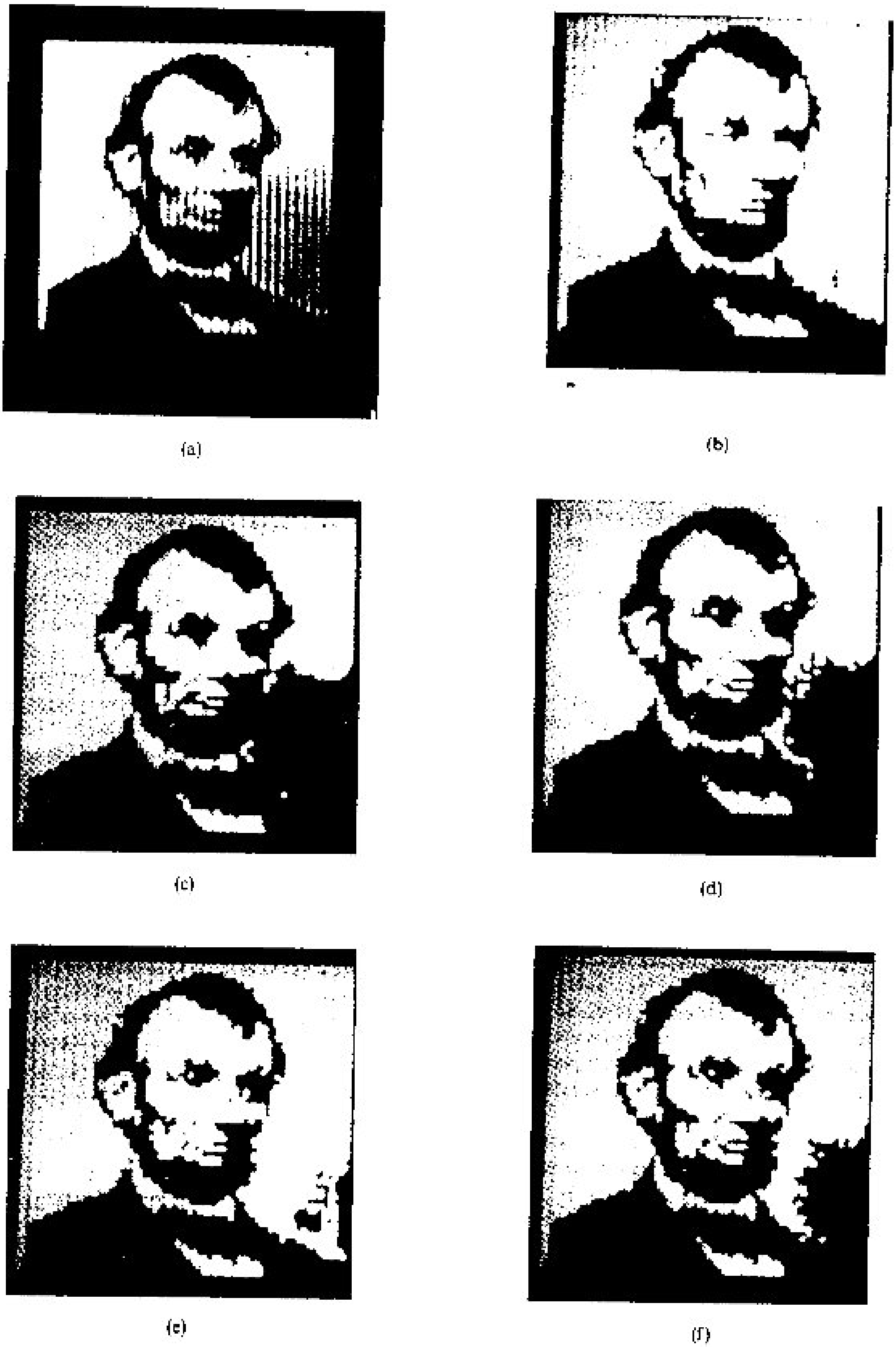

Fig. 2. Image of Abraham Lincoln: (a) input; (b) output by algorithm of Pal and Bhandari;[71] (c) output by algorithm of Pun;[65] (d) output by algorithm of Kapur *et al.*;[49] (e) output by algorithm of Pal and Pal;[6] (f) output by algorithm of Otsu.[52]

while the last two are, two input images have been used. Figure 2(a) is an image of Abraham Lincoln and Fig. 3(a) is a synthetic noisy image with geometric objects. Needless to say the first six algorithms fail for this image. We have applied the first six thresholding algorithms on Fig. 2(a) and the last two algorithms on Fig. 3(a). Figures 2(b)–(f) represent different segmented images produced by different thresholding methods

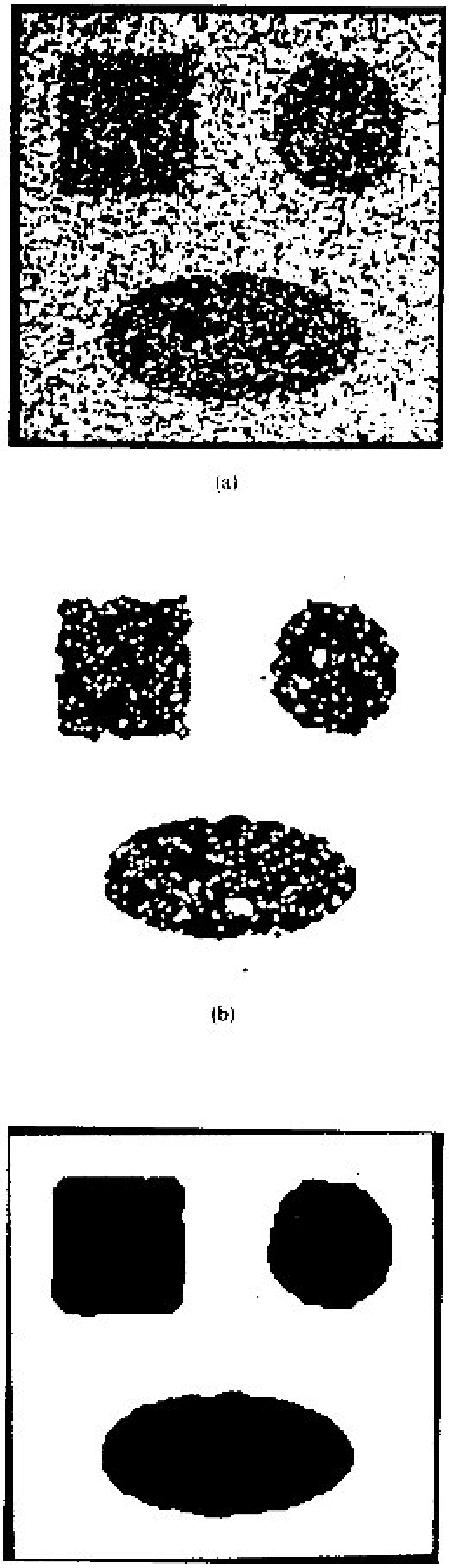

Fig. 3. Noisy image of geometric objects: (a) input; (b) output by the relaxation algorithm;[94] (c) output by neural net method.[112]

for the image of Lincoln. We tried different initial approximate thresholds for the method of Kittler and Illingworth, but it failed completely to produce any meaningful threshold. The algorithm either does not converge or converges towards the end of the gray scale. On the other hand, the algorithm in reference (71) which essentially uses the concept of Kittler and Illingworth but with Poisson distributions to model the histogram, produces a good thresholded image (Fig. 2(b)). The segmentation results produced by the methods of Pun[65] and Kapur *et al.*[49] are displayed in Figs 2(c) and (d), respectively. Both of these methods are based on entropy maximization. The parametric method in reference (8) which uses the Poisson distribution based model (derived considering the image formation process) produces Fig. 2(e). The result produced by the method of Otsu (Fig. 2(f)) is better than Figs 2(c) and (d); but this result is also not as good as those produced by the Poisson distribution based methods. For the noisy image (Fig. 3(a)) the probabilistic relaxation method produces a reasonably good segmentation (Fig. 3(b)). The neural network based method[167] which uses the GRF to model the noisy scene and then uses a network to obtain the MAP estimate of the scene (the segmented image) also produces a good segmentation (Fig. 3(c)) of Fig. 3(a).

9. OBJECTIVE EVALUATION OF SEGMENTATION RESULTS

We have already discussed several methods of image segmentation. It is known that no method is equally good for all images and all methods are not good for a particular type of images. Here an important problem remains to be discussed, how to make a quantitative evaluation of segmentation results. Such a quantitative measure would be quite useful for vision applications where automatic decisions are required. Also this will help to justify an algorithm. Unfortunately, a human being is the best judge to evaluate the output of any segmentation algorithm. However, some attempts have already been made for the quantitative evaluation. Levine and Nazif[178] used a two dimensional distance measure that quantifies the difference between two segmented images, one proposed by a human being the other by an algorithm. Later on they[95] defined another set of performance parameters such as region uniformity, region contrast, line contrast, etc. These measures have also been used for quantitative evaluation of segmentation algorithms. Lim and Lee[172] attempted to do this by computing the probability of error between the manually segmented image and the segmentation result. Pal and Bhandari[68] used the higher order local entropy as an index to measure the quality of the output. They also suggested the use of symmetric divergence between two probability distributions, one for the output generated by an algorithm and the other for the manually segmented image. The correlation measure[61] between the original image and the segmented one has also been used for the purpose of quantitative evaluation.[98] We have already mentioned that a human being is the ultimate judge to make an evaluation of the result. However, one can use a vector of such measures for objective evaluation. For example, if for some segmented image, the correlation, uniformity, and entropy are all high and divergence is low then one can consider the output to be good.

10. CONCLUSION

This paper reviews and summarizes some existing methods of segmentation. The literature is not so much rich on color image segmentation. Enough scope also exists for the fuzzy set theoretic approaches to segmentation. Neural network model based algorithms seem to be very promising as they can generate output in real time. Moreover, these algorithms are robust also. Selection of an appropriate segmentation technique largely depends on the type of images and application areas. An interesting area of investigation is to find methods of objective evaluation of segmentation results. It is very difficult to find a single quantitative index for this purpose because such an index should take into account many factors like homogeneity, contrast, compactness, continuity, psycho-visual perception, etc. Possibly the human being is the best judge for this. However, it may be possible to have a small vector of attributes which can be used for objective evaluation of results.

Acknowledgment—The authors gratefully acknowledge the referees for their thoughtful comments that have helped to improve the paper significantly.

REFERENCES

1. M. Osteaux, K. De Meirleir and M. Shahabpour, eds, *Magnetic Resonance Imaging and Spectroscopy in Sports Medicine*. Springer, Berlin (1991).
2. D. Marr, *Vision*. Freeman, San Francisco (1982).
3. A. Rosenfeld and A. C. Kak, *Digital Picture Processing*. Academic Press, New York (1982).
4. E. L. Hall, *Computer Image Processing and Recognition*. Academic Press, New York (1979).
5. R. C. Gonzalez and P. Wintz, *Digital Image Processing*. Addison-Wesley, Reading, Massachusetts (1987).
6. S. L. Horowitz and T. Pavlidis, Picture segmentation by directed split and merge procedure, *Proc. 2nd Int. Joint Conf. Pattern Recognition*, pp. 424–433 (1974).
7. A. Perez and R. C. Gonzalez, An iterative thresholding algorithm for image segmentation, *IEEE Pattern Analysis Mach. Intell.* **9**, 742–751 (1987).
8. N. R. Pal and S. K. Pal, Image model, Poisson distribution and object extraction, *Int. J. Pattern Recognition Artif. Intell.* **5**, 459–483 (1991).
9. P. C. Besl and R. C. Jain, Segmentation using variable surface fitting, *IEEE Pattern Analysis Mach. Intell.* **10**, 167–192 (1988).
10. D. Marr and E. Hildreth, Theory of edge detection, *Proc. R. Soc. Lond. B* **207**, 187–217 (1980).
11. C. K. Chow and T. Kaneko, Automatic boundary detection of the left-ventricle from cineangiograms, *Comput. Biomed. Res.* **5**, 388–410 (1972).
12. Y. Nakagawa and A. Rosenfeld, Some experiments on variable thresholding, *Pattern Recognition* **11**, 191–204 (1979).
13. S. D. Yanowitz and A. M. Bruckstein, A new method for image segmentation, *Comput. Vision Graphics Image Process.* **46**, 82–95 (1989).
14. N. Yokoya and M. D. Levine, Range image segmentation based on differential geometry: a hybrid approach, *IEEE Trans. Pattern Analysis Mach. Intell.* **11**(6), 643–694 (1989).
15. E. H. Al-Hujazi and A. K. Sood, Integration of edge and region based techniques for range image segmentation, *SPIE*, Vol. 1381, *Intelligent Robots and Comput. Vision—IX: Algorithms and Techniques*, pp. 589–599 (1990).
16. S. K. Pal and N. R. Pal, Segmentation based on measures of contrast, homogeneity, and region size, *IEEE Trans. Syst. Man Cybern.* **17**, 857–868 (1987).
17. M. K. Kundu and S. K. Pal, Thresholding for edge detection using human psychovisual phenomena, *Pattern Recognition Lett.* **4**, 433–441 (1986).
18. K. S. Fu and J. K. Mui, A survey on image segmentation, *Pattern Recognition* **13**, 3–16 (1981).
19. T. Pavlidis, *Structural Pattern Recognition*. Springer, New York (1977).
20. W. Doyle, Operations useful for similarity-invariant pattern recognition, *J. Ass. Comput. Mach.* **9**, 259–267 (1962).
21. S. Watanabe and the CYBEST Group, An automated apparatus for cancer preprocessing: CYBEST, *Comput. Graphics Image Process.* **3**, 350–358 (1974).
22. R. B. Ohlander, Analysis of natural scenes, Ph.D. Dissertation, Department of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania, April (1975).
23. J. F. Brenner *et al.*, Scene segmentation techniques for the analysis of routine bone marrow smears from acute lymphoblastic leukemia patients, *J. Histochem. Cytochem.* **25**, 601–613 (1977).
24. F. Tomita, M. Yachida and S. Tsuji, Detection of homogeneous regions by structural analysis, *Proc. 3rd IJCAI*, pp. 564–571 (1973).
25. S. Tsuji and F. Tomita, A structural analyzer for a class of textures, *Comput. Graphics Image Process.* **2**, 216–231 (1973).
26. I. T. Young and I. L. Paskowitz, Localization of cellular structures, *IEEE Trans. Biomed. Engng* **BME-22**, 35–40 (1975).
27. J. Keng, Syntactic algorithms for image segmentation and a special computer architecture for image processing, Ph.D. Thesis, Purdue University, West Lafayette, Indiana, December (1977).
28. R. M. Haralick and I. Dinstein, A spatial clustering procedure for multi-image data, *IEEE Trans. Circuits Syst.* **CAS-22**, 440–450 (1975).
29. R. K. Aggarwal and J. W. Bacus, A multi-spectral approach for scene analysis of cervical cytology smears, *J. Histochem. Cytochem.* **25**, 668–680 (1977).
30. J. R. Yoo and T. S. Huang, Image segmentation by unsupervised clustering and its applications, TR-EE 78-19, Purdue University, West Lafayette, Indiana (1978).
31. A. Rosenfeld and M. Thurston, Edge and curve detection for visual scene analysis, *IEEE Trans. Comput.* **C-20**, 562–569 (1971).
32. A. Rosenfeld, M. Thurston and Y. Lee, Edge and curve detection further experiments, *IEEE Trans. Comput.* **C-21**, 677–715 (1972).
33. M. Hueckel, An operator which locates edges in digital pictures, *J. Ass. Comput. Mach.* **18**, 113–125 (1971).
34. R. M. Haralick and L. G. Shapiro, Survey, image segmentation techniques, *Comput. Vision Graphics Image Process.* **29**, 100–132 (1985).
35. R. M. Haralick, Edge and region analysis for digital image data, *Comput. Graphics Image Process.* **12**, 60–73 (1980); in *Image Modeling*, A. Rosenfeld, ed., pp. 171–184. Academic Press, New York (1981).

36. R. M. Haralick, Zero-crossing of second directional derivative edge operator. *Proc. Soc. Photo-optical Instrumentation Engrs Tech. Symp. East*, Arlington, Virginia, Vol. 336, 3–7 May (1982).
37. Y. Yakimovsky, Boundary and object detection in real world image, *J. Ass. Comput. Mach.* **23**, 599–618 (1976).
38. M. D. Levine and S. I. Shaheen, A modular computer vision system for picture segmentation and interpretation, *IEEE Trans. Pattern Analysis Mach. Intell.* **3**, 540–556 (1981).
39. T. C. Pong, L. G. Shapiro, L. T. Watson and R. M. Haralick, Experiments in segmentation using a facet model region grower, *Comput. Vision. Graphics Image Process.* **25** (1984).
40. T. Pavlidis, Segmentation of pictures and maps through functional approximations. *Comput. Graphics Image Process.* **1**, 360–372 (1972).
41. G. Nagy and J. Tolaba, Nonsupervised crop classification through airborne multispectral observations, *IBM J. Res. Dev.* **16**, 138–153 (1972).
42. P. K. Sahoo, S. Soltani, A. K. C. Wong and Y. C. Chen, A survey of thresholding techniques, *Comput. Vision Graphics Image Process.* **41**, 233–260 (1988).
43. S. Boukharouba, J. M. Rebordao and P. L. Wendel, An amplitude sedimentation method based on the distribution function of an image, *Comput. Vision Graphics Image Process.* **29**, 47–59 (1985).
44. S. Wang and R. M. Haralick, Automatic multithreshold selection, *Comput. Vision Graphics Image Process.* **25**, 46–67 (1984).
45. R. Kohler, A segmentation system based on thresholding, *Comput. Graphics Image Process.* **15**, 319–338 (1981).
46. T. Taxt, P. J. Flynn and A. K. Jain, Segmentation of document images, *IEEE Trans. Pattern Analysis Mach. Intell.* **11**(12), 1322–1329 (1989).
47. N. R. Pal and S. K. Pal, Object–background segmentation using new definitions of entropy, *IEE Proc., Pt. E* **136**, 284–295 (1989).
48. N. R. Pal and S. K. Pal, Entropic thresholding, *Signal Process.* **16**, 97–108 (1989).
49. J. N. Kapur, P. K. Sahoo and A. K. C. Wong, A new method for gray level picture thresholding using the entropy of histogram, *Comput. Graphics Vision Image Process.* **29**, 273–285 (1985).
50. L. S. Davis, A. Rosenfeld and J. S. Weszka, Region extraction by averaging and thresholding, *IEEE Trans. Syst. Man, Cybern.* **5**, 383–388 (1975).
51. R. Nevatia, Locating object boundaries in textured environments, *IEEE Trans. Syst. Man Cybern.* **25**, 1170–1175 (1976).
52. N. Otsu, A threshold selection method from gray level histograms, *IEEE Trans. Syst. Man Cybern.* **9**, 62–66 (1979).
53. N. Ahuja and A. Rosenfeld, A note on the use of second order gray level statistics for threshold selection. *IEEE Trans. Syst. Man Cybern.* **8**, 895–898 (1978).
54. P. A. Dondes and A. Rosenfeld, Pixel classification on gray level and local "Busyness", *IEEE Trans. Pattern Analysis Mach. Intell.* **4**, 79–84 (1982).
55. A. Y. Wu, T. Hong and A. Rosenfeld, Threshold selection using quadtrees, *IEEE Trans. Pattern Analysis Mach. Intell.* **4**, 90–94 (1982).
56. A. K. Jain, S. P. Smith and L. E. Backer, Segmentation of muscle cell pictures: a preliminary study, *IEEE Trans. Pattern Analysis Mach. Intell.* **2**, 232–242 (1980).
57. J. Mantas, Methodologies in pattern recognition and image analysis: a brief survey, *Pattern Recognition* **20**, 1–6 (1987).
58. K. V. Mardia and T. J. Hainsworth, A spatial thresholding method for image segmentation, *IEEE Pattern Analysis Mach. Intell.* **10**, 919–927 (1988).
59. A Rosenfeld and R. C. Smith, Thresholding using relaxation, *IEEE Pattern Analysis Mach. Intell.* **3**, 598–606 (1981).
60. N. R. Pal and S. K. Pal, Entropy: a new definition and its applications, *IEEE Trans. Syst. Man Cybern.* **21**(5), 1260–1270 (1991).
61. A. B Brink, Gray level thresholding of images using a correlation criterion, *Pattern Recognition Lett.* **9**, 335–341 (1989).
62. J. Kittler and J. Illingworth, Minimum error thresholding, *Pattern Recognition* **19**, 41–47 (1986).
63. T. W. Ridler and S. Calvard, Picture thresholding using an iterative selection method, *IEEE Trans. Syst. Man Cybern.* **SMC-8**, 630–632 (1978).
64. T. Pun, Entropic thresholding: a new approach, *Comput. Graphics Image Process.* **16**, 210–239 (1981).
65. T. Pun, A new method for gray level picture thresholding using the entropy of the histogram, *Signal Process.* **2**, 223–237 (1980).
66. A. Rosenfeld, Survey, image analysis and computer vision, *Comput. Vision Graphics Image Process.* **46**, 196–250 (1989).
67. S. K. Pal and N. R. Pal, Segmentation using contrast and homogeneity measures, *Pattern Recognition Lett.* **5**, 293–304 (1987).
68. N. R. Pal and D. Bhandari, Object background classification: some new techniques, *Signal Process.* **33** (1993).
69. A. K. C. Wong and P. K. Sahoo, A gray level threshold selection method based on maximum entropy principle, *IEEE Trans. Syst. Man Cybern.* **19**, 866–871 (1989).
70. S. K. Pal and N. R. Pal, Object background classification using a new definition of entropy, *Proc. IEEE Int. Conf. Syst. Man Cybern.*, China, Vol. 2, pp. 773–776 (1988).
71. N. R. Pal and D. Bhandari, On object–background classification, *Int. J. Syst. Sci.* **23**(11), 1903–1920 (1992).
72. J. S. Weszka, Survey of threshold selection techniques, *Comput. Graphics Image Process.* **7**, 259–265 (1978).
73. A. Rosenfeld and L. S. Davis, Iterative histogram modification, *IEEE Trans. Syst. Man Cybern.* **8**, 300–302 (1978).
74. S. Peleg, Iterative histogram modification, *IEEE Trans. Syst. Man Cybern.* **8**, 555–556 (1978).
75. J. S. Weszka and A. Rosenfeld, Threshold evaluation techniques, *IEEE Trans. Syst. Man Cybern.* **8**, 622–629 (1978).
76. F. Deravi and S. K. Pal, Gray level thresholding using second-order statistics, *Pattern Recognition Lett.* **1**, 417–422 (1983).
77. B. Chanda, B. B. Chaudhuri and D. Dutta Majumder, On image enhancement and threshold selection using the gray level co-occurrence matrix, *Pattern Recognition Lett.* **3**, 243–251 (1985).
78. A. Rosenfeld and L. S. Davis, Image segmentation and image models, *Proc. IEEE* **67**, 764–772 (1979).
79. J. S. Weszka and A. Rosenfeld, Histogram modification for threshold selection, *IEEE Trans. Syst. Man Cybern.* **9**, 38–52 (1979).
80. A. S. Abutaleb, Automatic thresholding of gray level pictures using two-dimensional entropy, *Comput. Vision Graphics Image Process.* **47**, 22–32 (1989).
81. D. E. Lloyd, Automatic target classification using moment invariants of image shapes, Farnborough, U.K., Rep. RAE IDN AW126, December (1985).
82. R. Hoffman and A. K. Jain, Segmentation and classification of range images, *IEEE Trans. Pattern Analysis Mach. Intell.* **9**, 608–620 (1987).
83. R. O. Duda, D. Nitzan and P. Barrett, Use of range and reflectance data to find planar surface region, *IEEE Trans. Pattern Analysis Mach. Intell.* **1**(3), 259–271 (1979).
84. R. D. Rimey and F. S. Cohen, A maximum-likelihood approach to segmenting range data, *IEEE J. Robotics Automn* **4**(3), 277–286 (1988).
85. S. Baronti, A. Casini and F. Lotti, Variable pyramid structures for image segmentation, *Comput. Vision Graphics Image Process.* **49**, 346–356 (1990).
86. J. M. Beaulieu and M. Goldberg, Hierarchy in picture segmentation: a stepwise optimization approach, *IEEE Trans. Pattern Analysis Mach. Intell.* **11**, 150–163 (1989).

87. M. Pietikainen, A. Rosenfeld and I. Walter, Split and link algorithm for image segmentation, *Pattern Recognition* **15**, 287–298 (1982).
88. M. Amadasun and R. A. King, Low level segmentation of multispectral images via agglomerative clustering of uniform neighbours, *Pattern Recognition* **21**, 261–268 (1988).
89. B. Bhanu and B. A. Parvin, Segmentation of natural scene, *Pattern Recognition* **20**, 487–496 (1987).
90. S. Basu and K. S. Fu, Image segmentation by syntactic method, *Pattern Recognition* **20**, 33–44 (1987).
91. T. Hong, K. A. Narayanan, S. Peleg, A. Rosenfeld and T. Silberberg, Image smoothing and segmentation by multiresolution pixel linking: further experiments and extensions, *IEEE Syst. Man Cybern.* **12**, 611–622 (1982).
92. H. S. Don and K. S. Fu, A syntactic method for image segmentation and object recognition, *Pattern Recognition* **18**, 73–87 (1985).
93. S. Basu, Image segmentation by semantic method, *Pattern Recognition* **20**, 497–511 (1987).
94. A. Rosenfeld, R. A. Hummel and S. W. Zucker, Scene labeling by relaxation operations, *IEEE Trans. Syst. Man Cybern.* **6**, 420–433 (1976).
95. M. D. Levine and A. M. Nazif, Dynamic measurement of computer generated image segmentation, *IEEE Trans. Pattern Analysis Mach. Intell.* **7**, 155–164 (1985).
96. S. Peleg, A new probabilistic relaxation scheme, *IEEE Trans. Pattern Analysis Mach. Intell.* **2**, 362–369 (1980).
97. Y. T. Zhou, V. Venkateswar and R. Chellappa, Edge detection and linear feature extraction using a 2-D random field model, *IEEE Trans. Pattern Analysis Mach. Intell.* **11**, 84–95 (1989).
98. J. Kittler, J. Eggleton, J. Illingworth and K. Paler, An average edge detector, *Pattern Recognition Lett.* **6**, 27–32 (1987).
99. C. J. Jacobus and R. T. Chein, Two new edge detectors, *IEEE Trans. Pattern Analysis Mach. Intell.* **3**, 581–592 (1981).
100. W. H. H. J. Lunscher and M. P. Beddoes, Optimal edge detector design I: parameter selection and noise effects, *IEEE Trans. Pattern Analysis Mach. Intell.* **8**, 164–177 (1986).
101. W. H. H. J. Lunscher and M. P. Beddoes, Optimal edge detector design II: coefficient quantization, *IEEE Trans. Pattern Analysis Mach. Intell.* **8**, 178–187 (1986).
102. R. M. Haralick, Digital step edges from zero crossing of second directional derivatives, *IEEE Trans. Pattern Analysis Mach. Intell.* **6**, 58–68 (1984).
103. T. Peli and D. Malah, A study of edge detection algorithms, *Comput. Graphics Image Process.* **20**, 1–21 (1982).
104. A. Ikonomopoulos, An approach to edge detection based on the direction of edge elements, *Comput. Graphics Image Process.* **19**, 179–195 (1982).
105. J. F. Canny, A computational approach to edge detection, *IEEE Trans. Pattern Analysis. Mach. Intell.* **8**, 679–698 (1986).
106. Y. Lu and R. C. Jain, Behavior of edges in scale space, *IEEE Trans. Pattern Analysis Mach. Intell.* **11**, 337–356 (1989).
107. K. Paler and J. Kittler, Gray level edge thinning: a new method, *Pattern Recognition Lett.* **1**, 409–416 (1983).
108. J. F. Haddon, Generalized threshold selection for edge detection, *Pattern Recognition* **21**, 195–203 (1988).
109. L. S. Davis, A survey of edge detection techniques, *Comput. Graphics Image Process.* **4**, 248–270 (1975).
110. M. Kelly, Edge detection by computer using planning, *Machine Intelligence*, Vol. VI, pp. 397–409. Edinburgh University Press, Edinburgh (1971).
111. Y. P. Chien and K. S. Fu, Preprocessing and feature extraction of picture patterns, TR-EE 74-20, Purdue University, West Lafayette, Indiana (1974).
112. M. Gokmen and C. C. Li, Edge detection with iteratively refined regularization, *Proc. 10th ICPR*, pp. 690–693 (1990).
113. S. E. Reichenbach, S. K. Park and R. A. Gartenberg, Optimal, small kernels for edge detection, *Proc. 10th ICPR*, pp. 57–63 (1990).
114. A. K. Jain, Advances in mathematical models for image processing, *Proc. IEEE* **69**, 502–528 (1981).
115. F. R. Hansen and H. Elliott, Image segmentation using simple Markov random field models, *Comput. Graphics Image Process.* **20**, 101–132 (1982).
116. H. Elliott, H. Derin, R. Cristi and D. Geman, Application of the Gibbs distribution to image segmentation, *Proc. IEEE Int. Conf. Acoust. Speech. Signal Processing*, San Diego, California, March (1984).
117. H. Derin, H. Elliott, R. Cristi and D. Geman, Bayes smoothing algorithms for segmentation of binary images modeled by Markov random fields, *IEEE Trans. Pattern Analysis Mach. Intell.* **6**, 707–720 (1984).
118. S. Geman and D. Geman, Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images, *IEEE Trans. Pattern Analysis Mach. Intell.* **6**, 721–741 (1984).
119. F. R. Hansen, Application of Markov Random Field models to the design and analysis of image segmentation algorithm, M.S. Thesis, Colorado State University (1981).
120. M. Asker and H. Derin, A recursive algorithm for the Bayes solution of the smoothing problem, *IEEE Trans. Automat. Contr.* **AC-26**, 558–561 (1981).
121. H. Derin and H. Elliott, Modelling and segmentation of noisy and textured images using Gibbs Random Fields, *IEEE Trans. Pattern Analysis Mach. Intell.* **9**(1), 39–55 (1987).
122. J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York (1981).
123. A. Kandel, *Fuzzy Techniques in Pattern Recognition*. Wiley, New York (1982).
124. S. K. Pal and D. Dutta Majumder, *Fuzzy Mathematical Approach to Pattern Recognition*. Wiley (Halsted), New York (1986).
125. J. C. Bezdek and S. K. Pal, *Fuzzy Models for Pattern Recognition: Methods that Search for Structures in Data*. IEEE Press, New York (1992).
126. S. K. Pal and R. A. King, Image enhancement using fuzzy sets, *Electron. Lett.* **16**, 376–378 (1980).
127. S. K. Pal, R. A. King and A. A. Hashim, Automatic grey level thresholding through index of fuzziness and entropy, *Pattern Recognition Lett.* **1**, 141–146 (1983).
128. R. Jain and H. H. Nagel, Analysing a real world scene sequence using fuzziness, *Proc. IEEE Conf. Decision and Control*, New Orleans, pp. 1367–1372 (1977).
129. B. L. Deekshatulu, G. K. Rao and R. Krishnan, A quantitative index of information and a method of enhancement for digital imaging using fuzzy sets, *Proc. IEEE Int. Conf. Syst. Man Cybern.*, New Delhi, Vol. 2, pp. 1105–1109 (1983).
130. L. A. Zadeh, K. S. Fu, K. Tanaka and M. Shimura, eds, *Fuzzy Sets and their Applications to Cognitive and Decision Processes*. Academic Press, London (1975).
131. S. K. Pal and A. Rosenfeld, Image enhancement and thresholding by optimization of fuzzy compactness, *Pattern Recognition Lett.* **7**, 77–86 (1988).
132. C. A. Murthy and S. K. Pal, Fuzzy thresholding: mathematical framework, bound functions and weighted moving average technique, *Pattern Recognition Lett.* **11**, 197–206 (1990).
133. N. R. Pal, On image information measures and object extraction, Ph.D. Dissertation, Indian Statistical Institute, Calcutta (1990).
134. N. R. Pal and S. K. Pal, Higher order fuzzy entropy and hybrid entropy of a set, *Inf. Sci.* **61**, 211–231 (1992).
135. E. Backer, *Cluster Analysis by Optimal Decomposition of Induced Fuzzy Sets*. Delft University Press, Delft (1978).
136. A. Kandel, *Fuzzy Mathematical Techniques with Applications*. Addison-Wesley, Reading, Massachusetts (1986).
137. M. Trivedi and J. C. Bezdek, Low-level segmentation of

aerial images with fuzzy clustering, *IEEE Trans. Syst. Man Cybern.* **16**(4), 589–598 (1986).
138. L. O. Hall, A. M. Bensaid, L. P. Clarke, R. P. Velthuizen, M. Silbiger and J. C. Bezdek, A comparison of neural network and fuzzy clustering techniques in segmenting magnetic resonance images of the brain, *IEEE Trans. Neural Networks* 3(5), 672–682 (1992).
139. R. L. Cannon, R. L. Dave and J. C. Bezdek, Efficient implementation of fuzzy c-means clustering algorithms, *IEEE Trans. Pattern Analysis Mach. Intell.* **8**, 2 (1986).
140. J. M. Keller and C. L. Carpenter, Image segmentation in presence of uncertainty, *Int. J. Intell. Syst.* **5**, 193–208 (1990).
141. J. Keller, H. Qiu and H. Tahani, The fuzzy integral in image segmentation, *Proc. NAFIPS-86*, New Orleans, Louisiana, pp. 324–338, June (1986).
142. H. Qiu and J. Keller, Multispectral image segmentation using fuzzy techniques, *Proc. NAFIPS-87*, Purdue University, West Lafayette, Indiana, pp. 374–387 (1987).
143. B. Yan and J. Keller, Conditional fuzzy measures and image segmentation, *Proc. NAFIPS-91*, University of Missouri–Columbia, Missouri, pp. 32–36 (1991).
144. S. K. Pal and A. Ghosh, Index of area coverage of fuzzy image subsets and object extraction, *Pattern Recognition Lett.* **11**, 831–841 (1990).
145. W. X. Xie and S. D. Bedrosian, Experimentally derived fuzzy membership function for gray level images, *J. Franklin Inst.* **325**, 154–164 (1988).
146. S. K. Pal and A. Ghosh, Image segmentation using fuzzy correlation, *Inf. Sci.* **62**(3), 223–250 (1992).
147. A. Rosenfeld, The fuzzy geometry of image subsets, *Pattern Recognition Lett.* **2**, 311–317 (1984).
148. S. K. Pal and A. Dasgupta, Spectral fuzzy sets and soft thresholding, *Inf. Sci.* **65**, 65–97 (1992).
149. J. Keller, D. Subhangkasen and K. Unklesbay, Approximate reasoning for recognition in color image of beef steaks, *Int. J. General Syst.* **16**, 331–342 (1990).
150. J. M. S. Prewitt, Object enhancement and extraction, *Picture Processing and Psychopictorics*, B. S. Lipkin and A. Rosenfeld, eds, pp. 75–149. Academic Press, New York (1970).
151. N. R. Pal and S. K. Pal, Some information measures on fuzzy sets and their applications to image processing, *Proc. NACONECS-89*, pp. 94–96. Tata McGraw-Hill, New Delhi (1989).
152. V. Goetcherian, From binary to gray tone image processing using fuzzy logic concepts, *Pattern Recognition* **12**, 7–15 (1980).
153. S. K. Pal and R. A. King, On edge detection of X-ray images using fuzzy set, *IEEE Trans. Pattern Analysis Mach. Intell.* **PAMI-5**, 69–77 (1983).
154. Y. Nakagawa and A. Rosenfeld, A note on the use of local max and min operations in digital picture processing, *IEEE Trans. Syst. Man Cybern.* **SMC-8**, 632–635 (1978).
155. J. Hertz, A. Krogh and R. G. Palmer, *Introduction to the Theory of Neural Computation.* Addison-Wesley, New York (1991).
156. B. Kosko, *Neural Networks and Fuzzy Systems: a Dynamical Systems Approach to Machine Intelligence.* Prentice-Hall, Englewood Cliffs, New Jersey (1991).
157. T. Kohonen, *Self-organization and Associative Memory.* Springer, New York (1989).
158. Y. H. Pao, *Adaptive Pattern Recognition and Neural Networks.* Addison-Wesley, New York (1989).
159. W. E. Blanz and S. L. Gish, A connectionist classifier architecture applied to image segmentation, *Proc. 10th ICPR*, pp. 272–277 (1990).
160. N. Babaguchi, K. Yamada, K. Kise and T. Tezuka, Connectionist model binarization, *Proc. 10th ICPR*, pp. 51–56 (1990).
161. B. S. Manjunath, T. Simchony and R. Chellappa, Stochastic and deterministic network for texture segmentation, *IEEE Trans. Acoust. Speech Signal Process.* **38**, 1039–1049 (1990).
162. A. Ghosh, N. R. Pal and S. K. Pal, Image segmentation using a neural network, *Biol. Cybern.* **66**, 151–158 (1991).
163. J. Shah, Parameter estimation, multiscale representation and algorithms for energy-minimizing segmentation, *Proc. Int. Conf. Pattern Recognition*, pp. 815–819 (1990).
164. A. Ghosh, N. R. Pal and S. K. Pal, Self-organization for object extraction using multilayer neural networks and fuzziness measure, *IEEE Trans. Fuzzy Syst.* **1**(1), 54–68 (1993).
165. C. Cortes and J. A. Hertz, A network system for image segmentation, *Proc. Int. Joint Conf. on Neural Network*, Vol. 1, pp. 121–125 (1989).
166. A. Ghosh, N. R. Pal and S. K. Pal, Neural network, Gibbs distribution and object extraction, *Intelligent Robotics*, M. Vidyasagar and M. Trivedi, eds, pp. 95–106. Tata McGraw-Hill, New Delhi (1991).
167. A. Ghosh, N. R. Pal and S. K. Pal, Object extraction using a self organizing neural network, *Intelligent Robotics*, M. Vidyasagar and M. Trivedi, eds, pp. 686–697. Tata McGraw-Hill, New Delhi (1991).
168. C. T. Chen, E. C. Tsao and W. C. Lin, Medical image segmentation by a constraint satisfaction neural network, *IEEE Trans. Nucl. Sci.* **38**(2), 678–686 (1991).
169. A. Ghosh, N. R. Pal and S. K. Pal, Object background classification using Hopfield type neural network, *Int. J. Pattern Recognition Artif. Intell.* **6**(5), 989–1008 (1992).
170. W. J. Ho and C. F. Osborne, Texture segmentation using multilayered back propagation, *Proc. Int. Joint Conf Neural Net.* Singapore, pp. 981–986 (1991).
171. R. D. Overheim and D. L. Wagner, *Light and Color.* Wiley, New York (1982).
172. Y. W. Lim and S. U. Lee, On the color image segmentation algorithm based on the thresholding and fuzzy c-means techniques, *Pattern Recognition* **23**, 935–952 (1990).
173. S. K. Pal, Fuzzy sets and decision making in color image processing, *Artificial Intelligence and Applied Cybernetics*, A. Ghosal, ed., pp. 89–98. South Asian Publishers, New Delhi (1989).
174. Y. Ohta, T. Kanade and T. Sakai, Color information for region segmentation, *Comput. Graphics Image Process.* **13**, 224–241 (1980).
175. R. Ohlander, K. Price and D. R. Reddy, Picture segmentation using a recursive region splitting method, *Comput. Graphics Image Process.* **8**, 313–333 (1978).
176. W. X. Xie, An information measure for a color space, *Fuzzy Sets Syst.* **36**, 157–165 (1990).
177. T. L. Huntsberger, C. L. Jacobs and R. L. Cannon, Iterative fuzzy image segmentation, *Pattern Recognition* **18**, 131–138 (1985).
178. M. D. Levine and A. M. Nazif, An experimental rule based system for testing low level segmentation strategy, *Multicomputers and Image Processing: Algorithms and Programs*, K. Preston and L. Uhr, eds. Academic Press, New York (1982).

About the Author—NIKHIL R. PAL obtained the B.Sc. (Hons.) in physics and M.B.M. (operations research) in 1979 and 1982, respectively, from the University of Calcutta. He received the M.Tech. and Ph.D. in computer science from the Indian Statistical Institute, Calcutta, in 1984 and 1991, respectively. He was with Hindusthan Motors Ltd, India, from 1984 to 1985 and with Dunlop India Ltd, India, from 1985 to 1987. At present he is associated with Machine Intelligence Unit of the Indian Statistical Institute, Calcutta, and

currently visiting the University of West Florida, Pensacola, Florida. He was also a guest lecturer at the University of Calcutta. His research interests include image processing, pattern recognition, artificial intelligence, fuzzy sets and systems, uncertainty measures and neural networks. He is an associate editor of *International Journal of Approximate Reasoning*, and the *IEEE Transactions on Fuzzy Systems*. He is a member of IEEE, New York.

About the Author—SANKAR K. PAL is a Professor in Machine Intelligence Unit, Indian Statistical Institute, Calcutta. He obtained M.Tech. and Ph.D. in radiophysics and electronics from Calcutta University, and Ph.D./DIC in electrical engineering from Imperial College, London, in 1974, 1979 and 1982, respectively. He worked at Imperial College during 1979–83 (as a Commonwealth Scholar and an MRC Post-doctoral Fellow), at the University of California, Berkeley and the University of Maryland, College Park, during 1986–87 (as a Fulbright Fellow), and at the NASA Johnson Space Center, Houston, Texas, during 1990–1992 (as an NRC–NASA Senior Research Fellow). He served as a Professor-in-Charge, Physical and Earth Sciences Division, ISI, during 1988–90. He was also a guest teacher in computer science, Calcutta University, during 1983–86. His research interests include pattern recognition, image processing, neural nets, and fuzzy sets and systems. He is a co-author of the books *Fuzzy Mathematical Approach to Pattern Recognition*, Wiley (Halsted Press), New York, 1986 and *Fuzzy Models for Pattern Recognition: Methods that Search for Structures in Data*, IEEE Press, New York, 1992. He has about 150 research papers to his credit including ten in edited books and about 100 in international journals. He is listed in *Reference Asia, Asia's Who's Who of Men and Women of Achievements*, Vol. IV, 1989. He received the 1990 Shanti Swarup Bhatnagar Prize (which is the highest and most coveted award for a scientist in India) in engineering sciences for his contribution in the field of pattern recognition and image processing, and the 1993 Jawaharlal Nehru Fellowship. He is an Associate Editor of the *International Journal of Approximate Reasoning*, North-Holland, *IEEE Transactions on Fuzzy Systems*, a reviewer of the *Mathematical Reviews* (American Mathematical Society), a Fellow of IEEE, New York, and a Life Fellow of IETE, New Delhi.