B. A Special Case

It is interesting to consider the special case of equal standard deviations, $D_1 = D_2 = \cdots = D_C$ because then the classification rule reduces to a minimum-distance type of rule. In this case, letting D denote the common standard deviation,

$$\begin{aligned}\max_c f(x \mid W_c) &= \max_c (kD^2)^{-IJ/2} \exp\left[-\left\{S + IJ(g - G_c)^2 + (b - B_c)^2 \sum\sum [z(i,j) - \bar{z}]^2\right\}/(2D^2)\right] \\ &= (kD^2)^{-IJ/2} \exp[-S/(2D^2)] \\ &\quad \times \exp\left[-(2D^2)^{-1}\left[\min_c\left\{IJ(g - G_c) + (b - B_c)^2 \sum\sum [z(i,j) - \bar{z}]^2\right\}\right]\right].\end{aligned}$$

Let

$$D(W_c) = IJ(g - G_c)^2 + (b - B_c)^2 \sum\sum [z(i,j) - \bar{z}]^2,$$

$c = 1, 2, \cdots, C$. The quantity $D(W_c)$ can be considered as the square of a distance between (g, b) and (G_c, B_c). The classification rule reduces to classifying the object into class W_c if

$$D(W_c) = \min\{D(W_{c'}), c' = 1, 2, \cdots, C\}.$$

IV. Discussion

The statistic $D(W_c)$ measures how well the class W_c fits the data, compared to the best fit, as achieved by the least squares estimators. Thus the statistics $D(W_c)$ seem reasonable for use in classification in a Markov model even when normality is not assumed.

The presentation in the present paper has been in terms of univariate observations, for a first-order Markov model with equal autoregressive coefficients. The discussion has been given in terms of this simplest case for ease of exposition. Note, however, that allowing for different autoregression coefficients and higher order autoregression causes no essential complication; the solution is still given by the normal equations, providing the least squares estimates for linear statistical models; the solution would be given by multiple rather than simple regression. In this connection, note that the particular, simple model used here for illustration is not necessarily realistic for common images. A more common (and more realistic) model, which is second-order in the sense of allowing dependence on $X(i-1, j-1)$ as well as on $X(i-1, j)$ and $X(i, j-1)$, as alluded to in Section II, is a model of the form

$$X(i,j) = A + BX(i-1,j) + CX(i,j-1) + DX(i-1,j-1) + U(i,j).$$

The particular choice of class of model (causal, semicausal, or noncausal) is important in the sense that the set of statistics sufficient for the classification varies from model to model.

References

[1] A. Habibi, "Two-dimensional Bayesian estimate of images," *Proc. IEEE*, vol. 60, pp. 878–883, July 1972.

[2] M. Hassner and J. Sklansky, "Markov random fields as models of digitized image texture," in *Proc. IEEE Computer Society Conf. on Pattern Recognition and Image Processing*, Chicago, IL, May 31–June 2, 1978, pp. 346–351.

[3] A. K. Jain, "A semi-causal model for recursive filtering of two-dimensional images," *IEEE Trans. Comput.*, vol. C-26, pp. 343–350, Apr. 1977.

[4] A. Jain, "Partial differential equations and finite difference methods in image processing. Part I: Image representation," *J. Optimiz. Theory Appl.*, vol. 23, pp. 65–91, Sept. 1977.

[5] A. K. Jain and J. R. Jain, "Partial differential equations and finite difference methods in image processing. Part II: Image restoration," *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 817–834, Oct. 1978.

[6] D. A. Landgrebe, "The development of a spectral/spatial classifier for earth observational data," in *Proc. IEEE Computer Society Conf. on Pattern Recognition, and Image Processing*, Chicago, IL, May 31–June 2, 1978, pp. 470–475.

A Min-Max Medial Axis Transformation

SHMUEL PELEG AND AZRIEL ROSENFELD

***Abstract*—Blum's medial axis transformation (MAT) of the set S of 1's in a binary picture can be defined by an iterative shrinking and reexpanding process which detects "corners" on the contours of constant distance from $\bar{S}$, and thereby yields a "skeleton" of S. For unsegmented (gray level) pictures, one can use an analogous definition, in which local MIN and MAX operations play the roles of shrinking and expanding, to compute a "MMMAT value" at each point of the picture. The set of points having high values defines a good "skeleton" for the set of high-gray level points in the given picture.**

***Index Terms*—Medial axis transformation (MAT), local MIN and MAX operations, skeletonization.**

I. Introduction

Let S be a subset of a picture, let P be a point of S, and let $D(P)$ be the largest "disk" (or neighborhood of some specified shape) centered at P that is contained in S. We call $D(P)$ a maximal disk of S if it is not contained in $D(Q)$ for any $Q \neq P$. Evidently, S is the union of its maximal disks. The "medial axis transform" (MAT) [1] of S consists of the centers of these disks together with their radii. In digital pictures "disks" are usually approximated by squares, whose orientation depends on the definition of distance in the grid. When the "chessboard" distance $[d((a,b),(c,d)) = \max(|a-c|, |b-d|)]$ is used, the "disks" are upright squares. When the "city block" distance $[d((a,b),(c,d)) = |a-c| + |b-d|]$ is used, the "disks" are an approximation of diagonal squares.

An equivalent definition of MAT uses paths from a point to the boundary. The distance of a point x in S from $\bar{S}$ is the length of a shortest path from x to the complement $\bar{S}$. The MAT can then be defined as the set of all points in S which do not belong to the minimal path of any other point, together with their distances. It has been shown [2] that for digital pictures using discrete distance metrics the points in the MAT are those points whose distances from $\bar{S}$ are local maxima. The MAT can be regarded as a generalized axis of symmetry of a figure, and constitutes a kind of "skeleton."

Several generalizations of the MAT have been proposed, based on these definitions, which allow a MAT to be defined for a gray level digital picture, rather than for a two-valued picture representing a set S (1's at points of S, 0's elsewhere). One generalization [3], the *S*patial *P*iecewise *A*pproximation by *N*eighborhoods (SPAN), is defined in terms of maximal homogeneous disks; the given picture can be approximated if we are given the set of centers, radii, and average gray levels

Manuscript received January 30, 1980; revised March 28, 1980. This research was supported by the Defense Advanced Research Projects Agency and the U.S. Army Night Vision Laboratory under Contract DAAG-53-76C-0138 (DARPA Order 3206).

The authors are with the Computer Vision Laboratory, Computer Science Center, University of Maryland, College Park, MD 20742.

of these disks [5]. If the picture is two-valued, and "homogeneous" means "constant-valued," the SPAN reduces to the MAT. Another generalization, the GRAYMAT [4], is based on the concept of gray-weighted distance: the gray-weighted length of a path is proportional to the sum (or integral) of the gray levels along the path; the gray-weighted distance between two points is the lowest gray-weighted length of any path between them. The GRAYMAT is the set of all points which do not belong to any minimal gray-weighted path from any other point to the zero-valued background, together with the corresponding distance. This too reduces to the MAT in the two-valued case. Still another generalization, the GRADMAT [5], computes a score for each point P of a picture based on the gradient magnitudes at all pairs of points that have P as their midpoint; thus these scores are high at points that lie midway between pairs of antiparallel edges, so that they define a weighted "medial axis."

Each of these generalizations has disadvantages. The SPAN is costly to compute, since it involves testing neighborhoods of all sizes at every point for homogeneity. The GRAYMAT is defined relative to the set of 0's in the picture, so that it requires the picture to be segmented into 0's ("background") and non-0's ("objects"). The GRADMAT turns out to be rather sensitive to noise and to irregularities in region edges.

This correspondence proposes a new gray-scale generalization of the MAT which is inexpensive to compute, does not require the picture to be segmented, and is insensitive to noise. Its definition is based on the fact that the MAT of a set S can be constructed by a process of iteratively shrinking and reexpanding S [6]. For grayscale pictures, the operations of local MIN and local MAX are generalizations of shrinking and expanding, respectively [7]. Thus if we use iterated local MIN and MAX instead of shrinking and expanding, we obtain a "MAT" construction that is applicable to grayscale pictures. The resulting "MAT" will be called the MMMAT (short for "min–max MAT").

Section II reviews the shrink/expand construction of the MAT and defines its min/max generalization. Section III shows that this MMMAT construction yields reasonable "medial axes" in a variety of cases.

II. The MMMAT

The MAT can also be defined by a propagation process starting at the contour of the figure, and propagating toward the inside of the figure. The contour is the initial wavefront of the propagation process, and the propagation velocity is fixed. Wavefront superposition is not allowed, and wavefront intersection points are the points of the MAT. The gray level extension of this "grass fire" definition was given in [6], where the propagation velocity is inversely proportional to the gray level.

Following the above definition, the propagation of a wavefront in a binary digital picture can be modeled by a sequence of "shrink" operations, and the MAT can be constructed by a simple process of iterated shrinking and reexpanding using the appropriate neighborhood (4-neighborhood for the city block distance, 8-neighborhood for the chessboard distance). Let $S^{(k)}$ denote the result of "expanding" S k times, where a single expansion step ($S^{(1)}$) means that all points of $\overline{S}$ which are neighbors of points in S are adjoined to S. Similarly, let $S^{(-k)}$ be the result of "shrinking" S k times; a single shrinking step means that all points of S which are neighbors of points in $\overline{S}$ are deleted from S. Shrinking S is evidently equivalent to expanding $\overline{S}$, and vice versa. A point is in $S^{(-k)}$ if its distance from $\overline{S}$ is at least k; here the distance is city block if we use only horizontal and vertical neighbors in the definition of shrinking, and the distance is chessboard if we also use diagonal neighbors.

It can be shown that for all nonnegative i and j we have $(S^{(-i)})^{(j)} \subseteq S^{(j-i)} \subseteq (S^{(j)})^{(-i)}$; thus in particular, for all nonnegative k we have $(S^{(-k)})^{(1)} \subseteq S^{(-k+1)}$. The difference set $D_k \equiv S^{(-k+1)} - (S^{(-k)})^{(1)}$ consists of points whose distances from $\overline{S}$ are $k-1$, and which have no neighbor at distance k or greater; hence the discrete case D_k is just the set of distance maxima at distance $k-1$ from $\overline{S}$. Thus $\bigcup_k D_k$ is the set of all distance maxima, i.e., of MAT points.

Shrinking S is equivalent to performing a local MIN operation on the two-valued picture that has 1's at the points of S, and expanding S is equivalent to performing a local MAX operation on this picture, where "local" is defined in terms of the appropriate set of neighbors. For a gray level digital picture Σ, let $\Sigma^{(k)}$ be the result of applying k iterations of local MAX to Σ, and let $\Sigma^{(-k)}$ be the result of k iterations of local MIN. It can be shown [7] that for all nonnegative i and j we have $(\Sigma^{(-i)})^{(j)} \leqslant \Sigma^{(j-i)} \leqslant (\Sigma^{(j)})^{(-i)}$; thus in particular, for all nonnegative k we have $(\Sigma^{(-k)})^{(1)} \leqslant \Sigma^{(-k+1)}$, so that the difference picture $\Delta_k \equiv \Sigma^{(-k+1)} - (\Sigma^{(-k)})^{(1)}$ is everywhere nonnegative (all picture operations are performed pointwise). If Σ is a two-valued picture and S is its set of 1's, then the set of 1's of Δ_k is just D_k.

In the two-valued case, when we shrink S, a given point P of S remains unchanged until $k = d(P, \overline{S})$, and then changes to 0; but in the general case, when we iterate local MIN, the value of P may change many times. Let $Z_k(P)$ be the lowest gray level within distance $\leqslant k$ of P; thus $Z_0 \geqslant Z_1 \geqslant, \cdots$, where Z_0 is P's gray level in $\Sigma = \Sigma^{(0)}$. Readily, $Z_k(P)$ is the gray level $\Sigma^{(-k)}(P)$ of P in $\Sigma^{(-k)}$. If $Z_k(P) = Z_{k-1}(P)$, Δ_k must be 0 at P, since the max of $Z_k(P)$ and its neighbors in Δ_k is at least (hence exactly) $Z_{k-1}(P)$; but if $Z_k(P) < Z_{k-1}(P)$, Δ_k may be >0 at P.

The MMMAT value of P can be defined in terms of the $\Delta_k(P)$ values ($k = 1, 2, \cdots$) in several ways. One possibility is to use their maximum; another is to use their sum. As we shall see in the next section, both of these definitions yield MAT-like loci of high MMMAT values. It is evident that the max definition yields values in the same range $[0, Z]$ as the picture's gray scale, since $0 \leqslant \Delta_k \equiv \Sigma^{(-k+1)} - (\Sigma^{(-k)})^{(1)} \leqslant \Sigma^{(-k+1)} \leqslant Z$ for all k. For the sum definition too, we have $0 \leqslant \Delta_k \leqslant \Sigma \Delta_k$. On the other hand,

$$\sum_{k=0}^{N} (\Sigma^{(-k)} - \Sigma^{(-k-1)}) = \Sigma^{(0)} - \Sigma^{(-N-1)} \leqslant Z;$$

and since

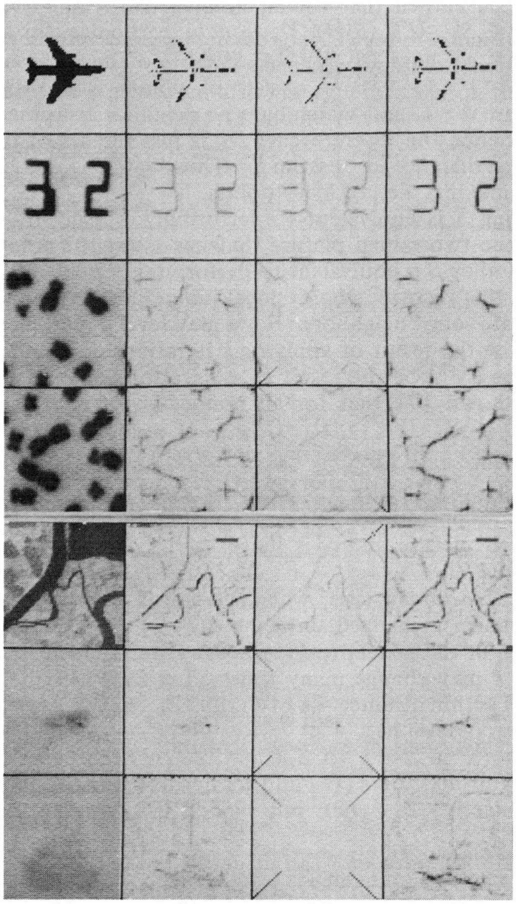
$$(\Sigma^{(-k-1)})^{(1)} \geqslant \Sigma^{(-k-1)},$$

this implies $\Sigma (\Sigma^{(-k)} - (\Sigma^{(-k-1)})^{(1)}) \leqslant Z$.

When the local MIN operation is iterated many times, border effects become a serious problem. In the two-valued case, if we require that S be interior to the picture, then the border of the picture consists entirely of 0's, and we can treat the outside of the picture as consisting of 0's without creating any artifacts. In the gray-scale case, however, whatever value(s) we use outside the picture will have effects on their neighbors inside it, and as the process is iterated, these effects propagate, as we will see in the next section.

III. Examples and Concluding Remarks

Fig. 1 shows eight pictures and their MMMAT's computed in three different ways: max Δ_x using eight- and four-neighbor operations, and $\Sigma \Delta_k$ using eight-neighbor operations. The four-neighbor version contains artifacts due to border effects, resulting from the fact that the outside of the picture is treated as consisting of 0's. In all cases, the high MMMAT values constitute very reasonable "skeletons" of the dark points in the given picture.



(a) (b) (c) (d)

Fig. 1. Some pictures and their MMMAT's: (a) Originals (64 by 64 pixels). (b) 8-neighbor MMMAT's using $\max_k$. (c) 4-neighbor MMMAT's using $\max_k$. (d) 8-neighbor MMMAT's using Σ_k. The number of iterations used in each case is somewhat greater than the radius of the largest dark region.

In the two-valued cases, S can be reconstructed from its MAT by a reexpansion process; in fact, S is the union of the disks centered at the MAT points and with radii equal to the distance values of those points. In the gray-scale case, analogously, if we know $\Delta_1(Z), \Delta_2(Z), \cdots, \Delta_m(Z)$ for each Z, we can reconstruct Σ from $\Sigma^{(-m)}$ by an interated local MAX process, where at each step we add the appropriate Δ value back into the picture. Specifically, given $\Sigma^{(-m)}$, we have $\Sigma^{(-m+1)} = (\Sigma^{(-m)})^{(1)} + \Delta_m$; $\Sigma^{(-m+2)} = (\Sigma^{(-m+1)})^{(1)} + \Delta_{m-1}; \cdots; \Sigma \equiv \Sigma^{(0)} = (\Sigma^{(-1)})^{(1)} + \Delta_1$. Note, however, that this reconstruction process requires a large amount of information, namely m arrays of Δ values, unlike the two-valued case where we only need a single distance value for each point. This is a consequence of the fact that in the MAT construction process, the value of a point changes from 1 to 0 only once (for k equal to its distance from $\overline{S}$), whereas in MMMAT construction, the value of a point may change at every iteration. In any case, the picture cannot be reconstructed from its MMMAT values, since these are maxes or sums of Δ_k's, and we need all of the individual Δ_k values for correct reconstruction.

It has been suggested [1] that biological visual systems compute MAT's and use them to extract perceptually significant features from shapes and patterns (e.g., lobes on a shape correspond to branches on its MAT). However, it seems implausible that visual systems threshold their input, which would be necessary for MAT computation. The MMMAT provides a possible alternative approach in which medial axes can be computed from unthresholded input.

Comparisons with the other methods mentioned in Section I are not given here, but some idea of relative performance can be gotten by comparing our results with those in [3]-[5]. All of the methods produce "disconnected" skeletons, just as the MAT itself does (local maxima of distance cannot be adjacent unless they have the same value). If the MMMAT algorithm is applied for only a few iterations, it yields a "partial" skeleton arising from thin parts of the objects in the image, or extending part way along the bisectors of corners. The time complexity of the algorithm is proportional to the image size times the number of iterations, since each iteration requires another local MIN and local MAX applied to the entire image; on a cellular array (one processor per pixel), the time would be proportional to the number of iterations.

References

[1] H. Blum, "A transformation for extracting new descriptors of shape," in *Models for the Perception of Speech and Visual Form*, W. Wathen-Dunn, Ed. Cambridge, MA: MIT Press, 1967, pp. 363-380.

[2] A. Rosenfeld and J. L. Pfaltz, "Sequential operations in digital picture processing," *J. Ass. Comput. Mach.*, vol. 13, pp. 471-494, 1966.

[3] N. Ahuja, L. S. Davis, D. L. Milgram, and A. Rosenfeld, "Piecewise approximation of pictures using maximal neighborhoods," *IEEE Trans. Comput.*, vol. C-27, pp. 375-379, 1978.

[4] G. Levi and U. Montanari, "A grey-weighted skeleton," *Inform. Contr.*, vol. 17, pp. 62-91, 1970.

[5] S. Wang, A. Rosenfeld, and A. Y. Wu, "A medial axis transformation for grayscale pictures," Comput. Sci. Cent., Univ. Maryland, Tech. Rep. TR-843, Dec. 1979.

[6] J. C. Mott-Smith, "Medial axis transformations," in *Picture Processing and Psychopictorics*, B. S. Lipkin and A. Rosenfeld, Eds. New York: Academic, 1970, 267-283.

[7] Y. Nakagawa and A. Rosenfeld, "A note on the use of local min and max operations in digital picture processing," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-8, pp. 632-635, 1978.

Construction of a Distributed Associative Memory on the Basis of Bayes Discriminant Rule

KENJI MURAKAMI AND TSUNEHIRO AIBARA

***Abstract*—The purpose of this correspondence is to propose a new construction method of distributed associative memory which operates with discrete-valued signals. In this method, memorized pairs of vectors (cue vectors and data vectors) are recorded in the form of a matrix W and a vector T. From an input vector X, the data vector is recalled by an operation $u(XW + T)$ where X is a cue vector or a noisy cue vector and u is a quantizing function. The methods of memorization and recall are similar to the Associatron; however, the proposed model can recall the data vectors optimally in Bayesian sense even when noisy cue vectors are given as the input vectors.**

Manuscript received April 3, 1978; revised March 30, 1980.

The authors are with the Department of Electronics Engineering, Ehime University, Matsuyama, Japan.

0162-8828/81/0300-0210$00.75 © 1981 IEEE