

## **Fuzzy medial axis transformation (FMAT): Image representation, skeleton extraction and uncertainty management in computer vision**

SANKAR K PAL\* and LUI WANG

Software Technology Branch/PT4, National Aeronautics & Space Administration, Lyndon B Johnson Space Center, Houston, Texas 77058, USA

\*Present address: Electronics and Communication Sciences Unit, Indian Statistical Institute, Calcutta 700 035

**Abstract.** The problem of extraction of medial axis transformation of a gray image with reference to skeletonization, image representation and uncertainty management in a vision system has been addressed. The Fuzzy Medial Axis Transformation (FMAT) of a fuzzy set  $f$  is a set of fuzzy disks whose sup is  $f$ . Unfortunately, specifying the FMAT sometimes requires more storage space than specifying  $f$  itself. The present paper describes some techniques to improve the compact representation of FMAT; thereby making it practically useful to an image for its skeleton extraction and compact representation, for shape analysis and template matching, for representation and retrieval, for uncertainty management in recognition and for creating new images of various poses. The algorithms involve reduction of redundancy in FMAT, its approximation, and reduction of the searching spaces for its computation. Computational aspects for the convenience of writing an efficient program have been described. Some applications of the FMAT have also been mentioned.

**Keywords.** Fuzzy medial axis transformation (FMAT); fuzzy disk; fuzzy geometry; image processing/representation; skeleton extraction; uncertainty management.

### **1. Introduction**

The task of pattern recognition by a computer can be viewed as a transformation from the measurement space  $M$  to the feature space  $F$  and finally to the decision space  $D$ , i.e.,

$$M \rightarrow F \rightarrow D.$$

When the input pattern is a gray tone image, the measurement space involves some

important processing tasks such as enhancement, filtering, noise reduction, segmentation, contour extraction and skeleton extraction, in order to derive salient features from the image pattern. This is what is basically known as *image processing*. The ultimate aim is to use data contained in the image to enable the system to understand, recognize and interpret the processed information available from the image pattern. Such a complete image recognition/interpretation system is called a *vision system* which may be viewed as consisting of three levels, namely, low level, mid level and high level.

The relevance of fuzzy set theory in pattern recognition problems has adequately been addressed in the literature (Bezdek 1981; Kandel 1982; Pal & Dutta Majumder 1986; Bezdek & Pal 1992). It is seen that the concept of fuzzy sets can be used (i) 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, (ii) at the classification level in representing multi-class membership of an ambiguous pattern, and (iii) in providing an estimate (or a representation) of missing information in terms of membership values. In other words, fuzzy set theory may be incorporated in handling uncertainties (arising from deficiencies of information available from a situation; 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 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 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 (Marr 1982; Rosenfeld & Kak 1982; Gonzalez & Wintz 1987) 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) to make regions, features, primitives, properties, relations and interpretations crisp.

Since the regions in an image are not always crisply defined, uncertainty can arise within every phase of the 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 etc.) 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 problem of extracting the medial axis transformation plays a key role in the field of image processing, analysis and recognition because of the simplicity of image (and hence object) representation and skeleton extraction it allows. There has been extensive research done (Rosenfeld & Kak 1982, p. 191) in extracting the medial axis of a region and skeleton of elongated objects from a two-tone image. The medial axis transformation 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 likely (and hence further increase uncertainty in the system), if we compute the MAT from the aforesaid 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.

In this paper, we explain the role of fuzzy set theory to the problem of medial axis transformation of a gray tone image for its representation and skeleton extraction along with the management of uncertainties in these processes. The medial axis (MA) of a subset  $S$  of a binary picture is the set of those points in  $S$  whose distances from  $\bar{S}$  (complement of  $S$ ) are local maxima. The distance of a point  $P$  in  $S$  from  $\bar{S}$  is the length of a shortest path from  $P$  to  $\bar{S}$ . The MAT of  $S$  consists of the aforesaid local maxima points together with their distances from  $\bar{S}$ . It provides a compact (economical) representation of  $S$  so that  $S$  can be reconstructed from its MAT. The MA of  $S$  can be regarded as a generalized axis of its symmetry and constitutes also a kind of skeleton, if  $S$  is elongated. Note that the word "kind of" is used to mean that such a skeleton may be disconnected, since the MA is a union of the local maxima points.

From the aforesaid definitions it is seen that the MAT of a region in a binary picture is determined with respect to its boundary. As mentioned before, the boundaries among various regions in a gray tone image are *not* well defined. A few generalizations of MAT have been proposed in order to make them applicable to gray tone images. These include SPAN (Spatial Piecewise Approximation by Neighborhoods) (Ahuja *et al* 1978), GRAYMAT (Levi & Montanari 1970) and MMMAT (Peleg & Rosenfeld 1981).

SPAN provides an approximated representation of an image in terms of centres, radii and average gray level of homogeneous disks. GRAYMAT is based on the concept of gray-weighted distance of a point from the background and therefore it requires an image to be segmented into zeros (background) and non-zeros (object). MMMAT is based on iterative min and max operations, and does not require the picture to be segmented. However, the number of iterations depends on the application and the border effects become a serious problem when the local min operation is iterated several times. All these techniques reduce to the MAT in the case of a two-tone image. They provide a reasonable gray medial axis (and skeleton) for the darker pixels in the case of gray tone image but the original (input) gray image, in any case, cannot be reconstructed from its MMMAT or SPAN or GRAYMAT values. This property of image reconstruction is supposed to be an important characteristic of the MAT representation. Furthermore, the requirement of an image to be segmented into zero (background) and non-zero (object) needs the selection of the threshold to be made judiciously because the resulting skeleton is dependent on that selection.

There have also been some attempts made for extracting a gray skeleton of an image, but without using the concept of MAT. These include fuzzy skeletonization technique (Pal 1989) and ridge seeking method (Salari & Siy 1984). Fuzzy skeletonization of Pal (1989) needs an initial fuzzy segmentation of the image space so that the membership of a pixel for the subset skeleton can be computed with respect to the  $\varepsilon$  edge (edge points of object after which its class membership value is less than or equal to  $\varepsilon$ ,  $0 \leq \varepsilon \leq 1$ ) of the object region. The ridge seeking method also requires that the pixels with gray value below a certain threshold be set to zero. The skeletons produced by these methods do not depend much on the boundary selection. Since these techniques are not formulated from the point of MAT construction, the question of reconstruction of original gray image did not get any attention.

A fuzzy medial axis transformation (FMAT) based on the concept of fuzzy disks has recently been defined by Pal & Rosenfeld (1991) by making a straightforward generalization of the MAT definition to the fuzzy subsets of a metric space. Since the gray level (scaled in  $[0, 1]$ ) of a pixel can be regarded as its degree of membership in the set of high-valued ("bright") pixels, a gray scale image can be regarded as a fuzzy set. A gray image, for the computation of its FMAT, does not need to be thresholded (e.g., to a binary image), whereas the MAT requires the image to be binary. The FMAT therefore has the ability to get the image back from the skeleton.

Unfortunately, for a gray image  $X$ , specifying the FMAT may require more storage space than specifying the image itself. The FMAT is seen to be redundant in the sense of representation and reconstruction of  $X$ . Moreover, its computational aspects were not addressed in Pal & Rosenfeld (1991). In fact, it is found to be very expensive if one needs to check the inclusion (subset) condition of a fuzzy disk  $g_p$  by another disk  $g_q$  for all the points in an image.

The present work describes some algorithms on these issues. The problem of reducing redundancy is tackled by selecting a set of fuzzy disks which together contain a fuzzy disk  $g_p$  for its deletion from the medial axis (MA) output. The computational algorithm developed in this context is based on the concept of "sponsoring" capability of a pixel and it involves the process of successive deletion of pixels from exterior to interior of the image. The algorithm can also be used for two tone (binary) images. Three different techniques have then been developed to provide approximated FMAT versions so that the image representation becomes economical without affecting much

of its reconstruction and it also provides good skeletons of the object regions (darker pixels). A set of criteria is finally established which allows one to work with a much smaller subset of an image space and therefore reduces significantly the time of computation of FMAT. Their graphical representation has also been provided to facilitate the task of programming. The performance of the algorithms is demonstrated on a gray image.

Section 2 describes the distance measures  $d_4$  and  $d_8$  in a digital picture, the definition of MAT of a binary (or crisp) image and its various characteristics for the convenience of understanding the FMAT of a gray image. The definition of FMAT (as defined by Pal & Rosenfeld 1991) and some remarks on it are given in § 3. Section 4 explains the redundancy removal techniques. The methods of reducing the searching space and the time of computation are explained in § 5. Various approximation techniques are described in § 6. Section 7 demonstrates the experimental results. Some key features and applications of the FMAT are described in § 8. Section 9 contains the discussion and conclusions.

## 2. Medial axis transformation (MAT)

### 2.1 Distance measures

(Rosenfeld & Kak 1982, p. 209). Let  $X$  be a digital picture (image) of size  $M \times N$ . The city block distance between two points  $P(x, y)$  and  $Q(u, v)$ ,  $x, u = 0, 1, 2, \dots, M-1$ ;  $y, v = 0, 1, 2, \dots, N-1$  in a digital picture is defined as

$$d_4(P, Q) = |x - u| + |y - v|. \quad (1)$$

The chessboard distance between them is defined as

$$d_8(P, Q) = \max(|x - u|, |y - v|). \quad (2)$$

These measures satisfy the metric properties and are simpler than the Euclidean distance to work on a digital image.

Points at a city block distance  $d_4 = 1$  and  $d_8 = 1$  from  $P$  respectively represent 4-neighbours and 8-neighbours of  $P$ . The points at a chessboard distance  $d_4 \leq r$  from  $P$  form a diamond (i.e., diagonally oriented square) centred at  $P$ . Similarly, the points at  $d_8 \leq r$  from  $P$  form an upright square of odd side length centred at  $P$ . These are shown in figure 1 when  $M = N = 5$  and  $r = 1$ . Let us call them diamond disk and upright square disk of radius 1. The distance between a point  $P$  and a set  $S$  is defined to be shortest distance between  $P$  and any point of  $S$ .

### 2.2 Medial axis transform of a binary (two tone) image

Let  $X$  be a binary (two tone) image i.e., each point in  $X$  can have a value either 0 or 1. In the previous section it is seen that the points within a given chessboard distance ( $d_8$ ) of  $P$  form an upright square of odd side length centred at  $P$ . Let us associate with each point  $P \in X$  a set of such upright squares of constant value. (Though a set of upright squares formed by chessboard distance is considered here, the discussion may be generalized to any family of mutually similar shapes formed by other distance measures.) For every point  $P$ , let  $D_P$  be the largest square (or disk)

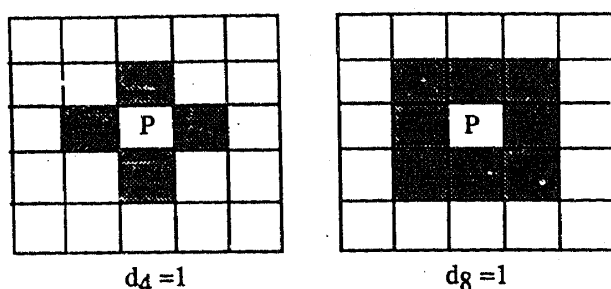


Figure 1. Disks corresponding to  $d_4 = 1$  and  $d_8 = 1$ .

that is contained in  $X$  and has a constant value  $v_P$ . Let  $r_P$  be the radius of  $D_P$ . Then  $X$  is the union of the  $D_P$ 's. Let  $C$  be a subset of  $X$  such that, for all  $P \in X$ , there exists  $Q \in C$  so that  $D_Q \supseteq D_P$ . (In other words,  $C$  is a set of points  $Q$  at which  $D_Q$  is a local maximum i.e., for any neighbour  $P$  of  $Q$ ,  $D_Q \supseteq D_P$ .) We call  $C$  a sufficient subset of  $X$ . Evidently, for any such  $C$ ,  $X$  is the union of the  $D_Q$ 's,  $Q \in C$ .

The image  $X$  can therefore be completely represented by the centres  $P$ , radii  $r_P$  and values  $v_P$  of the disks  $D_P$ ,  $P \in C$ , because any point of  $X$  must lie in at least one of these disks (blocks). The set  $C$  is called the medial axis (or symmetric axis) of  $X$ , and the set of  $D_P$ 's i.e., the set of all  $P$ ,  $r_P$  and  $v_P$  for  $P \in C$  is called the medial axis transformation of  $X$ .

Figures 2 and 3a show the MAT representations of two binary pictures of sizes  $8 \times 8$  and  $5 \times 8$  using upright square disks. The centres of maximal blocks with value  $v = 1$  are underlined. The centre coordinates  $(x, y)$  and radii of the maximal blocks are shown considering that the lower left corners of the images have coordinates  $(1, 1)$ . The images can therefore be reconstructed once the values of  $(x, y)$ ,  $r$  and  $v$  of the maximal blocks are given.

### 2.3 Some remarks

i) Note that for the binary images (figures 2 and 3), one needs to specify the maximal blocks of only one value (of 1's). The points not covered by these blocks must have the value of 0's.

If  $X$  is an  $L$  level image i.e., each point in  $X$  can take a value from  $\{0, 1, 2, \dots, L-1\}$ , then  $X$  can be viewed as consisting of  $L$  constant valued regions. In that case, one needs to specify the maximal blocks for  $(L-1)$  of the values; the points not covered by any of the maximal blocks must have the omitted value.

ii) The MAT representation may still be redundant in the sense of reconstruction of image i.e., some blocks (disks) may be contained in unions of others. For example,

|   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
| 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |

| <u>x</u> | <u>y</u> | <u>r</u> |
|----------|----------|----------|
| 3        | 3        | 2        |
| 3        | 7        | 0        |
| 4        | 3        | 2        |
| 6        | 4        | 1        |
| 6        | 6        | 0        |

Figure 2. MAT representation of an  $8 \times 8$  binary image (Rosenfeld & Kak 1982).

|   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

| <u>x</u> | <u>y</u> | <u>r</u> |
|----------|----------|----------|
| 3        | 3        | 1        |
| 4        | 3        | 1        |
| 5        | 3        | 1        |
| 6        | 3        | 1        |

|   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

(a)

(b)

Figure 3. (a) MAT representation of a  $5 \times 8$  binary image. (b) MAT after removing redundancy.

|                               |                                      |
|-------------------------------|--------------------------------------|
| 0 0 0 0 0 0 0 0               | 0 0 0 0 0 0 0 0                      |
| 0 0 <u>1</u> 0 0 0 0 0        | 0 0 <u>1</u> 0 0 0 0 0               |
| 0 0 0 0 0 <u>1</u> 0 0        | 0 0 0 0 0 <u>1</u> 0 0               |
| 1 1 1 1 1 1 1 0               | <u>1</u> 1 1 1 1 <u>2</u> 1 0        |
| 1 2 2 2 2 <u>2</u> 1 0        | 1 <u>2</u> 2 2 2 <u>2</u> 1 0        |
| 1 2 <u>3</u> <u>3</u> 2 1 1 0 | 1 2 <u>3</u> <u>3</u> <u>3</u> 2 1 0 |
| 1 2 2 2 2 1 0 0               | 1 <u>2</u> 2 2 2 1 0 0               |
| 1 1 1 1 1 1 0 0               | <u>1</u> 1 1 1 1 <u>1</u> 0 0        |
| (a)                           | (b)                                  |

Figure 4. Distance to  $\bar{R}$ , the MA (local maxima) points are underlined, (a) using  $d_8$  metric, (b) using  $d_4$  metric.

the disks  $D(4, 3)$  and  $D(5, 3)$  at the points  $(4, 3)$  and  $(5, 3)$  of figure 3a are redundant, because these are contained in the union set  $(D(3, 3) \cup D(6, 3))$ . If these are removed (as shown in figure 3b), one can still reconstruct the image. However, there does not seem to be any simple way of reducing this redundancy without carrying out a lengthy search process.

iii) When  $X$  is considered to be partitioned into  $R$  and  $\bar{R}$  (i.e., object and background, say) the medial axis of  $R$  may be viewed as consisting of only those points of  $R$  whose distances from  $\bar{R}$  are local maxima. If  $d_8$  is used as the metric, local maximum means that no 8-neighbour of the point has greater distance from  $\bar{R}$ . This is explained in figure 4a when the regions of 1's in figure 2 are considered as  $R$  and that of 0's as  $\bar{R}$ . The entries denote the  $d_8$  values of pixels from  $\bar{R}$ .

Similarly, if one uses city block distance ( $d_4$ ) for obtaining a MAT of  $R$ , the local maxima then refer to 4-neighbours. The corresponding  $d_4$  values and the medial axis thus obtained are shown in figure 4b. The MAT operation is therefore seen to be influenced by the choice of a distance measure.

iv) The concept of "maximal block" or "local distance maxima" locates the centres of the MAT at mid-points or along local symmetry axes of the regions of constant value in  $X$ . For example, if we consider the point  $P(5, 2)$  as an MA point instead of  $P(6, 4)$  in figure 4a, or consider the point  $P(7, 4)$  instead of  $P(6, 4)$  in figure 4b, the changed MAT will still be able to represent and to reconstruct the image of figure 2 with the bit requirement the same as before. Note that the disks centred at the replacing points  $(5, 2)$  and  $(7, 4)$  are not maximal (or their distances from the border are not local maxima) and therefore selection of these disks locates the MA off-centre from the uniform region.

v) Since the MA is the set of local maxima of the distances to  $\bar{R}$ , it is usually disconnected, and is two-pixel thick at the places where the region  $R$  has even width. These factors are to be taken care of while extracting the skeleton of an elongated region  $R$  using its MAT.

### 3. Fuzzy medial axis transformation

(Pal and Rosenfeld 1991). Let  $D$  be a metric space with metric  $d$  and let  $f$  be a fuzzy subset of  $D$ . For each  $P \in D$ , let  $g_P^f$  be the fuzzy subset of  $D$  with membership at each point  $Q \in D$  defined by

$$g_P^f(Q) = \inf_{d(P, R) = d(P, Q)} f(R), \quad (3)$$

with  $f(R)$  being the membership value of  $R$  in the fuzzy set  $f$ . Evidently,  $g_P^f$  is a fuzzy disk with  $g_P^f(P) = f(P)$  and  $g_P^f \leq f$  (i.e.,  $g_P^f$  is a subset of  $f$ ). Moreover, it is easy to see that

$$\sup_{P \in D} g_P^f = f.$$

Let  $C$  be any subset of  $D$  such that for all  $P \in D$ , there exists  $Q \in C$  for which  $g_P^f \leq g_Q^f$ . We call such a set  $C$  an  $f$ -sufficient subset of  $D$ . Evidently, for any such subset  $C$  we have

$$\sup_{P \in C} g_P^f = f. \quad (4)$$

In other words, the set  $\{g_P^f | P \in C\}$  can determine  $f$ . If  $D$  is finite, and we make  $C$  as small as possible,  $\{g_P^f | P \in C\}$  is a compact representation of  $f$ .

In particular, let  $D$  be a digital image  $X$  and  $f$  be a fuzzy representation of  $X$  i.e.,  $X = \{f(P), f(P) \in [0, 1]\}$ . For example,  $f(P)$ , the normalized gray value of a pixel  $P$  may be viewed as a degree of its belonging to  $f$  (or the degree to which the pixel  $P$  is bright, say). We say that  $P \in D$  is a (nonstrict) local maximum of  $f$  if  $P$  has no neighbour  $Q$  such that  $g_P^f < g_Q^f$ . Let  $C_f$  be the set of such (nonstrict) local maxima of  $f$ . Evidently, the set  $C_f$  is an  $f$ -sufficient subset of  $D$ , so that  $f$  is the sup of the  $g_P^f$ 's for all  $P \in C_f$ . In other words, the pixel intensity  $f(t)$  at a point  $t$  of the image can be obtained from

$$f(t) = \max_{P \in C_f} g_P^f(t), \quad (5)$$

because  $t$  has its maximum membership value ( $= f(t)$ ) to one of the disks  $g_P^f$ ,  $P \in C_f$ . The set  $C_f$  is called the *fuzzy medial axis* (FMA) of  $f$ , and  $\{g_P^f | P \in C_f\}$  is called the *fuzzy medial axis transformation* (FMAT) of  $f$ . If  $f$  is a crisp subset of  $D$  i.e.,  $f \in \{0, 1\}$ , the aforesaid definitions reduce to the standard definitions of the MAT (as described in § 2).

The definition of the FMAT is thus seen to involve natural extensions (generalization) of the concepts of maximal disk, union, inclusion and symmetry for an ordinary set to a fuzzy set. For a gray tone image  $X$  (the gray values being normalized in the range  $[0, 1]$ ), it computes, first of all, the various fuzzy disks centred at the pixels and then retains a few (as small as possible) of them, as designated by  $g_Q^f$ 's, so that their union can represent the entire image  $X$ .

### 3.1 Example

Consider a  $5 \times 5$  image  $X$  as shown in figure 5. The lower left pixel of intensity 4 has coordinate  $(1, 1)$ . Fuzzy disks are computed with  $d_8$  metric. The centre pixel has  $g_P = \{7, 6, 4\}$ . (Note that the gray values are not normalized in  $[0, 1]$ . Pixels outside the image are assumed to be of zero intensity while computing the disk values. The superscript  $f$  is omitted.) Here, the first entry denotes the non-normalized membership value of the pixel itself to the disk  $g_{(3,3)}$  i.e.,  $g_P$  value at zero distance or at radius  $r = 0$ . The second and third entries denote, respectively, the membership values of the pixels, which are at distances 1 and 2 from the centre pixel, for the disk  $g_{(3,3)}$ . For pixels having intensity 6,  $g_P$  has pairs of values  $\{6, 5\}$  except the one at  $(2, 2)$  for which it is  $\{6, 4\}$ . These entries correspond to  $r = 0$  and 1; the  $g_P$  value at  $r = 2$  is



|   |                 |                 |                 |   |
|---|-----------------|-----------------|-----------------|---|
| 5 | 5               | 5               | 5               | 5 |
| 5 | <b><u>6</u></b> | 6               | <b><u>6</u></b> | 5 |
| 5 | <b><u>6</u></b> | <b><u>7</u></b> | 6               | 5 |
| 5 | 6               | <b><u>6</u></b> | <b><u>6</u></b> | 5 |
| 4 | 5               | 5               | 5               | 5 |

**Figure 5.**  $5 \times 5$  digital image. Pixels belonging to fuzzy medial axis are marked bold. Pixels belonging to reduced fuzzy medial axis are underlined.

zero. For the pixels on the border of the image,  $g_P$  is specified by the single value  $f(P)$ . The  $g_P$  values are zero for  $r = 1$  and 2.

The pixels constituting the fuzzy medial axis  $C_f$  are marked bold in figure 5. The disks  $g_P$  around these pixels define the fuzzy medial axis transformation of the image  $X$ . Note that for any pixel  $P \in X$ , there exists a pixel  $Q \in C_f$  which satisfies the (subset) criterion  $g_P(t) \leq g_Q(t)$  for all points  $t$  in the image. Since the  $g_P$  of the centre pixel is defined by the triple of values, there is no other pixel for which the above subset condition can be satisfied; so it is treated as a member of  $C_f$ . Similarly, for the  $g_P$ 's of all the pixels having value 6, except the one at (2, 2), the subset condition is not satisfied when the centre pixel is considered as  $Q$ . For example, consider the pixel at the location (2, 4) as  $P$ . Let the pixel at the location (1, 5) be considered  $t$ . Then the values of  $g_{(2,4)}(t)$  and  $g_{(3,3)}(t)$  denoting the membership values of  $t$  corresponding to the fuzzy disks centred at the points (2, 4) and (3, 3) are 5 and 4 respectively. Therefore, the criterion  $g_{(2,4)}(t) \leq g_{(3,3)}(t)$  for all  $t \in X$  is not satisfied. For the border pixels of intensity 5, the subset condition is satisfied by one of the 6's in  $C_f$ , whereas it is the centre pixel which the condition satisfies for the lower-left pixel 4.

In order to reconstruct the input image i.e., to restore the deleted pixels, simply put back all the disk values of FMA pixels at those locations. In case a location has more than one such value, select the largest one. For example, if we want to get the intensity back at the point (2, 2), then put all the disk values (which are 6, 5, 5, 0, 0, 0, 0, 0) of the eight FMA pixels at that point. Since there are eight values, select the largest one i.e., 6 as the intensity at the point (2, 2). Similarly, to get the intensity back at the location (1, 1), take the largest among  $\{4, 0, 0, 0, 0, 0, 0, 0\}$ .

### 3.2 Some remarks

i) The number of the FMA pixels in  $X$ , denoted by  $|C_f|$  or  $|FMA|$ , is dependent on both  $r$ , the radius (or extent) of fuzzy disk used and the crispness of  $X$ . With increase of the crispness (contrast between object and background, say) in  $X$ ,  $|C_f|$  tends to decrease. Its minimum value corresponds to the two tone version of  $X$ . As  $r$  decreases,  $|C_f|$  increases because the possibility of a disk  $g_P$  being maximal increases. In other words, the likeliness of the condition  $g_P \leq g_Q$ ,  $P \neq Q$  being satisfied for any  $P$  decreases with decrease in  $r$ . (In the limiting case when  $r = 0$ , FMA of an  $M \times N$  dimensional image is the image itself and  $|C_f| = MN$ .) These were explained in Pal & Rosenfeld (1991).

Again, the increase in  $r$  after a certain value, say  $r'$ , may not cause further reduction in  $|C_f|$  by making some more pixels  $P$  satisfy the subset property  $g_P \leq g_Q$ . The higher the fuzziness in  $X$ , the greater will be the value of  $r'$ .

ii) It is seen from the figure 5 that, specifying the FMAT requires 17 values (1 disk requires 3 values, 7 require 2 each), as compared to 25 values required by  $X$  itself. For real images, the situation becomes even worse (Pal & Rosenfeld 1991). (Note that if we had the pixel intensity 4 of  $X$  replaced by 5, the FMAT would have been reduced to only one disk with  $g_{(3,3)} = \{7, 6, 5\}$ .)

iii) It is found to be computationally very expensive if one needs to check the inclusion (subset) condition of a fuzzy disk  $g_P^f(t)$  by another disk  $g_Q^f(t)$  for all  $P, Q$  and  $t$  in the image  $f$ .

In the following sections, we will explain some concepts and algorithms developed on these issues, namely, reducing redundancy in FMAT, extracting various approximated versions of FMAT and the related computational aspects in order to make these tasks efficient, so that the FMAT can be made practically feasible.

#### 4. Redundancy removal (RFMAT)

The removal of redundancy in pixels (fuzzy disks) from the fuzzy medial axis output is made by considering the criterion

$$g_P^f(t) \leq \sup g_{Q_i}^f(t), \quad i = 1, 2, \dots, \quad (6)$$

instead of

$$g_P^f(t) \leq g_Q^f(t). \quad (7)$$

For example, the representation of the FMAT in figure 5 is seen to be redundant. Note that the point at location (3, 4) can be removed because it is contained in the union of the fuzzy disks around (3, 3) and (2, 4) (or (4, 4)) i.e.,

$$g_{(3,4)} \leq g_{(3,3)} \cup g_{(2,4)} \quad (\text{or } \leq g_{(3,3)} \cup g_{(4,4)}), \quad (8)$$

or

$$g_{(3,4)} \leq \max\{g_{(3,3)}, g_{(2,4)}\} \quad (\text{or } \leq \max\{g_{(3,3)}, g_{(4,4)}\}), \quad (9)$$

for all pixels in  $X$ . Similar is the case with the pixel at location (4, 3) which can also be removed. The pixels representing the final reduced FMA are underlined in figure 5. Let RFMAT denote the FMAT after reducing its redundancy.

It therefore appears that one needs to go through a lengthy searching procedure to determine, for every pixel  $P$ , a set of  $Q$ 's (local maxima) which can satisfy the criterion (6) of containment in order to delete  $P$ . An efficient way of performing this task is explained below.

The algorithm checks in a preferential sequence (from exterior to interior) with every pixel, how many pixels it is sponsoring (representing) and whether the current pixel along with the ones it is sponsoring can all be replaced by some other (or a combination of others). A pixel  $P$  is said to be sponsored (or represented) by a pixel  $Q$  at a distance  $r$  if its normalized intensity  $f(P) = g_Q(P)$ , the membership value of  $P$  to the fuzzy set around  $Q$ . Obviously,  $P$  can have more than one sponsoring pixel other than itself. In such a case,  $P$  can be deleted because  $f(P)$  can be reconstructed from its sponsoring pixels i.e.,

$$f(P) = \sup_{k \in N} g_k(P), \quad (10)$$

$N$  being the set of neighbouring pixels of  $P$  at a distance  $r$ .

Now, satisfying the criterion (6) means that the fuzzy disk  $g_P$  is a subset of the union or combination of some other disks  $g_{Q1}, g_{Q2}, \dots$ , and therefore the pixel  $P$  and the ones it is sponsoring (representing) can all be sponsored by a set of pixels  $Q1, Q2, \dots$ . In such a case,  $P$  can be deleted i.e., not treated as a member of the MA because  $Q1, Q2, \dots$  can take care (in the sense of representation) of both  $P$  and its sponsoring pixels.

By rank of a pixel  $P$  we mean the extent to which the values of  $g_P(r)$ ,  $r = 1, 2, \dots$ , are nonzero. The higher the rank of a pixel (or disk), the greater is its possibility of being treated as a maximal fuzzy disk or a strong sponsor (i.e., as a member of MA). The aforesaid ranking system provides a spatial preference among the disks (or pixels) in order to facilitate the process of considering a pixel for its possible deletion. That is why the algorithm checks from exterior to interior with every pixel for its possible deletion. This also enables one to provide a compact representation of the (object) region of our interest, which may be assumed to be in the interior part of the image.

#### 4.1 Algorithm

The aforesaid concept of deletion of pixels is further explained in figure 6 for the convenience of writing a computer program. The arrows from a few border pixels point to their respective sponsors. In case a pixel gets more than one sponsor, we select the one which comes at left most (for horizontal) or upper most (for vertical) position. The central point is seen to sponsor all the pixels having intensity 6 at  $r = 1$ , plus the pixel (which has intensity 4) at  $r = 2$ .

Let us now consider the circled pixel for its possible deletion. This pixel is sponsoring only the triangled pixel. It is further seen that the triangled pixel can also be sponsored by the squared pixel (denoted by the dotted arrow) which is currently sponsoring three pixels, and the circled pixel itself can be replaced by the centre pixel having the highest intensity. Therefore, the circled pixel can be deleted from the MAT representation because this pixel and the pixel it was sponsoring can now be covered by the centre pixel and the squared pixel together; thereby making them more responsible (i.e., stronger sponsor). Successive deletion of the pixels, in this way, from exterior to interior will ultimately result in the reduced FMAT (RFMAT).

Although the technique for reducing redundancy has been developed here for a gray tone image, it is well applicable for a binary image also. It is further to be noted that, for checking of the criterion (6) for a particular pixel  $P$ , one need not search over the entire image space for determining a set of  $Q$ 's. Similar is the case with checking the criterion (7) for all  $Q$  and  $t$ . In the following section, we will be explaining

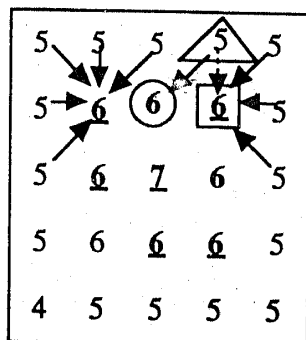


Figure 6. Sponsoring and deletion procedure.

how the time of computation of FMAT can be reduced significantly by restricting the searching space.

### 5. Reducing the search space

Let  $n_P$  and  $n_Q$  be the number of membership values of the fuzzy disks centred at  $P$  and  $Q$  respectively. Let  $d(P, Q)$  denote the distance ( $d_8$  metric of (2)) between  $P$  and  $Q$ . A careful investigation of the definition of FMAT then reveals the following criteria

$$n_P < n_Q, \quad (11a)$$

$$d(P, Q) < n_Q, \quad (11b)$$

and

$$t \in \{\text{the supporting points } S_P \text{ of } P \text{ i.e., the points for which } g_P^f(t) > 0\}, \quad (11c)$$

for checking with every pixel  $P$  the condition  $g_P^f(t) \leq g_Q^f(t)$ . The underlying concept is that the set of supporting points of the upright square fuzzy disk (since we are using  $d_8$  metric) centred at  $Q$  should include those of  $P$  i.e.,  $S_P \subset S_Q$ . These conditions allow working with a much smaller subset of an image and therefore reducing the time of computation.

Similarly, for determining a sponsor  $Q$  for a given  $P$ , while reducing the redundancy in FMAT, the criterion followed is simply

$$d(P, Q) \leq n_Q. \quad (12)$$

It is to be noted from figure 6 that the consideration of  $t$ 's does not arise here.

The significant reduction of searching space while computing the FMAT of an image and reducing its redundancy is explained graphically in figures 7a and b respectively. The shaded portions represent the domain of  $Q$  to be only considered for a particular  $P$ . The square at the bottom of figure 7a indicates the domain of  $t$  corresponding to  $P1$ .

Note that the number of rows  $M$  and columns  $N$  considered here are different. There are four bisectors drawn from four corners of the image. They meet at the centre and split the image space into two triangles and two trapezoids. (The splitting may result in four triangles or four trapezoids depending on the values of  $M$  and  $N$ .) For a given  $P$  in figure 7a, two lines are then drawn at an angle of  $\pm 45^\circ$  with the horizon. The space covered by these lines with the two closest bisectors constitutes the domain of  $Q$ . Similarly, the square drawn with length equal to twice the distance of  $P$  from its nearest boundary constitutes its domain of  $t$ .

In the case of figure 7b, first of all, draw two horizontal and two vertical lines placed at half the distances of  $P$  from the boundaries. The rectangle thus formed by their intersections constitutes the maximum possible domain of  $Q$ . This rectangular

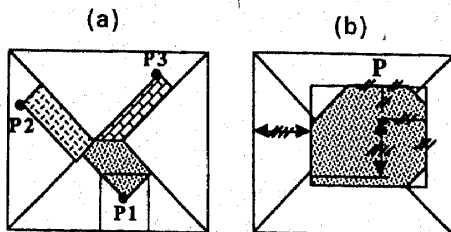


Figure 7(a). Searching space of  $Q$  and  $t$  for computing FMAT. (b). Searching space of  $Q$  for computing RFMAT.

region is further reduced by truncating the corners according to the criterion (12). These graphical representations are given for the convenience of writing an efficient programme.

## 6. Methods of approximation (AFMAT)

The RFMA output of an image  $X$  as obtained in the previous section can be regarded as a generalized fuzzy axis of symmetry and provides a kind of skeleton (of the elongated regions) of  $X$  by allowing it to be a fuzzy subset of the image. The set of disks  $\{g_P | P \in \text{RFMA}\}$  provides a compact representation of the fuzzy set  $f$  characterizing the image  $X$ . The value of  $f(P)$ ,  $P \in \text{RFMA}$  may be viewed as denoting the degree of belonging of  $P$  to the fuzzy subset "medial axis" of  $X$  as far as its compact representation (and hence exact reconstruction) is concerned.

Since specifying the FMAT, even after reducing redundancy, needs more storage, we will describe here three different types of procedures for approximating RFMAT and/or reconstructed output (for more economical representation without affecting the reconstruction much) of an image.

### 6.1 Approximation during reconstruction (method 1)

The approximation technique described here has the following two steps:

Step 1: Construct RFMAT of an image  $X$  using fuzzy disks of radius  $r$ , say.

Step 2: Reconstruct the image with disk values from 0 to  $z$ ,  $z < r$ .

It means, the storage space of  $X$  is reduced simply by keeping a few lower radii (truncated) disk values for the purpose of its reconstruction. The extent of the exactness of reconstruction is obviously dependent on the value of  $z$ . The smaller the value of  $z$ , the lesser is the storage requirement and the lower will be the exactness in reconstruction. (Note that  $z = 0$  corresponds to the RFMA output.)

### 6.2 Approximating RFMAT (method 2)

The previous algorithm does not involve any approximation on the RFMAT output; it only truncates the disk values during reconstruction. That is, the number of disks remains the same. The present algorithm, on the other hand, makes an approximation on the FMAT itself by truncating off a set of disks, while keeping the radius of disks unchanged. The truncation is achieved based on the importance of an individual RMA pixel (or RMA fuzzy disk) on the basis of its sponsoring capability. The higher the sponsoring capability of a pixel (or disk), the greater is its importance in representing the image.

Let us consider, for example, the sponsoring capability map (figure 8) of the RFMA output of figure 5. Here, the set of 0's means that the corresponding pixels are deleted because these have been taken care of by others. The number "1" implies that the corresponding MA pixel sponsors only itself. "2" corresponds to a pixel which can sponsor some one other than itself, and so on for other numbers in figure 8. The histogram of figure 8 will therefore represent the number of MA pixels with different sponsoring capabilities, e.g., it has 19, 0, 1, 1, 1, 2 & 1 pixels with capability 0, 1, 2, ..., 6 respectively. (The sum of these values is equal to the total number of pixels in  $X$ .)

|   |   |   |   |   |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 |
| 0 | 6 | 0 | 5 | 0 |
| 0 | 2 | 5 | 0 | 0 |
| 0 | 0 | 3 | 4 | 0 |
| 0 | 0 | 0 | 0 | 0 |

Figure 8. Sponsoring capability map of RFMAT in figure 5.

For real images, the histogram is usually seen to have monotonic non-increasing behaviour. This leads one to eliminate a large number of low sponsoring pixels from the MA, thereby making the image representation economical without affecting much of its reconstruction. For example, if we delete the MA pixel at location (3,2) of sponsoring capability 2, it will produce an approximated reconstruction with the original pixel value 5 at the point (2,1) being changed to 4.

The algorithm therefore has the following three steps:

- Step 1: Construct RFMAT of an image  $X$  using fuzzy disks of radius  $r$ , say.
- Step 2: Construct sponsoring capability map of RFMAT output (i.e., RMA pixels).
- Step 3: Select a sponsoring capability number  $T$ , say, and delete all the MA pixels of sponsoring capability  $\leq T$ .

The degree of exactness in reconstruction will obviously depend on the number of deleted MA pixels and their sponsoring capability. The higher these values are, the lower is the exactness. It will also depend, to some extent, on the value of  $r$ . These have been demonstrated in §7.

### 6.3 Extracting an optimum compact RFMAT (method 3)

The techniques described in the previous two algorithms affect the pixel intensities over the entire image by making the approximation distributed. In many applications, we are more interested (or only interested) in the MA of the object regions represented by darker pixels. In that case, one may delete from the RFMA output some of the pixels (or disks) representing the medial axis of background (i.e., lighter) regions; thereby resulting in an approximate medial axis of  $X$ .

A technique is described here to perform this task automatically where the deletion process is guided by the principle of minimizing ambiguity in the geometry (or in spatial domain) of RFMA. In other words, considering a fuzzy subset  $\mu$  defined over the FMA output of  $X$ , its optimum (in the sense of minimizing ambiguity in geometry) version can be extracted from one of its  $\alpha$ -cuts having maximum "compactness" value. This optimum version may be viewed as an approximate MAT representation of the image  $X$ , because it keeps only a set of darker MA pixels (to an optimum amount in the sense of maximizing compactness) for representing the object regions of our interest and ignores the rest.

The extent of reconstruction, particularly the background region, will be determined by the radius  $r$  of the remaining disks. The larger the disk radius, the more will be the coverage of background region. Before describing the method of obtaining an

optimum FMAT, let us give the definition of compactness of a fuzzy set and the way of determining its optimum version in terms of  $\alpha$ -cuts.

*Compactness of a fuzzy set:* Let  $\mu$  denote a piecewise constant fuzzy representation of an image  $X$ . Compactness of  $\mu$  is defined as (Rosenfeld 1984; Pal & Dutta Majumder 1986; Pal 1989; Pal & Ghosh 1992)

$$\text{Comp}(\mu) \triangleq a(\mu)/p^2(\mu), \quad (13)$$

where

$$a(\mu) \triangleq \sum_i \mu_i, \quad (14)$$

and

$$p(\mu) = \sum_i \sum_j \sum_k |\mu_i - \mu_j| |A_{i,j,k}|, \quad (15)$$

$$i, j = 1, 2, \dots, r; \quad i < j,$$

$$k = 1, 2, \dots, r_{ij},$$

denote the area and the perimeter of  $\mu$  respectively.  $a(\mu)$  is the weighted sum of the area of the regions on which  $\mu$  has constant value, weighted by these values.  $p(\mu)$  is the weighted sum of the length of the arcs  $|A_{ijk}|$  along which the  $i$ th and  $j$ th regions having constant values  $\mu_i$  and  $\mu_j$  respectively meet, weighted by the absolute difference of these values.

*Optimum  $\alpha$ -cut:* Let  $\mu$  denote the fuzzy set representation of RFMA output of an image  $X$ . Let each element of this output be normalized so that its membership value  $\mu(P)$  can lie in  $[0, 1]$ . The  $\alpha$ -cut of such a  $\mu$  plane is defined as

$$\mu_\alpha = \{P \in \text{RFMA} \mid \mu(P) \geq \alpha\}, \quad 1 > \alpha > 0. \quad (16)$$

Modification of  $\alpha$  will therefore result in different approximated FMA planes with varying  $\text{Comp}(\mu)$  value. As  $\alpha$  increases, the  $\text{Comp}(\mu)$  value initially increases to a certain maximum (peak) and then for a further increase in  $\alpha$ , the  $\text{Comp}(\mu)$  measure decreases.

The initial increase in  $\text{Comp}(\mu)$  value can be explained by observing that for every value of  $\alpha$ , the background medial axis pixels having  $\mu$ -value less than  $\alpha$  are not taken into consideration. So both area (14) and perimeter (15) are less than those for the previous value of  $\alpha$ . But the decrease in perimeter is greater (since the lighter pixels representing the MA of background are highly disconnected, they have high perimeter value) than the decrease in its area and hence the compactness (13) increases (initially) to a certain maximum corresponding to a value  $\alpha = \alpha'$ , say.

Further increase in  $\alpha$  (i.e., for  $\alpha > \alpha'$ ), results in a  $\mu_\alpha$  plane consisting mainly of the object (darker) MA pixels which are not much disconnected (because most of the lighter or background MA pixels which were highly disconnected have already been dropped out). As a result, the decrease in area here is more than the decrease in perimeter, and  $\text{Comp}(\mu)$  decreases. The  $\mu_\alpha$  plane having a maximum compactness value can be taken as an optimum (in the sense of compact representation of  $\mu$ ) FMA (and hence FMAT) for the (object) regions of our interest in the image  $X$ . This optimum version may also be treated as an approximate FMAT representation of the image  $X$ ,

because it retains only an optimal (in the sense of maximizing compactness) set of darker MA pixels for representing the object region and ignores the rest.

Note further that the  $\mu_x$  plane, as obtained above, can be regarded as a subset denoting the optimum fuzzy skeleton of elongated regions of the image  $X$ . The plane  $\mu_x$  is most compact and has minimum spatial ambiguity as far as its skeleton extraction is concerned.

## 7. Experimental results

Figure 9 shows a  $36 \times 60$ , 32 level input image of "S". Figure 10a is its FMAT output, after reducing redundancy in the medial axis, when the fuzzy disks are computed only up to the radius  $r = 2$  (i.e., disk values at radii 0, 1 and 2 only are considered for computing the FMAT). The number of resulting disks in figure 10a is 943. This consists of 781 disks with 3 membership values and 81 disks each with 2 and 1 membership values; thereby making the total number of values equal to 2586. (Note that specifying the original image (figure 9) itself needs only 2160 ( $= 36 \times 60$ ) values.) The situation becomes worse for larger values of  $r$  which lead to lesser numbers of disks but with larger numbers of total membership values for its exact representation. This is illustrated in table 1. For example, the RFMAT output corresponding to  $r = 9$  needs 879 disks with total number of values equal to 5449 (271 disks, 45 disks, 48 disks, 58 disks, 61 disks, 80 disks, 73 disks, 81 disks, 81 disks and 81 disks with membership values 10, 9, 8, ..., 1 respectively). If we increase the disk radius to 17 (maximum value), the image still needs 879 disks, but the aforesaid 271 disks, each of 10 membership values, have now been split up into 44, 40, 36, 36, 28, 27, 26, 18 and 16 disks with membership values 10, 11, 12, ..., 18 respectively. As a result, the total number of membership values to specify the image increases to 6326. The

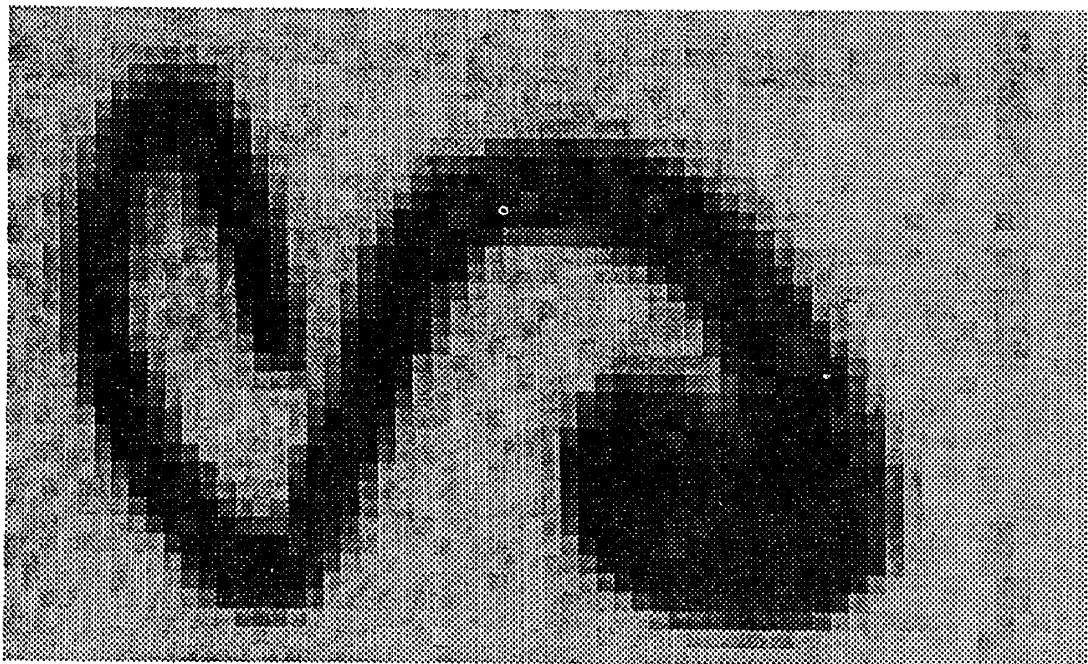


Figure 9.  $36 \times 60$  "S" input image.





Figure 10. RFMAT output of figure 9. (a)  $r = 2$ , (b)  $r = 17$ .

corresponding RFMAT output is shown in figure 10b. (For  $r = 0$ , the RFMAT is just the entire set of pixels in the image.)

Figures 11 and 12 show the sponsoring capability map (number of pixels being sponsored by an individual MA pixel) and the sponsoring capability histogram (number of occurrences of the MA pixels of different sponsoring capabilities) of the RFMAT output in figure 10a. As expected, figure 12 shows monotonic non-increasing behaviour.

Figure 13 shows the reconstructed output using method 1 for three possible combinations of  $r$  and  $r_1$ . Comparing figure 13a with figure 13b, it is seen that

**Table 1.** Variation of number of fuzzy disks and membership values with radius  $r$  for computing RFMAT of "S" image.

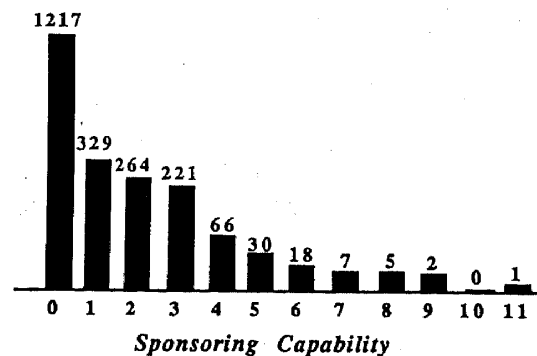
| Radius | Fuzzy disks | Membership values |
|--------|-------------|-------------------|
| 17     | 879         | 6326              |
| 15     | 879         | 6276              |
| 14     | 879         | 6216              |
| 10     | 879         | 5676              |
| 9      | 879         | 5449              |
| 8      | 880         | 5187              |
| 6      | 883         | 4526              |
| 5      | 885         | 4112              |
| 4      | 890         | 3648              |
| 3      | 901         | 3118              |
| 2      | 943         | 2586              |
| 1      | 1126        | 2165              |
| 0      | 2160        | 2160              |

```

1 1 11 11 11 11 1 111 11 1 1 11 1 111 1 1
1 3112 31 223 4 3 2 2 322 42 223 333 242 31
1 51 23 4 23231 25221 143 5 42421115 3112 212 4 2 11 1
24 32 7 13 2 5 32 22 11 72221 2111151111321
2 733 4 321 15 1 228 1 12 1 1 12 11 61 21122 41
4 11222 1 21 1 1 4 1 2 511 22
1 222213 2 1 22 2 31211 42 216
1 323 1143 21 4 21 3 1 1 31 315
1 323 2243 2 1 334323 3 4 2 5 2 1 13
533 32 2 11 5233 2113231 5 9 1 1 2 3
62 3 2 2232 3 2 32 1322 26 23 4 11 122
63 2422 12 137 3211 3 42 3 5 3 1 31
13 24 33 6 2 33 3 2 2 3 2 3 231 14 73 2
1242 33 22 3 33 333 62 2 1 51
11112231 41 24 23333 3 3 5 3 233 7 51311
1 42 314 4332 1 5 23 413 6 4 3111
4 2226 232 12 131 3 3 4 14222 2 1
2 31 5 233 243 41 11 23123 22 2 11
12 3211 3 233 231 6 2 3 11 2 1 25 3 2 312 7
312 2 23 53 3 4 1 2 1 33 11 312 11
6 2 312221 3 6 3 411 3 42 1 21326 122 1 2 1
32 1 4 2233 4 2 3222 11 1 111
1 2 232 1 4 1 3 3 4 4 1 3 32 12118 8 1311
1 23 11 222 211 33 42 211 222 322121 4
223 233 1221 2 1 3 23 11 2 4 2222 223
1 33 333 5115 4 3 443111 3222 22 1 412
1 1 2 3333 2332 123 6 32 833 B 1 1 3 111
12 2 3333 322 2 2 33 21 623213311 1511
121 3232 333 2 4 2 3112 311 1323 21 1 24
112 11 3132 2 3 3 2 1 231 221 3 215 6
112 6 4 333 3233 1 9 3 331 11 3143 5 3 52
11 2 2 34233 4 12 1 2 5 3321111 2 21 14 1
1 211 43 22212 44 1 3342133 21 2
11 211 1 32 21 4 3 22 6 833 45 3213 2 2
111 111 5133 11 211 1 1 3 1 11 13 4
1111 11 11 1 1 1111 11111111 11 1 1111 1

```

**Figure 11.** Sponsoring capability map of figure 10a. The letters A and B correspond to the numbers 10 and 11.

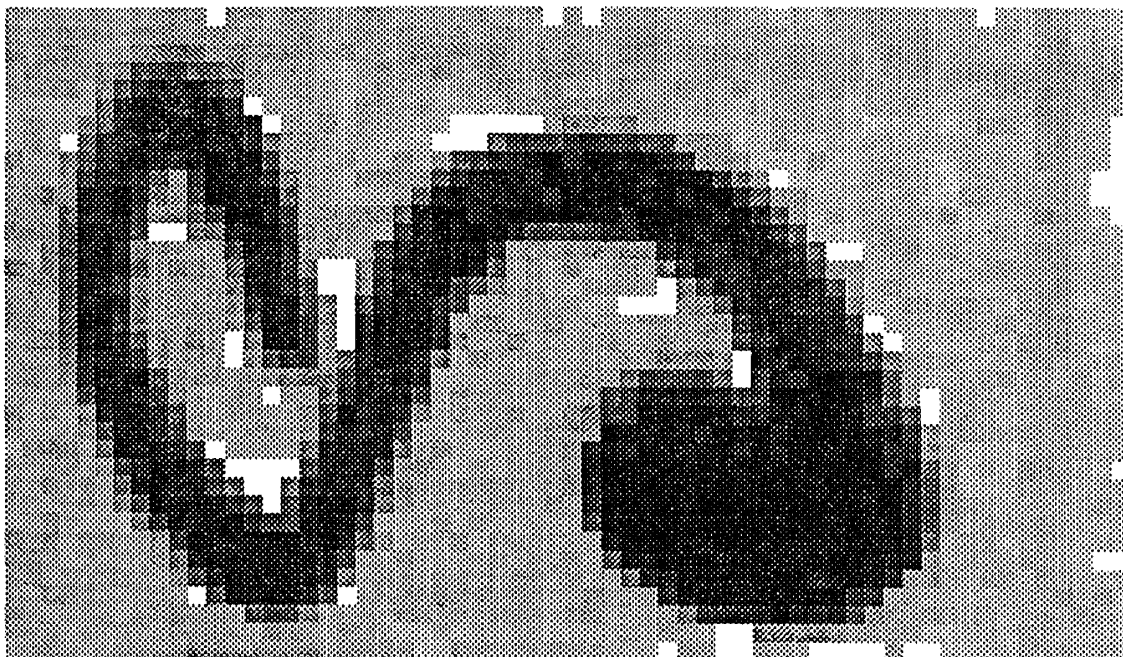


**Figure 12.** Sponsoring capability histogram of figure 10a.

figure 13a has less pixels missing (blank) because it has been produced with a comparatively larger number of disks. Furthermore, most of the uncovered pixels are seen to lie around the edge of the object region. This is obvious from their RFMAT output (figure 10) which shows that the corresponding pixels were taken care of by others, particularly by object pixels at distance  $d_g \geq 2$ . Since all the disks after radius  $r \geq 2$  have been truncated, those lost pixels could never be recovered. Even keeping one more radius (i.e., truncating disks after  $r \geq 3$ ) could not entirely avoid the appearance of these blanks (see figure 13c). The total number of membership values cut down by this approximation method is explained in table 2.

Figure 14 shows the reconstructed output using method 2 for four possible

(a)



(b)

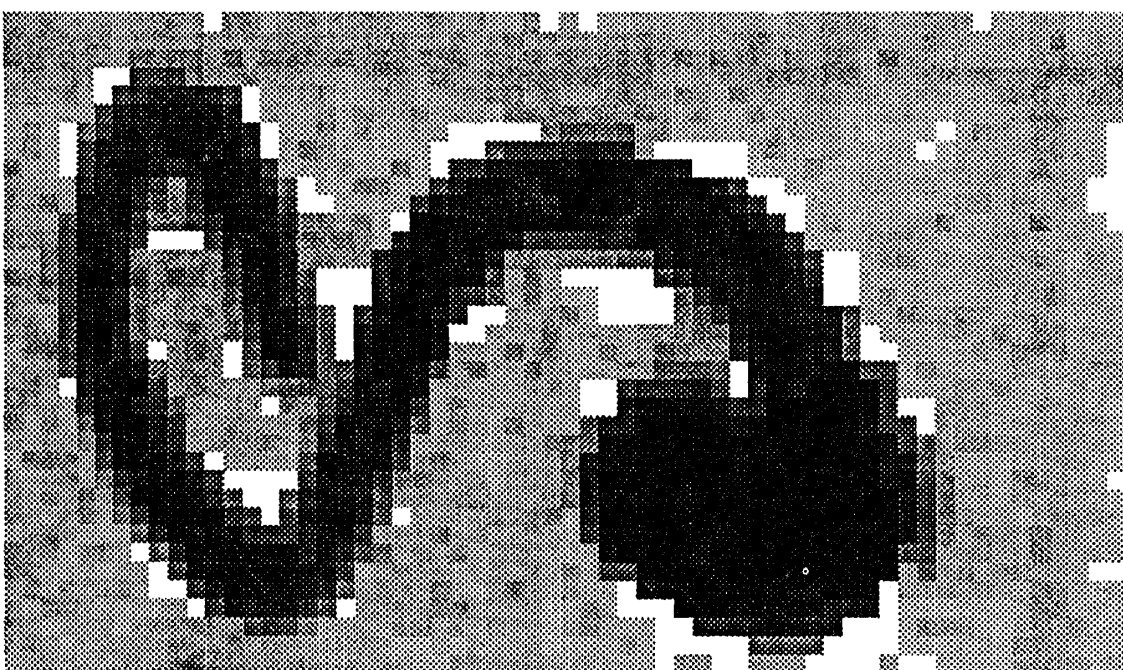
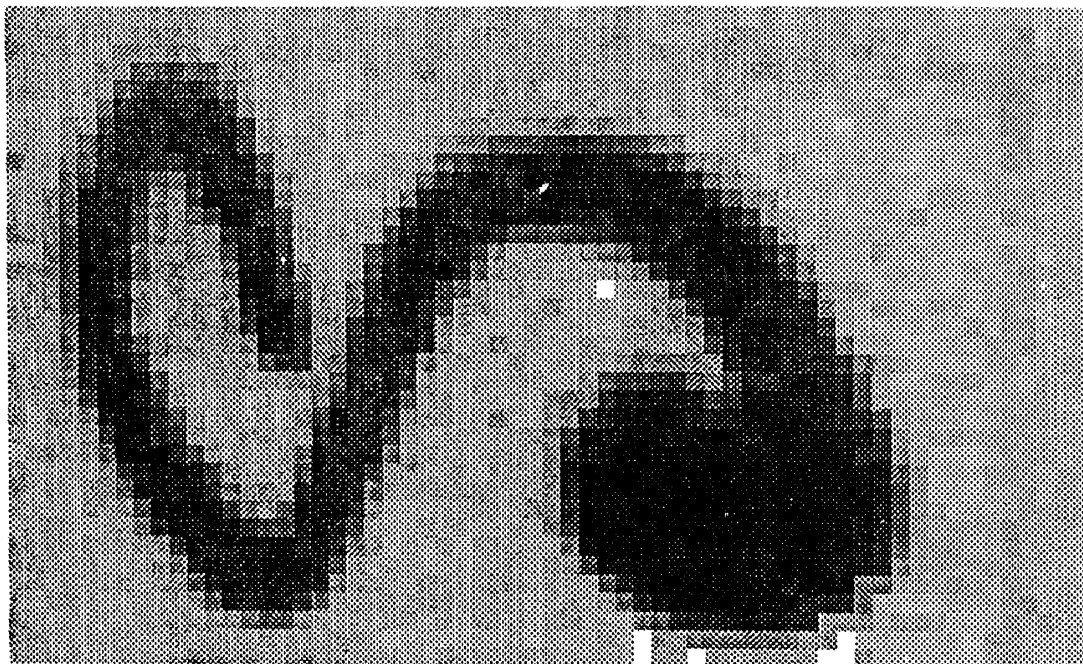


Figure 13. (a) & (b) (continued on p. 270)

(c)



**Figure 13.** (Continued) Reconstructed output using method 1. (a)  $r = 2$ ,  $z = 1$ ; (b)  $r = 17$ ,  $z = 1$ ; (c)  $r = 17$ ,  $z = 2$ .

combinations of  $r$  and  $T$ . Figure 14a corresponds to the case when all the 329 MA pixels of sponsoring capability 1 (figure 12) have been deleted. The image (figure 14a) thus reconstructed from the truncated RFMAT still looks very similar to the original input. Note that the number of fuzzy disks remaining after truncation is 614. This consists of 568 disks with 3 membership values and 46 disks with 2 membership values, thereby making the total number of values in the truncated FMAT equal to 1796. Figure 14b shows another approximated output when all the 264 MA pixels with sponsoring capability 2 are further deleted. This brings the total number of membership values in the truncated FMAT further down to 1023 (323 fuzzy disks with 3 membership values and 27 disks with 2 membership values). The approximated RFMAT (AFMAT) versions corresponding to  $r = 17$  (figures 14c and d) have 512 and

**Table 2.** Number of fuzzy disks and membership values for different approximated FMAT (AFMAT) versions of "S" image.

| AFMAT* | Fuzzy disks | membership values |
|--------|-------------|-------------------|
| 9a     | 943         | 1805              |
| 9b     | 879         | 1677              |
| 9c     | 879         | 2394              |
| 10a    | 614         | 1796              |
| 10b    | 350         | 1023              |
| 10c    | 512         | 4384              |
| 10d    | 345         | 3136              |
| 11a    | 368         | 1095              |
| 11b    | 343         | 3567              |

\*Numbers pertain to the respective figure numbers.

345 disks (table 2). Since these disks use maximum possible radius, the approximated representation of FMAT does not turn out to be economical.

Table 3 shows the variation of  $\text{Comp}(\mu)$  with  $\alpha$  (method 3) when figures 10a and b respectively are considered inputs. The maximum (peak) values are asterisked. The optimum RFMA versions corresponding to the global maximum  $\text{Comp}(\mu)$  values are shown in figure 15 along with the reconstructed images (figure 16) obtained from them.

The reconstructed images demonstrate the capability of the algorithm in providing automatically an optimum FMA, which provides a good skeleton and is very effective and economical (in the sense of storage requirement) in representing the object regions. Here again, the use of the maximum possible disk radii (figure 16b) makes the image representation expensive.

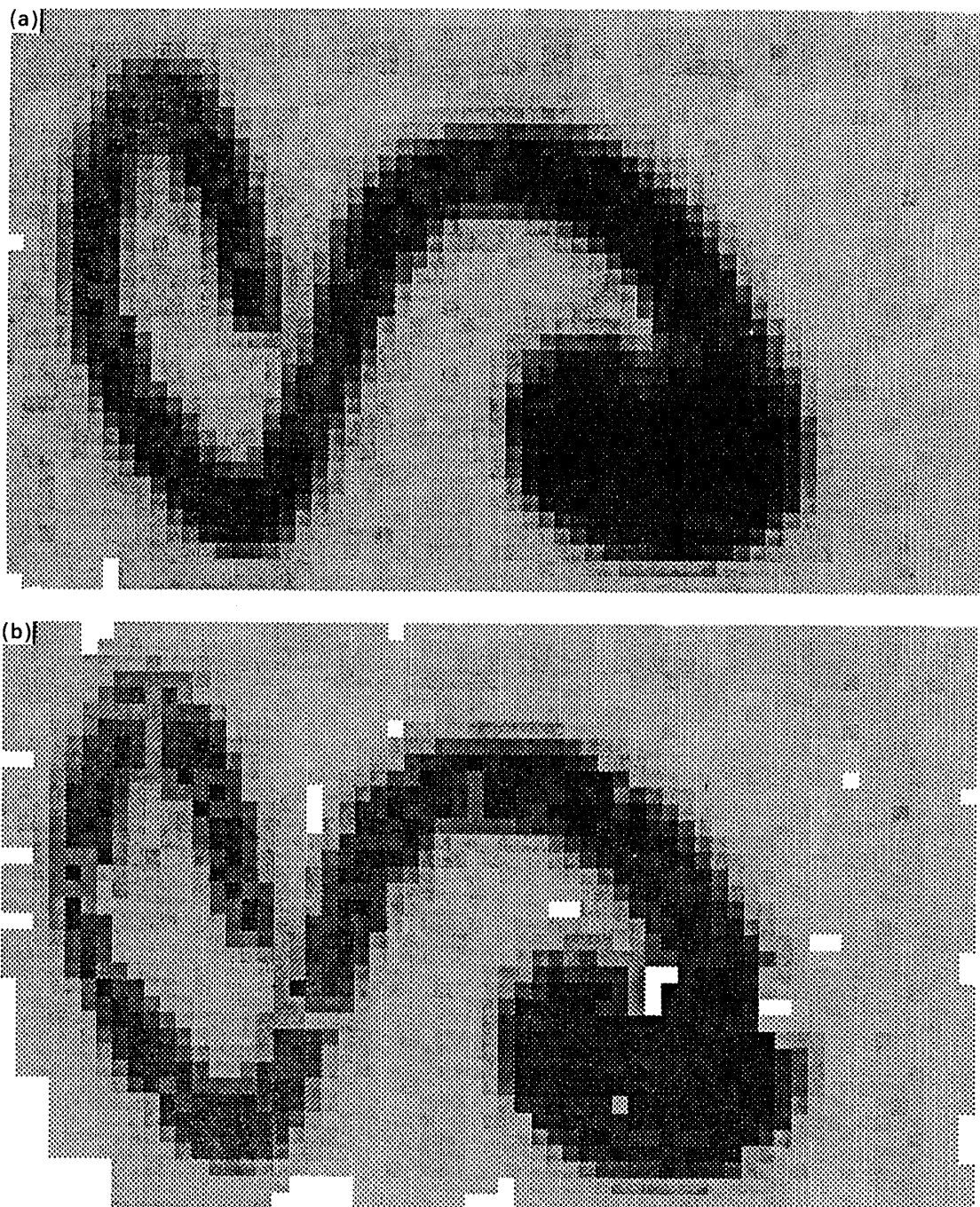


Figure 14. (a) & (b) (continued on p. 272)



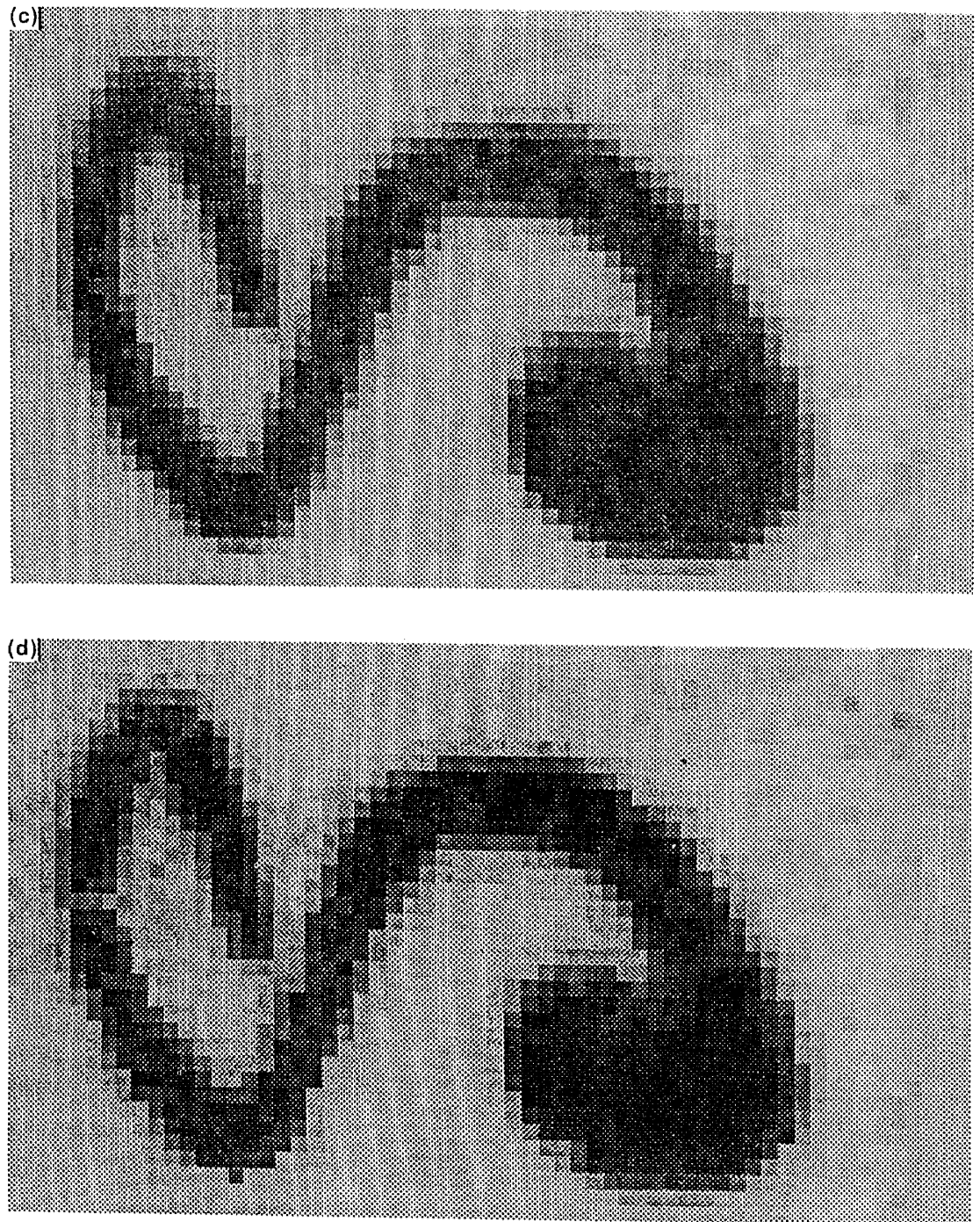


Figure 14. (Continued) Reconstructed output using method 2. (a)  $r = 2$ ,  $T = 1$ ; (b)  $r = 2$ ,  $T = 2$ ; (c)  $r = 17$ ,  $T = 1$ ; (d)  $r = 17$ ,  $T = 2$ .

## 8. Significance and applications

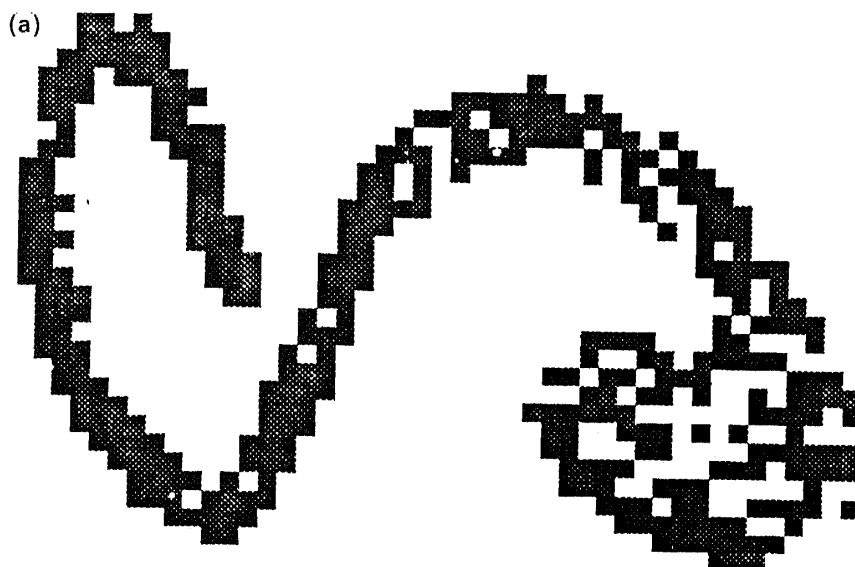
A few example applications of the FMAT are described below.

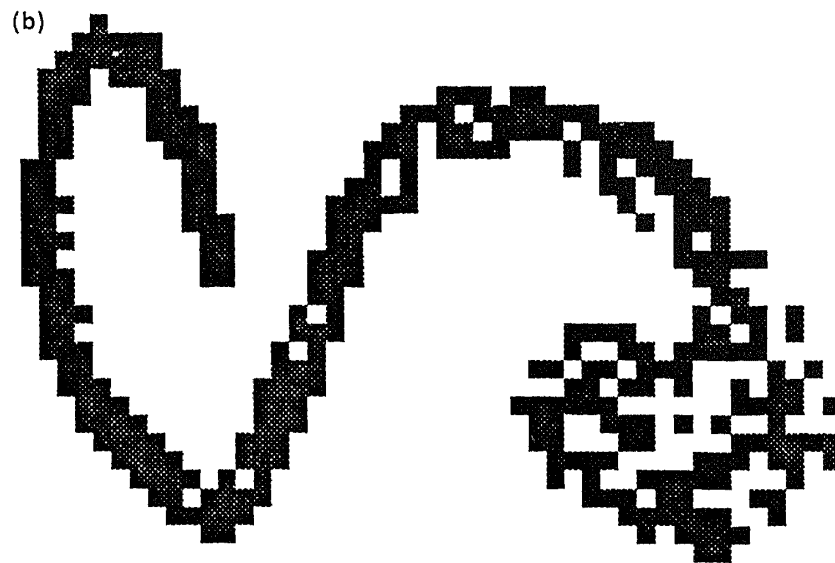
The FMAT is seen to be capable of extracting the fuzzy skeleton of an ill-defined object directly from the input image without involving the concept of image segmentation. This means, the uncertainty in selecting a proper threshold, as in conventional procedures, does not arise. Furthermore, the fuzzy skeleton has provision to retain the information content of the original image for making efficient decisions at the

**Table 3.** Variation of  $\text{Comp}(\mu)$  with  $\alpha$  for figures 10a and b.

| Non-normalized<br>$\alpha$ | $10^4 \text{ Comp}(\mu)$<br>Figure 10a | $10^4 \text{ Comp}(\mu)$<br>Figure 10b |
|----------------------------|--|--|
| 0                          | 7.35                                   | 7.87                                   |
| 1                          | 7.35                                   | 7.87                                   |
| 2                          | 7.35                                   | 7.87                                   |
| 3                          | 7.35                                   | 7.87                                   |
| 4                          | 7.43                                   | 7.95                                   |
| 5                          | 7.62                                   | 8.16                                   |
| 6                          | 8.10                                   | 8.67                                   |
| 7                          | 9.65                                   | 10.34                                  |
| 8                          | 10.56                                  | 11.39                                  |
| 9                          | 10.84                                  | 11.61                                  |
| 10                         | 11.15                                  | 11.93                                  |
| 11                         | 11.24                                  | 12.07                                  |
| 12                         | 11.52                                  | 12.39                                  |
| 13                         | 11.61                                  | 12.43                                  |
| 14                         | 11.94                                  | 12.57                                  |
| 15                         | 11.98                                  | 12.61                                  |
| 16                         | 12.51                                  | 12.82                                  |
| 17                         | 12.61*                                 | 12.92*                                 |
| 18                         | 12.55                                  | 12.88                                  |
| 19                         | 12.58                                  | 12.90                                  |
| 20                         | 12.97*                                 | 13.14                                  |
| 21                         | 12.96                                  | 13.31*                                 |
| 22                         | 12.43                                  | 12.78                                  |
| 23                         | 12.11                                  | 11.93                                  |
| 24                         | 12.76                                  | 12.85                                  |

\* Denotes maximum

**Figure 15a.**



**Figure 15.** (Continued) Optimum  $\alpha$ -cut version of RFMAT output. (a)  $r = 2, \alpha = 20$ ; (b)  $r = 17, r = 21$ .

higher levels (e.g., extracting/defining fuzzy primitives, relations and production rules) for the purpose of recognition. This will cause the ultimate output (result) of the system to be associated with least uncertainty.

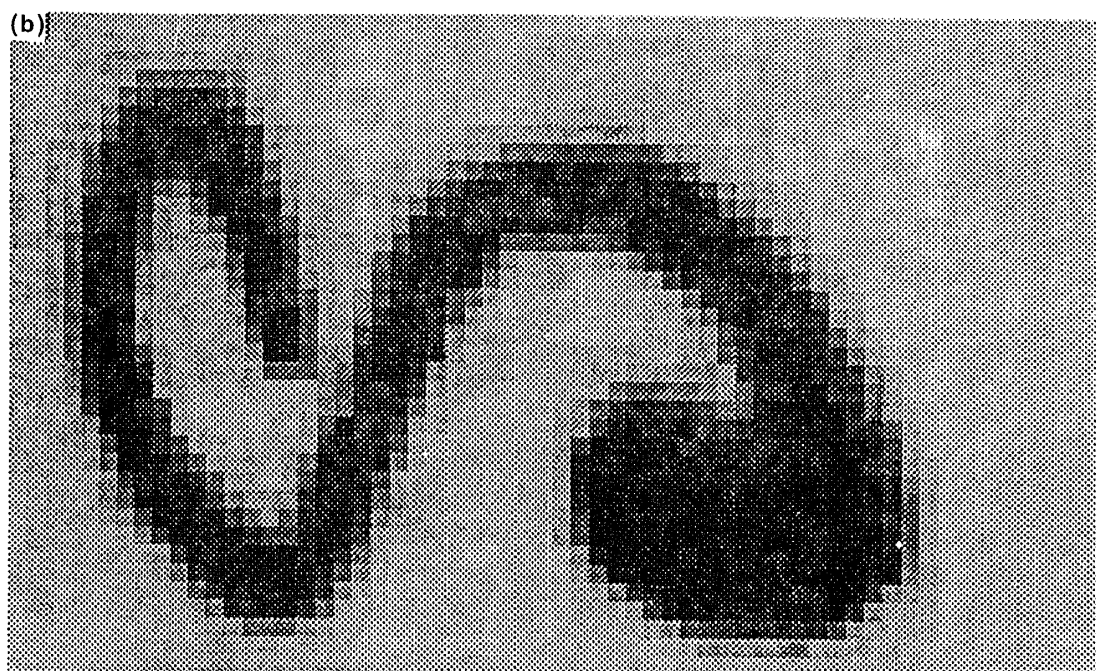
Since the FMA (i.e., the medial axis pixels) represents the structural shape of a region in a reduced form, it can be used as a core line or template of the image for its shape analysis. Therefore, the FMA of an image can also be applied directly to a pattern recognition procedure (without going through the stages of extracting fuzzy primitives and generating fuzzy grammars) that compares directly an unknown FMA to a set of stored fuzzy medial axes.

The third application is based on the property that the FMAT can represent and reconstruct the image *exactly*. This property can be explored to an image storage



**Figure 16a.** (Continued)





**Figure 16.** (Continued) Reconstructed output using method 3, (a)  $r = 2$ ,  $\alpha = 20$ ;  
(b)  $r = 17$ ,  $\alpha = 21$ .

retrieval application. The unique characteristic of this process is that the reconstruction procedure is very robust and can degrade gracefully with respect to the decreasing number of FMA disks and/or disk radii. From the primary FMAT definition, the FMA pixels are seen to be highly redundant. Each FMA pixel not only stores its own membership value, but also stores the membership values of its neighbours. With this coupled disk environment, the end user can then have the opportunity to make the representation economical by selecting only a set of disks based upon the individual's storage and/or representation requirements.

The notion of image reconstruction can further be extended to include creating a new image from the original image. This idea can be realized from the fact that the FMAT transforms the original image from a single valued pixel space into a multi-valued disk space. With the multi-valued disk representation the FMA pixels are aware of their surroundings. In addition, the FMA pixels also represent the core lines for elongated objects. These core lines may be moved, rotated, cut and coupled from one part of the FMA region to another. After an editing procedure, a new image can easily be created by applying the reconstruction procedure of the FMAT. One can therefore create many poses of the same image without actually storing all the different configurations.

## 9. Discussion

The problem of extraction of medial axis transformation of a gray image with reference to skeletonization, image representation and uncertainty management in a vision system has been addressed. The relevance of fuzzy set theory in this context has been highlighted. Some algorithms for reducing redundancy in the FMAT, as defined by Pal & Rosenfeld (1991), for generating its approximate versions and reducing the

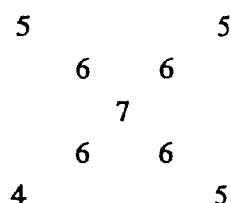
searching spaces for its computation have been described. The ultimate aim is to be able to make the FMAT practically feasible and useful to a gray image for its skeleton extraction and compact representation, for shape analysis and template matching, for representation and retrieval, for uncertainty management in recognition and for creating new images of various poses.

Computational aspects for the convenience of writing an efficient program have also been addressed. The approximation algorithms initially require the removal of redundancy in the FMAT. Two kinds of approximations are made, namely, distributed approximation and optimum fuzzy compactness version. The distributed approximation involves the concept of the sponsoring capability of MA pixels. On the other hand, the optimum compactness version can be obtained automatically based on the criterion of minimum spatial ambiguity (maximum fuzzy compactness) from the various  $\alpha$ -cuts of the fuzzy MA plane. It is to be noted that for some images, there may not be any optimum (peak) attained in its  $\text{Comp}(\mu)$  variation with  $\alpha$ . In such cases, the  $\alpha$ -cut for which there is a "maximum increase" in  $\text{Comp}(\mu)$  value of the image may be taken as the optimum one.

The optimum FMA output of the "S" image can be regarded as its optimum fuzzy skeleton (in the sense of minimum ambiguity) where the gray value of a pixel  $P$  (i.e.,  $g_p(0)$  value),  $P \in \text{RFMA}$  denotes the degree of its belonging to the fuzzy subset "skeleton". Such skeletons are seen, in a few parts, to be disconnected. The reason (as explained in §2.3) is that these are constituted by extracting the maximal fuzzy disks. If a non-fuzzy single-pixel width skeleton is desired, it can be obtained by a contour tracking algorithm (Pal *et al* 1983; Pal & Dutta Majumder 1986) which takes into account the direction of contour, multiple crossing pixels, lost path due to spurious wiggles etc. based on the octal chain code. Connectivity can further be preserved either by putting back the maximal disks which were initially deleted for reducing the redundancy in the FMAT representation or by inserting pixels having intensity equal to the minimum of those of pairs of neighbours in the object. In such a case, the compactness in representation would be further improved.

It would be of interest if some quantitative analysis can be made on the degree of exactness in reconstruction. The measure "higher order image entropy" based on the probability of co-occurrence of pixels (and using logarithmic gain function (Pal & Pal 1989a) or exponential gain function (Pal & Pal 1989b)) can be used in such an analysis.

The results shown in the paper were computed using  $d_8$  metric (2). If instead of using  $d_8$ , we use  $d_4$  (1) distance measure, the fuzzy disks will have diamond-like shapes (diagonal square). As a typical illustration, the RFMA output of figure 5 using  $d_4$  metric is shown in figure 17. Note that the number of MA disks in figure 17 is increased, because the uniform regions in figure 5 are approximated better by disks of square shape than by those of diamond shape.



5 Figure 17. RFMA output of figure 5 using  $d_4$  metric.

This work was done while SKP held an NRC-NASA Senior Research Associateship at the Lyndon B Johnson Space Center, Houston, Texas. He is grateful to Dr Robert N Lea for his interest in this work and to the Indian Statistical Institute, Calcutta, for allowing him leave for the same.

## References

- Ahuja N, Davis L S, Milgram D L, Rosenfeld A 1978 Piecewise approximation of picture using maximal neighborhoods. *IEEE Trans. Comput.* C-27: 375-379
- Bezdek J C 1981 *Pattern recognition with fuzzy objective function algorithm* (New York: Plenum)
- Bezdek J C, Pal S K (eds.) 1992 *Fuzzy models for pattern recognition: Methods that search for structures in data* (New York: IEEE)
- Gonzalez R C, Wintz P 1987 *Digital image processing*, 2nd edn (Reading, MA: Addison-Wesley)
- Kandel A 1982 *Fuzzy techniques in pattern recognition* (New York: Wiley Interscience)
- Levi G, Montanari U 1970 A grey-weighted skeleton. *Inf. Control* 17: 62-91
- Marr D 1982 *Vision* (San Francisco: W H Freeman)
- Pal N R, Pal S K 1989a Entropic thresholding, *Signal Process.* 16: 97-108
- Pal N R, Pal S K 1989b Object-background segmentation using new definitions of entropy. *Proc. Inst. Electr. Eng.* E136: 284-295
- Pal S K 1989 Fuzzy skeletonization of an image. *Pattern Recogn. Lett.* 10: 17-23
- Pal S K, Dutta Majumder D 1986 *Fuzzy mathematical approach to pattern recognition* (New York: John Wiley & Sons)
- Pal S K, Ghosh A 1992 Fuzzy geometry in image analysis. *Fuzzy Sets Syst.* 48: 23-40
- Pal S K, King R A, Hashim A A 1983 Image description and primitive extraction using fuzzy set. *IEEE Trans. Syst., Man Cybern.* SMC-13: 94-100
- Pal S K, Rosenfeld A 1991 A fuzzy medial axis transformation based on fuzzy disks. *Pattern Recogn. Lett.* 12: 585-590
- Peleg S, Rosenfeld A 1981 A min-max medial axis transformation. *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-3: 208-210
- Rosenfeld A 1984 The fuzzy geometry of image subsets. *Pattern Recogn. Lett.* 2: 311-317
- Rosenfeld A, Kak A C 1982 *Digital picture processing*, 2nd edn (New York: Academic Press) vol. 2, chap. 11
- Salari E, Siy P 1984 The ridge-seeking method for obtaining the skeleton of digital images. *IEEE Trans. Syst., Man Cybern.* SMC-14: 524-528