Feature Extraction: A Survey

MARTIN D. LEVINE

Abstract—A survey of computer algorithms and philosophies applied to problems of feature extraction and pattern recognition in conjunction with image analysis is presented. The main emphasis is on usable techniques applicable to practical image processing systems. The various methods are discussed under the broad headings of microanalysis and macroanalysis.

I. Introduction

ANY DISCUSSION of the vagaries of feature extraction and processing leads immediately to rather questionable arguments about the definitions of pattern recognition. This may then develop from the sublime to the ridiculous into various dialogues on the subject of artificial intelligence. The author does not intend to enter the argument here although admittedly it is difficult to restrain oneself. Some illuminating discussions on this topic which contain a large number of references and which might serve as an introduction are found in papers by Greene [33], Kelly and Selfridge [55], Minsky [76], Glushkov [30], Tonge [117] and Uhr [121 (see particularly pp. 291-294 for a comparison of the work of psychologists and computer people)]. The book *Cognitive Psychology* [86] is also of considerable interest. An excellent survey article on pattern recognition with an emphasis on mathematical techniques can be found in Nagy [82].

Munson [79] describes the pattern recognition process in terms of the following three stages where each is considered as an independent component:

TRANSDUCER → PREPROCESSOR → CLASSIFIER

However, considering the first step, "it is doubtful if recognition occurs before the eyes are directed toward a (known) object, since otherwise we would not bother to look at the object" [51]. Thus it would seem to be reasonable to use the raw data in the form of the hard-copy image as bulk memory and allow the transducer to search for "regions of interest." In this way the transduction and preprocessing stages could be integrated with the result being a more meaningful and efficient operation. It is less evident in what manner the last two stages could be associated. One suggestion [56] is to consider both numerical and pictorial data as outputs from the preprocessor and the classifier and introduce a man–machine communication environment (such as a computer graphics display) to provide a link.

Manuscript received September 18, 1968; revised April 23, 1969. The work reported here was jointly sponsored by the National Research Council of Canada under Grant A4156 and the Medical Research Council of Canada under Grant MA-3236.

The author is with the Department of Electrical Engineering, McGill University, Montreal, Canada.

In considering the three stages, the literature overwhelmingly concentrates on the various aspects of classification. In his review of the book by Sebestyen [103], David [12] states that " . . . this monograph reduces recognition to a decision process and concentrates on it from both analytical and algorithmic points of view. It is here that a substantial objection can be raised. Is not the more significant part of the problem that of characterizing the world by a set of properties that provide the desired discriminations? Many of us believe that it is." In fact Selfridge [104] defines pattern recognition solely in terms of "the extraction of the significant features from a background of irrelevant detail." It is emphasized, and this is an important point, that the significance is a function of both context and the experience of the pattern recognizer. On the subject of feature extraction, Nilsson [89] comments that:

1) No general theory exists to allow us to choose what features are relevant for a particular problem.
2) Design of feature extractors is empirical and uses many ad hoc strategies.
3) We can get some guidance from biological prototypes.

These points are supported by Selfridge and Neisser [105]:

> At present the only way the machine can get an adequate set of features is from a human programmer. The effectiveness of any particular set can be demonstrated only by experiment. In general there is probably safety in numbers. The designer will do well to include all the features he can think of that might plausibly be useful.

Considering point 3), Kazmierczak and Steinbuch [54] state that "the human visual system is capable of selecting features or criteria from a pattern where the statement of the description would be independent of registration, skew, size, contrast, deformation, or other noise effects." What is needed, according to Duda [19], are "rugged features." "A rugged feature is one whose presence is not changed, and whose characteristics are not greatly altered, by normal variations in the image of a character in a given category." In this context, the problems of noise, both random and burst, are discussed by Yamada and Fornanago [129]. However the importance of structures, substructures, and relationships among these is also pertinent to feature extraction and therefore pattern recognition [67]. Kolers [58] mentions that "the understanding of perception must look to ordering principles and coding rules that give structure and organization to this input."

With these comments in mind, the object of this paper is to present a compendium and discussion of computer methods and algorithms used in conjunction with image analysis. Techniques of feature extraction and pattern recognition (not pattern classification) which appear in the literature are surveyed. Essentially we have staked out the middle territory that neither involves the initial data transducer or acquisition stage nor the final classification or abstraction stage. Methods of determining the attributes as well as their different types will be discussed. In addition, certain operations of scene analysis [82] and noise filtering are given. Having taken a middle position, it is quite likely that both boundaries may in some cases be overstepped. In fact, it is felt that none of these subproblems should really be treated in isolation and that any one stage should be completely integrated with another.

Completeness is difficult since most of the operations are found in applications articles not specifically devoted to feature extraction. The methods and philosophies are roughly divided into those conceptually utilizing microanalysis (local processing) and those utilizing macroanalysis (global processing) as emphasized by Pfaltz *et al.* [92]. The distinction for particular concepts is often a tenuous one and the classification can well be argued.

Usually we shall consider the transduced image as being an $m \times n$ (most often a square) matrix $T(x, y)$ of points represented by the coordinates (x, y). Associated with each coordinate is a grey level a_{ij} which is an element of the matrix $A(x, y)$. In many cases a_{ij} is defined on a two level or binary scale of black (taken as being a 1) and white (taken as a 0). Operations on the matrix T will transform it into a matrix V with a grey level distribution B whose components are labeled b_{ij}.

II. Microanalysis and Microoperations

A. Smoothing

If one considers real data, then immediately one is confronted with scenes or lines containing noise in the form of the absence or appearance of the signal. It is important in most cases to eliminate isolated islands of black whether they are generated by the actual source of the data or the processes of transduction into a computer compatible format. The problem of eliminating or disregarding large blotches rather than small islands or specks is a difficult one. One approach is to incorporate the connectivity algorithm of Rosenfeld and Pfaltz [99] which has been used to advantage by Reisch [96] in operating on biological data with respect to a method concerned with measuring the internal surface area of the human lung. Complimentary to this situation is the requirement for filling in isolated gaps or notches in otherwise black areas. Special procedures for lines will be considered in Section II-*B*.

Consider first the matrix T with an associated two-level grey scale where 0 and 1 represent white and black, respectively, and suppose that an $m \times n$ window, centered on a_{ij}, can take on either of these two values. If V is the resulting transformed matrix with elements b_{ij} then we can use the following method proposed by Dinneen [15] and later discussed by Unger [123], Yamada and Fornanago [129], and Nilsson [89].

If

$$\sum_{p=-m}^{+m} \sum_{q=-n}^{+n} a_{i+p,j+q} \geq \theta,$$

then

$$b_{ij} = 1;$$

if

$$\sum_{p=-m}^{+m} \sum_{q=-n}^{+n} a_{i+p,j+q} < \theta,$$

then

$$b_{ij} = 0.$$

Here θ is an arbitrary threshold which has a great effect on the resulting transformed picture, and this is amply demonstrated pictorially by Dinneen [15]. A particular choice would have to be made on the basis of an examination of the data.

The $m \times n$ window is usually taken to be a 3×3 symmetrically centered on a_{ij}. The effect of the averaging process is directly dependent on the size of the window. Obviously, small windows looking at local neighborhoods will tend to make smaller changes than larger windows which might destroy some of the important detail. This is analogous to the narrow-band and wide-band filtering of electrical signals. A low-pass spatial filter or averager suggested by Graham [31] and Forsen [24] could be used to produce a picture without any sharp gradations in brightness or grey level variation. This operation yields the analogy of a picture slightly out of focus, which gives the general trends but dampens out the sharp detail. The extent of this is governed by the window size of the filter.

Although the method described above is easy to apply and usually works well, in general it would be desirable to invoke an adaptive technique with a window size and threshold depending on the neighborhood properties of the scene. Also one is usually considering pictures with many grey levels so that the above algorithm is not directly applicable. Smoothing procedures for this more interesting situation have not been discussed in the literature. One way of handling the threshold problem has been suggested by Reisch [96]: the picture matrix is decomposed into many submatrices and an average grey level is determined for each area. A threshold proportional to this average is then utilized as a basis for comparison.

B. Filtering Operations on Lines

Since a large proportion of the literature in pattern recognition is concerned with character recognition, both hand and machine printed, a considerable amount of the discussions on possible operations involves lines or line segments.

These transformations have applicability in other areas where one is interested mainly in contour information. Examples of this are the discussions on line drawings [129] and the interesting work on the representation of 3-dimensional figures by Roberts [98].

Many routines for smoothing data appear in the numerical analysis literature [87] but these are generally too complicated. The main criteria in making a choice in a particular case are simplicity and the amount of computer time required. This latter aspect is especially important for pattern recognition performed in real time. Groner [35] has considered such an application involving handprinted text and used the following operation:

$$x_S(k) = \tfrac{3}{4}x_S(k-1) + \tfrac{1}{4}x_R(k)$$
$$y_S(k) = \tfrac{3}{4}y_S(k-1) + \tfrac{1}{4}y_R(k)$$

where

$x_R(k), y_R(k) \triangleq$ coordinates of the kth sample raw data point,

$x_S(k), y_S(k) \triangleq$ coordinates of the kth sample smoothed data point.

Another essentially similar approach introduces the concept of mechanical backlash in terms of a threshold θ. As part of a reading machine for the blind, Mason and Clemens [70] have invoked the following:

If

$$x_S(k-1) < x_R(k) - \theta$$

then

$$x_S(k) = x_R(k) - \theta;$$

if

$$x_R(k) - \theta \le x_S(k-1) \le x_R(k) + \theta$$

then

$$x_S(k) = x_S(k-1);$$

if

$$x_R(k) + \theta < x_S(k-1)$$

then

$$x_S(k) = x_R(k) + \theta;$$

a similar operation is used for the y direction.

Smoothing has the effect of eliminating noise and reducing sharp variations. In order to limit the amount of information to manageable proportions, it is often necessary to use thinning as well. This will also tend to dampen the effect of small perturbations. In any event, the thinned representation will usually describe the essential aspects of the picture, particularly in the case of characters or sketches. Groner [35] first applies a smoothing operation followed by thinning. The position of the previous point in a thinned track is compared with the position of the latest smoothed point, and if these are sufficiently apart, then the smoothed data point becomes an element of the thinned track. Thus

If

$$|x_S(k) - x_T(l-1)| \ge \theta$$

and/or

$$|y_S(k) - y_T(l-1)| \ge \theta,$$

then

$$x_T(l) = x_S(k)$$

and

$$y_T(l) = y_S(k)$$

where $(x_T(l), y_T(l))$ are the coordinates of the lth thinned data point and θ is a threshold.

The concept of connectivity was introduced by Sherman [106] as a basis for a thinning operation and later applied by Minneman [75] to the problem of character recognition (not in real time). The philosophy is to discard an element a_{ij} if it does not result in a line being shortened or if it does not create a gap in a string of connected 1's. A 3×3 window was used [106]. Note that this operation is more widely applicable than that of Groner [35] since the lines in this case have a finite thickness which must also be thinned.

C. Contour Tracing

Smoothing and thinning is one approach to reducing the amount of data in a picture. Another alternative involves scanning a picture and tracing a contour or outline of a figure and then basing the recognition or classification decision on this information [103]. It is well known [61], that "contours carry a significant fraction of the information required for recognition of image objects." Examples of this approach applied to character recognition are discussed in the literature [123], [110], [32], [18], [9]. The tracing techniques may naturally also be used as a first step in the coding of line data and an obvious application of this is the coding of architectural drawings. The second step could then possibly involve the chain encoding schemes of Freeman [25]–[28]. The advantage of using a contour description is that the latter is independent of shape, translation, size, and rotation [52].

Let us consider the tracing algorithm [31] used to achieve efficient image transmission and communication. A raster scan is made in order to find a point a_{ij} whose grey level exceeds some threshold. The latter will define which elements do and do not belong to the contour. A 3×3 window is centered on this point, and the next point on the contour is chosen as that point (out of seven possibilities) not already on a contour, having the maximum value on the grey scale. The contour point is labelled and other points not already on a contour or not immediately connected to the new point are set to zero. When no next point is found, the contour is

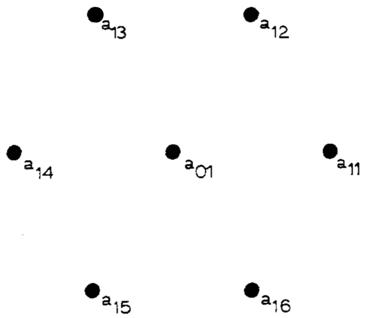

Fig. 1. An hexagonal array.

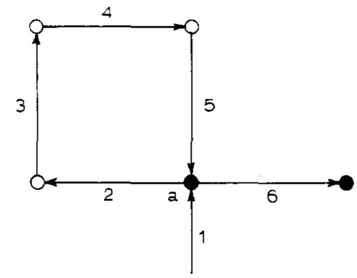

Fig. 2. A typical scan pattern for the line follower of Mason and Clemens [70].

terminated. This operation tends to sharpen the contour and acts as a differentiator.

A simpler situation arises with images defined by a two-level grey scale. Kasvand [52] employs a hexagonal window as shown in Fig. 1. The patterns studied were artificially generated and therefore the method requires further testing on real data. The logic states that $a_{01}=1$ and $\Sigma_{j=1}^{6} a_{ij} \leq 5$, then the central point is labelled as being on the contour. A slight modification would make this useful for a rectangular grid.

The work of Mason and Clemens [70] on an experimental reading machine for blind people involves the use of an edge tracer as a preliminary to the coding of the positions of the vertical and horizontal extremities of a particular character. The tracer proceeds from one grid point in the raster to a neighboring point, continuously moving right or left until some neighborhood of the starting point of the contour is reached. A choice is made to turn right after a white point is encountered and left after a black point as shown in Fig. 2. Note that because of noise effects it is possible that upon return to point *a*, a white reading could be made which would result in the scanner continuously running around the loop. To avoid this situation, one may use the ad hoc rule that after three successive similar turns (either right or left), the fourth one is automatically chosen to be the other. This method is more attractive than those described above in that the logic is extremely simple and therefore the algorithm will consume less computer time.

An interesting line scanning system which involves man–computer communication via a graphic console is discussed by Krull and Foote [60]. With this system, ambiguous situations in which the tracing routine may find itself, are resolved by the intervention of a human operator. The scanner attempts to follow a line by "looking ahead" using the mechanism of "scan feeler vectors" of a given arbitrary length. This type of man–machine interaction should be considered as an asset and not a failure of the program. It is idealistic to expect the computer to perform as well as the eye–brain system of the human.

A similar concept is invoked by Guzman-Arenas [36] who uses a "box" consisting of three parallel rectangles as shown in Fig. 13. "The innermost rectangle is termed the acceptance box; the outer two rectangles are collectively termed the looking box." This is a line follower with learning capabilities in that the width and length of the box are "functions of the length of the portion of line already found and of the noise present in the picture." Conceptually this method is superior, but its performance would be considerably enhanced by the use of analog techniques.

D. Line Drawings and Sketches

Not very much has been reported in the literature on the subject of feature extraction from line drawings, sketches, or cartoons. However, the discussions on contour tracing in Section II-*C* and filtering operations on lines in Section II-*B* are of direct interest here. An interesting and promising area for investigation is the relationship between this problem and computer graphics languages.

The pattern recognizer devised by Harmon [38] recognizes line drawings of geometrical figures such as circles, triangles, pentagons, and hexagons. A dilating circular scan is used and a plot of the radius versus position in each ring is derived. This expanding array or wavefront approach results in a recognition independent of the figure's rotation, size, or registration. Geometrically similar figures result in similar transformations. These concepts will be discussed further in Section III-*C*.

Uhr and Vossler [119] briefly mention the recognition of cartoons. Use is made of a program initially developed for character recognition and applied to two examples of two different facial cartoons plotted on a 20×20 grid.

A short discussion from the point of view of what features are of interest is presented by Yamada and Fornanago [129]. Although no actual experiments are reported, it is suggested that in order to recognize a face it is necessary to separate out the various parts such as the ears, eyes, nose, mouth, and neck. This could be done by recognizing nodes, branches, vertices, and terminations.

Another associated (and probably more significant) problem involves the recognition of actual photographs of scenes of objects. McCormick [71] presents an articular or linguistic approach and this will be discussed in more detail in Section III-*D*. He suggests that the image may be idealized into a line drawing by using local operators of the type already discussed and then producing a linear graph description of the object in terms of the characteristics mentioned by Yamada and Fornanago [129]. This will lead to a syntactic model of the visual description. Narasimhan

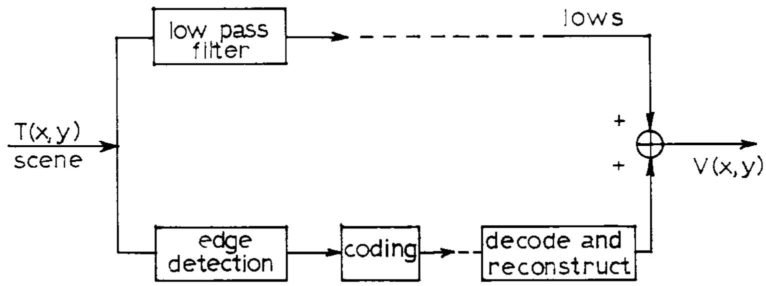

Fig. 3. An image communication system incorporating separate transmission of high- and low-frequency content.

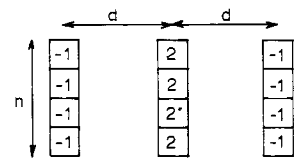

Fig. 4. A typical receptor configuration employed by Deutsch [14] in his short line extractor neuron (SLEN).

and Fornanago [84] have also approached this problem and have produced a "sketch" from a photograph by first retaining only points above a certain threshold grey level and then retaining the boundaries between regions in the remaining grey levels above the threshold. "The major problem is determining the actual values of the parameters that enter into the labelling algorithms for a given class of pictures [84]."

The work of Roberts [98] is interesting in this context although it does not directly involve pattern recognition. First a local differential operator is used to produce a line drawing from a photograph. Then by comparison with standard polygon structures employed as models (only three are used), a 3-dimensional object list consisting of model structures, and transformed and compound versions of the models, is prepared from the drawing. With this description, it is then possible to rotate and translate the object in three dimensions and view its 2-dimensional projections.

E. Edging

The properties of the visual system as they affect communication systems have been discussed by Graham [31] and Schreiber [101]. Efficient transmission of 2- and 3-dimensional scenes are of concern in this context. Typically a 2-dimensional image contains structures and substructures which are delineated by sharp changes in grey level or brightness: we may refer to this data as high frequency in nature. It has been shown that the visual system subjectively enhances sharp differences in grey level [31]. Generally, a scene will also contain broad areas of relatively even texture or brightness. This low-frequency data can be transmitted with fewer samples than the high-frequency information and Graham [31] suggests a method of communication shown in Fig. 3. It is not meant, however, to imply that for the purposes of pattern recognition it is sufficient or even reasonable to separate these two types of data.

Because of the simplicity of concept from the point of view of a computer algorithm and because there is evidence that the mamalian visual cortex behaves in such a way as to detect straight edges [46], [47], [61], [14] the edge feature has been used as a basis for many pattern recognition schemes. Although the latter thesis is of intrinsic interest, it does not directly bear upon the topic under discussion. Readers are referred to the review paper of Chung [10] on the electrophysiology of the visual system and to Part IV of the book by Uhr [121] where some results of neurophysiology and psychology are discussed.

Deutsch [14] has applied the hypothesis of Hubel [47] and conceived of a SLEN (short line extractor neuron) which acts as an optical edge detector. The SLEN (see Fig. 4) is constructed of a column of ON receptors paralleled on both sides by a column of OFF receptors. This model will produce zero output if it is subjected to uniform excitation (the sum of the weights equals zero) or if it is convolved with long and thin lines orthogonal to the receptor columns. Note that the maximum output occurs when the SLEN is presented with a long column parallel to itself. Deutsch [14] has considered the simplest SLEN ($n=2$, $d=1$) as shown in Fig. 5 and arranged a grid of horizontal and vertical receptor groups to be used in order to extract features in the recognition of digits. A similar type of window edge detector is discussed by Nadler [80], Hawkins [39], and Munson [78] and is shown in Fig. 6. The representatives of this class of operators are essentially correlators of local phenomena. Note that if the data is noisy, and especially if the edges are irregular and occur at various angles, the choice of the dimensions for this window presents a difficult problem. These correlation techniques will be discussed further in Section II-*G*.

Another approach is to consider an $n \times n$ aperture array. Kirsch *et al.* [56] discusses a man–computer environment

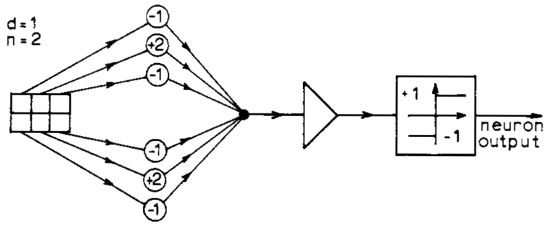

Fig. 5. The simplest SLEN.

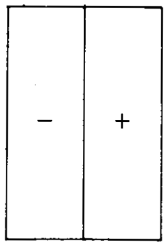

Fig. 6. A window edge detector.

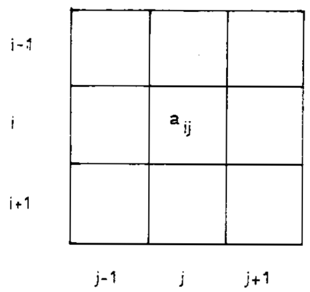

Fig. 7. Using the above 3×3 window, two-dimensional differentiation can be achieved with the following operator:

$$\Delta x = a_{i,j} - a_{i,j-1}$$
$$\Delta y = a_{i-1,j} - a_{i+1,j}$$
$$\text{grad at } (i, j) = |x| + |y|.$$

which would allow for the consultation of and reference to various subroutines for testing different processing ideas. After a computer manipulation of the data using a particular algorithm, the information could be displayed in either numerical or pictorial form. A computationally simple example of one of these operators is suggested as a means of calculating the first derivative of a pattern: if all n^2 grid points of the array are black, then the central point is made white.

In a more complicated vein, Dinneen [15] has produced an interesting algorithm which was later discussed by Nilsson [89]. Although the method requires a greater amount of computer time, it is more powerful than the window methods discussed above. Consider an $n \times n$ (n odd) window such that the element a_{ij} is centered on a black square. Start at element $(i-(n-1)/2),\ j-(n-1)/2))$ and move around the square "ring." When a 1 (or black) is observed, the three diagonally opposite elements in the ring are examined. If these are all 0 (or white), count 1, otherwise ignore. The complete window is checked ring by ring with the exception of the center element. The total count N is compared with a threshold θ: if

$$N > \theta, \qquad b_{ij} = 1;$$
$$N \leq \theta, \qquad b_{ij} = 0.$$

Dinneen [15] suggests that this operation may be made adaptive by setting θ equal to a proportion of the total number of 1's in the window. It is also possible to weight each one of the "rings" with a different value.

Obviously, an edge detector is essentially performing 2-dimensional differentiation. A simple operator to perform this task is described by Holmes *et al.* [42], Forsen [24], and Graham [31] and shown in Fig. 7. An alternate and better [42] version uses the following equations to calculate the perturbations:

$$\Delta x = \tfrac{1}{3}(a_{i-1,j+1} + a_{i,j+1} + a_{i+1,j+1}) - \tfrac{1}{3}(a_{i-1,j-1} + a_{i,j-1} + a_{i+1,j-1})$$
$$\Delta y = \tfrac{1}{3}(a_{i-1,j-1} + a_{i-1,j} + a_{i-1,j+1}) - \tfrac{1}{3}(a_{i+1,j-1} + a_{i+1,j} + a_{i+1,j+1}).$$

These relations differ from the others discussed above in that they are applicable to pictures described by a grey level scale greater than two.

Another similar approach uses the Laplacian operator. If $A(x, y)$ represents the grey level distribution, then we have [61], [31]:

$$\nabla^2 A(x, y) = \frac{\partial^2 A(x, y)}{\partial x^2} + \frac{\partial A(x, y)}{\partial y^2}.$$

An interesting application of this operator is discussed by Hawkins [40] in connection with a method of parallel processing of scenes using electrooptical techniques.

F. Shape and Curvature

Kolers [58] has stated that "... inflection points on a contour are its informationally richest parts." It is also interesting to observe that the visual system of the frog contains convexity detectors which respond only to objects which exhibit curvature in the receptive field [10]. Because this has been recognized by many workers, curvature has often been used as a major feature in pattern recognition problems. For instance, Kazmierczak [53] represents characters using 2-dimensional fields of flow and then extracts shape descriptors such as gross convexity open towards the left or right. In lieu of this binary criterion, Kasvand [51] suggests the quantization of the curvature of a contour or edge.

An interesting application of this property is its use in detecting overlapping objects. Rintala and Hsu [97] define a cell as "an enclosed region or a group of enclosed regions surrounded by a smooth self-intersecting line." The avoidance of sharp curvature or the concept of "continuation of smoothness" is invoked in order to detect overlapping cells in an artificial image. A similar approach is employed for the pattern recognition problem of touching and overlapping chromosomes described by Ledley *et al.* [66]. Examples of the application of this method to scanned data are given in the reference.

Measuring curvature as a continuous variable along a given edge or curve leads to an excessive amount of resolution and requires more bits than would a discrete scale for the representation. It is usually hypothesized that general trends or gross changes of direction are of prime importance. Freeman [25]–[27] has suggested an eight direction scale (shown in Fig. 8) to quantize the directions along a curve which is assumed to be tracked by a line follower. This method of assigning directions has also been used by Tomita and Nouguchi [116] in their work on the recognition of handwritten Katakana characters. Based on this concept, Freeman [25]–[27] has devised a method of chain encoding and has demonstrated many interesting and potentially useful properties of the code. The technique has been used by Rintala and Hsu [97], as well as Freeman [28] in his unusual work on apictorial jigsaw puzzles. In order to represent the planar puzzle pieces, the latter has defined chainlets as portions of a larger chain defined by curves between slope discontinuities or inflection points. The shape of the pieces is described in terms of various features associated with the properties of the associated chains and chainlets. An alternate method of representing the shape of an object is skeletonization which will be discussed in detail in Section III-*C*. Pfaltz and Rosenfeld [91] make a comparison between the relative merits of this approach and that of chain encoding.

In his work on the real-time recognition of handprinted material, Groner [35] quantizes the curves into only four directions: up, down, left, and right. As the person "writes" (on a RAND tablet), the track is smoothed and thinned in order to reduce the amount of unnecessary data. One of the four quantized directions is chosen given a new point (the jth) in the previously determined thinned track (x_T and y_T represent the coordinates of the thinned track):

1) If $|x_T(j) - x_T(j-1)| \geq |y_T(j-1)|$
   a) direction is right if $x_T(j) - x_T(j-1) \geq 0$
   b) direction is left if $x_T(j) - x_T(j-1) < 0$.
2) If $|x_T(j) - x_T(j-1)| < |y_T(j) - y_T(j-1)|$
   a) direction is up if $y_T(j) - y_T(j-1) > 0$
   b) direction is down if $y_T(j) - y_T(j-1) < 0$.

If using these criteria at two successive points along the thinned track results in the same direction and if this direction differs from the last listed in the stored sequence, it is added to the list. If the test fails, the measurement is not considered any further.

Kasvand [52] discusses the importance of curvature as a descriptor and uses a hexagonal array (shown in Fig. 9) as a basis for his operations on monochromatic 2-dimensional objects. If $a_{01} = 1$, which means that the assembly is on a contour, then define

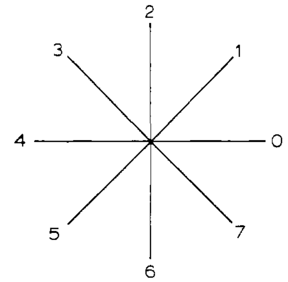

Fig. 8. An eight level scale for quantizing directions along a curve.

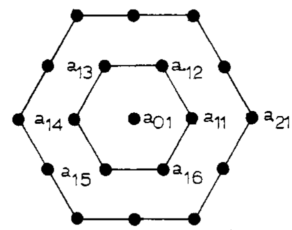

Fig. 9. An hexagonal array with two layers.

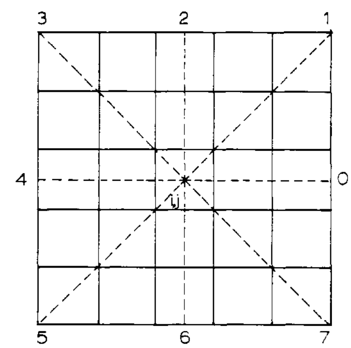

Fig. 10. The quantized directions in the 5 × 5 array employed by Yamada and Fornanago [129].

$$C = a_{01} + \sum_{j=1}^{6} a_{ij} + \sum_{j=1}^{12} a_{2j}.$$

It can be shown that C is directly (but nonlinearly) related to the angle of curvature and this is then used as the basis for the coding. No account is made for the effects of noise or otherwise imperfect data. However, this method is promising in that it is simple and does not require the services of a curve follower.

Yamada and Fornanago [129] quantize and label points in a scene according to a scale of eight directions of a line through the point. They consider the array shown in Fig. 10. The point (i, j) is given the labels d and $d+4$ if (i, j) and the associated second neighbors $((i-2, j-2), (i-2, j), (i-2, j+2), (i, j-2), (i, j+2), (i+2, j-2), (i+2, j), (i+2, j+2))$ along the directions d and $d+4$ are also black. Using a 3×3 array, after operating in this way on the complete picture, gaps are filled by labelling a point in the d direction if it has a d neighbor and either a $d+3$, $d+4$, or $d+5$ neighbor labelled in the initial processing; however new labels are not assigned to points already labelled. Some interesting examples of this procedure are presented. Although a follower is not used, this method would require more computer time than that of Kasvand [52].

G. Correlation Methods

The techniques of correlation in two dimensions can be used to compare a standard or representative object or part of an object with a given scene. Given a file of standard objects in the computer memory, a scene is correlated with each member of the file and it is concluded that it most resembles that object for which the correlation is a maximum. This type of approach is often termed "template matching" or "window methods" and has been suggested as a basic component of a model for the visual processing system [49]. Note that the correlation may also be performed using optical methods [125], [112].

Although seemingly straightforward in concept, these techniques of either digital or optical matching present certain difficulties in their application [41], [103], [105]. Sensitivity to translation, magnification, brightness, contrast, and orientation are characteristic of this approach [103]. In order to achieve a match with a stored reference item, it is necessary to prenormalize the image and eliminate these effects. This is often difficult since different examples of the same type of object may not really have similar shapes. In fact, even if they do, slight variations may result in poor correlation with the stored image. Highleyman [41] has suggested two methods for reducing position sensitivity in the case of character recognition. One approach aligns the center of gravity with a reference point while in the other the object is transformed with respect to the reference image until a maximum correlation is achieved. The former technique is faster and does not allow for the possibility of convergence to an incorrect extremum. General transformations of this type will be discussed in Section II-*I*.

If the dimensions of the objects under consideration do not vary to any great extent, it is theoretically possible to use window methods to extract the features. An interesting approach would be to design a window operator whose dimensions varied according to the local trends in the data; the line follower of Guzman-Arenas [36] is a good example of this.

The detection of either horizontal or vertical edges can be achieved using the operator shown in Fig. 6 [80], [39], [78], [89]. Considering a two-level grey scale, the sum of the black (or 1) elements on the left-hand side is subtracted from the sum on the right side and the result compared with a suitable threshold. If the difference exceeds the threshold, an edge has been detected. With reference to Fig. 11, note that the correlation of this aperture function when compared with an object exhibiting a certain radius of curvature increases monotonically with the radius and decreases with the angular rotation of the edge or border [39].

Another type of window or mask may be used to detect right-angle corners. A slight variation of the mask shown in Fig. 12 and suggested by Nilsson [89] is mentioned by Duda *et al.* [19] in a discussion on character recognition. It is questionable whether this type of aperture could be used to detect other edge orientations.

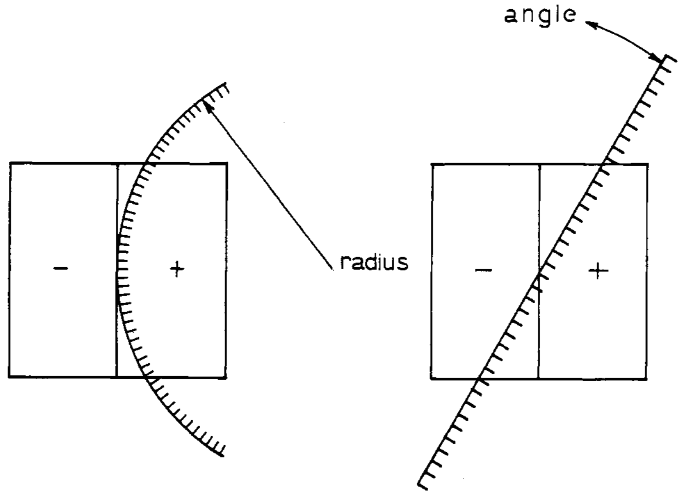

Fig. 11. The effect of orientation on correlations with the window edge detector.

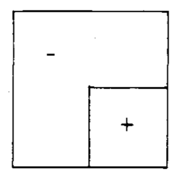

Fig. 12. A corner detector.

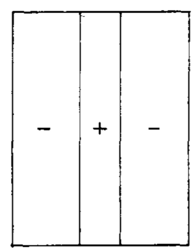

Fig. 13. A line segment detector.

Nilsson [89] has suggested a window, shown in Fig. 13, which could be used to detect line segments. Also the SLEN (short line extractor neuron), hypothesized by Deutsch [14] as a model for visual pattern recognition, uses this type of aperture as a basic building block. The two major shortcomings of this technique are that the width of the segments must be known a priori and the segments must have relatively similar dimensions. Difficulties arise if the line segments, in addition, occur at various orientations [96].

Other aperture functions suggested by Nilsson [89] can be used to detect spots or line intersections as shown in Fig. 14. No examples of their application appear in the literature. In general (as discussed by Reisch [96]), although

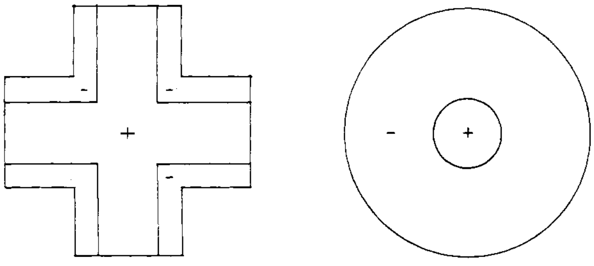

Fig. 14. Window detectors for determining (a) intersections of lines at right angles to each other and (b) spots of a given radius.

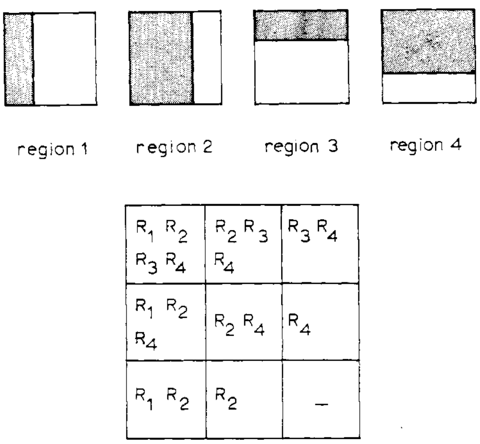

Fig. 15. Regions used to construct a property vector for the recognition in real time of handwritten characters [115].

these techniques are conceptually attractive, they are not useful for practical applications.

The methods of digital correlation or template matching have been applied to character recognition and have aroused controversy as to their usefulness [120]. Casey and Nagy [8] consider the interesting problem of the pattern recognition of Chinese printed characters while Highleyman [41], Horwitz and Shelton [44], and Clayden *et al.* [11] discuss the usual Arabic characters.

Straightforward correlation may also be used to extract local features. In his interesting discussion of a mobile vehicle controlled by a computer and capable of visual observations, Forsen [24] suggests the following feature extractor: "At each picture element location (7×7 window) we assign the feature of a set of computer generated features that produces the highest 'correlation' with the data in the neighborhood of that element." Eight sample features are used to describe the differentiated representation of the TV scan of an object. Although successful for simple objects, the method is inadequate for more complicated images such as a typical office scene shown in the reference. Instead of prescribing a set of features, Uhr and Vossler [119] have used a 5×5 matrix with a given number of randomly generated feature patterns or windows. Bledsoe and Browning [2] invoke 75 pairs of elements chosen at random from the scene matrix. These are determined for various letters of the alphabet and used for correlation purposes. This is equivalent to treating the exhibited pattern in the pair (four states are possible) as a feature. A similar approach has been proposed by Teitelman [115] for the recognition in real time of handwritten characters where a property vector is constructed which describes the position of the pen at time t according to the four regions (in a 3×3 matrix) shown in Fig. 15. However, here the method is enhanced as a result of the availability of the temporal information.

Finally, a very promising approach using a correlation technique does exist and is suggested and demonstrated by Hawkins [40]. The employment of an image intensifier tube achieves parallel processing of a picture in very short time intervals. Some striking results of the use of this method, which is basically the correlation of one image with another, are presented by the author.

H. Connectivity

Julesz [48] discusses the importance of proximity and uniformity as criteria for visual pattern discrimination of clusters of points or lines. These ideas and their relation to the perception of patterns and the coding of data, are referred to by Kolers [58] who stresses the principle of grouping by similarity, proximity and good continuation. A serious problem in applying any connectivity algorithm is the degree of noisiness in the data [83] since noise in the form of gaps may destroy the basic connectivity of lines or objects in a scene. One way of circumventing this difficulty is by averaging or smoothing the data before processing for connectivity [96]. Alternately, the use of larger windows or arrays would facilitate the process of jumping across small gaps or holes [89]. A locally adaptive procedure for adjusting the array size could be used for this purpose.

Nilsson [89] defines an isolation operation which is complimentary to that of connectivity. This transformation selects an object or a connected object in a picture and eliminates all other points not connected to it. A prescan chooses a black cell about which is centered an $n \times n$ window and black cells inside this region are labelled and retained. These, in turn, are used as centers of other windows and the labelling operation is continued until no new black points are encountered. Any points not labelled are set equal to zero.

The concept of connectivity has been applied to problems of character recognition [106], [78], [116]. The characters are approximated by a set of lines or line segments joining the nodes of structures where the nodes are chosen to be line endpoints, intersection points, and bend points. A connectivity matrix is then constructed and used as the basis for the pattern recognition. Rosenfeld and Pfaltz [99], in a comprehensive and significant paper, discuss a connectivity operation which incorporates sequential processing of the image matrix. The algorithm uses a 3×3 window and labels each connected set in the overall picture with a different symbol. Reisch [96] has used this program to identify and isolate alveolae walls in the human lung which are known to form a connected surface. In this way, unwanted noise can be ignored since the latter and the walls will be assigned different labels. Another interesting application has been presented by Yamada and Fornanago [129], who have grouped segments of lines having the same direction using connectivity.

It should be noted that the contour tracing algorithms discussed in Section II-*C* are implicitly invoking the principle of connectivity.

I. Moment Transformations

It is possible to use various transformations to facilitate shape recognition on the basis of a normalized representation of an object. These techniques generally assume that an object forms a connected set and has somehow been isolated in the context of a whole scene. Since this problem of object isolation is a major one, the applicability of this approach is restricted to reasonably well defined pattern classes. Hu [45] has presented a theory of 2-dimensional moment invariants for geometric shapes defined on a plane. A complete system of moment invariants under translation, similitude, and orthogonal transformations are derived. However only the first moment, which yields the centroid of an object, seems to have been incorporated into any pattern recognition application.

Kirsch *et al.* [56] has suggested and Tsai and Sheng [118] have used the center of gravity of the pattern as the origin of a polar coordinate system. If the picture contains $m \times n$ cells, then the centroid is defined by $(\bar{x}, \bar{y})$ such that

$$\bar{x} = \frac{1}{s} \sum_{j=1}^{m} \sum_{k=1}^{n} x_j a_{jk}$$

$$\bar{y} = \frac{1}{s} \sum_{j=1}^{m} \sum_{k=1}^{n} y_k a_{jk}$$

$$s = \sum_{j=1}^{m} \sum_{k=1}^{n} a_{jk}$$

where

$$a_{jk} = 0 \qquad \text{for white cells}$$

$$a_{jk} = 1 \qquad \text{for black cells.}$$

A plot of the logarithm of the radius to the contour of the object versus the angle is then used as a basis for the recognition scheme. These curves are shown to possess some interesting properties:

1) The curve is a periodic function of period 2π.
2) A translated representation of the objects results in an identical curve.
3) The effect of rotation of an object manifests itself in a leftward or rightward shift of the curve.
4) Two objects with identical shapes but of different size will be shifted vertically from one another.
5) The resulting transformed curve may be a multi-valued function of the angle.

No comparisons are made with object skeletonization (Section III-*C*) or the method of chain encoding (Section II-*F*) and artificial data was taken as the example.

The use of such a transformation presupposes that a scanning routine exists which can isolate the objects of interest. An obvious situation where this is so is character recognition. Highleyman [41], Minneman [75], and Spinrad [109] have normalized the characters by choosing the centroid as an arbitrary reference point. Another application is the photo interpretation of aerial photographs [42], [108]. In particular, Smith and Wright [108] are interested in automatic ship identification and have considered the i–jth generalized moment:

$$m_{ij} = \int\int f_{ij}(x, y) A(x, y) dx dy$$

where $f_{ij}(x, y)$ is some function of x and y and $A(x, y)$ may be chosen to be $x^i y^j$, Chebyshev polynomials in x and y, or sine and cosine functions of ix and jy. In the latter case the m_{ij} are the coefficients of a 2-dimensional spatial Fourier series expansion of the image. Template matching is a special case of the method of moments if $f_{ij}(x, y) = 1$ or 0 for $A(x, y)$ equal black or white, respectively. Note that the higher order moments can be made invariant to certain effects by normalization with respect to the lower order moments.

J. Local Topological Properties

Topological properties are attributes of images which would usually be invoked by humans when describing these objects. "... these features are primarily concerned with the geometrical and topological components and relationships of the character as a whole [19]." The primary application has been in the field of character recognition [9] where the main concern is line drawings.

The obvious feature of concern in a character is the number and position of any extremities [53], [113], [32], [18], [78], [116], [70]. These might be roughly positioned as occurring at the top, bottom, right, or left or may just be considered as nodes for a connection matrix. More detailed information about the actual strokes or line segments that make up a figure may be desirable [16], [21], [113], [109], [78], [116]. Closely related to these measurements is the detection of corners [53], [78] or curved shapes open to either the left or right [53]. A multitude of other properties have been used [21], [32]: bars, hooks, arches, loops, bays, arcs, notches, lakes, etc. Unger [123] lists 36 topological properties used by a spatial computer for the purpose of character recognition. Two interesting features, concavities and enclosures, have been suggested by Munson [78]. The concavities are found by first determining the outside contour and then the convex hull and are those connected regions adjacent to both the figure and the hull as shown in Fig. 16. The enclosures are defined as those connected regions of ground touching the figure but not the convex hull. An example is shown in Fig. 17. The

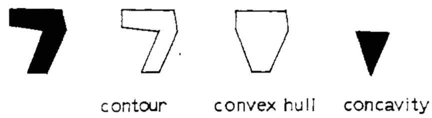

Fig. 16. A concavity operator as proposed by Munson [78] applied to the character "7."

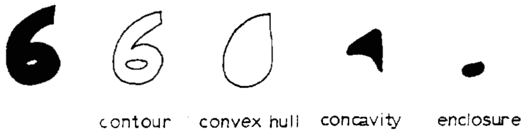

Fig. 17. The determination of the enclosures in the character "6" using a method suggested by Munson [78].

real-time recognition of handwritten characters has also involved the detection of corners and various strokes, [73], [35].

K. Feature Finders

Although in general it would seem to be preferable to choose features by a careful study of the data to be processed, some techniques have been presented which either generate their own features or use random ones. These have been restricted to very small image matrices and are of limited applicability.

Muchnik [77] describes a method for "forming the simplest features" and hypothesizes that they may be used as a basis for constructing more intricate features. He defines the term "informative fragments" as portions of an image which can be described in terms of simple words such as corners, curvature, and intersection. A major assumption is that the fragments are spaced well apart. An arbitrary reference window with weights, as shown in Fig. 18, is used to determine the value of a distance function ("informative function") for a given fragment of the overall picture. Using a gradient search technique, a local extremum of the distance function is found and thus isolates an "informative fragment." Hence the data will yield many reference attributes which may then be organized using a clustering program. Similarly, a predictive coding experiment, concerned mainly with the transmission of pictorial data, is presented by Wholey [128]. A statistical survey of a representative class of weather maps, traced onto a 70×100 matrix (7000 elements), was made to determine the most likely grey level (black or white) to follow an aperture with a particular grey level distribution. Given the statistics and an image to be coded, an error matrix can be generated which can later be used to regenerate the original picture. The method did not prove to be very successful. Kamentsky and Liu [50] discuss a design and search procedure for multifont print recognition logic, which is also based on an analysis of the statistics of the characters which are representative of those to be used. The computer program chooses the best logic configuration where the latter operates on a subset of the scanned image matrix and therefore represents the attributes which are employed as an input to the classification stage. Digital masks are used to restrict the number of points considered in the image to 64.

1	1	1	1	1	1
1	2	2	2	2	1
1	2	3	3	2	1
1	2	3	3	2	1
1	2	2	2	2	1
1	1	1	1	1	1

Fig. 18. A reference aperture with weighted elements used by Muchnik [77] to detect "informative fragments."

The results for numerous experiments were very good and it was found that the percent recognition could be related to the required number of logic circuits.

Another approach is to utilize random patterns. Uhr and Vossler [119] generate these in a 5×5 aperture and then invoke correlation to determine whether the patterns exist in the image. Bledsoe and Browning [2] choose 75 pairs of elements at random each of which may take on either one of four states. The characters are analyzed in this way and correlation then used.

It is possible to find representative features or attributes using the concept of learning machines [88], [7], [1]. Block *et al.* [3] incorporates a perceptron in the pattern recognition scheme and states the problem as: "Given a set of patterns, determine a set of features, minimal in number, such that each pattern can be formed by the superposition of a subset of these features." A 5×5 pattern array was chosen.

Generally speaking, it is not clear how the techniques discussed above could be extended efficiently to handle problems with large and complicated images of high resolution.

L. Retinal Maps

Chung [10] states that "the pattern of optic activity does not correspond to the pattern of stimulation on the receptor mosaic." The retina reorganizes and processes the visual image to a considerable extent. If one is to take into account the well-optimized system of nature, it would be expected that techniques which are solely based on the retinal map without any further processing would not prove to be fruitful. Kolers [58] has stated that "... human perceiving cannot be adequately understood, if understood at all, in terms of passive transmission of impulses from physically defined stimuli; rather perceiving is an active selective process that often has a significant amount of problem-solving about it, and is dependent upon various recoding mechanisms in the perceiver." And later, "Light rays focused on the eye, that is to say, constitute neither a necessary nor a sufficient condition for perception—they constitute only the most studied condition."

A property vector of 1's and 0's is constituted by scanning the character image [90], [69], [126]. This represents a retinal map and no further processing is done. Nadler [81] has strongly criticized this approach.

An interesting discussion of this technique, in the context of pattern recognition generally, is presented by Uhr [120].

M. Grey Scale Discrimination

Although in general most types of optical scanning equipment yield digital representations of pictures in terms of a grey scale with several levels (typically as many as 64), in fact not much use has been made of this additional information. Since a grey scale with more than two levels requires considerably more computer storage, it is common to reduce the data to this binary form. The grey scale is often used for preliminary preprocessing or filtering operations but is rarely intimately involved in the pattern decision process.

Deutsch [13] and Kasvand [52] suggest that grey level discrimination is not often necessary. The major requirement is for a significant contrast between the patterns or visual areas which must be differentiated. An interesting discussion of this problem as it affects human vision is presented by Julesz [48]. Graham [31] and Schreiber [101] examine this problem in the context of more efficient communication of 2-dimensional images. It seems reasonable to assume that for relatively simple geometrical shapes, a two-level grey scale would prove sufficient. Witness the facility with which we can easily distinguish political caricatures. However, for more complicated applications such as in the field of biomedical image processing, the multilevel grey scale will undeniably play a major role in the future.

No mathematical techniques or image operations of any major consequence have appeared in the literature to handle these more realistic images. However, Kirsch *et al.* [56] has suggested some simple routines which mainly involve counting the number of cells in a given picture matrix having a certain grey level. Guzman-Arenas [36] has discussed the use of an intensity histogram to aid in setting thresholds for the various grey levels in a picture. Some interesting experiments on reducing photographs to line sketches has been mentioned by Narasimhan and Fornanago [84] who used thresholding as the basic method for achieving this. Holmes *et al.* [42] suggests that targets may be distinguished from their background because of the general variations in their brightness.

In order to analyze cell images and their substructures using computers, it is necessary to consider the optical density of the patterns [72], [94], [95] as well as their size, shape, and texture. Although not strictly concerned with the computer recognition of images, Lorch [68] has discussed the use of computer generated displays to enhance grey level discrimination in order to present more clearly the tissue structure of a monkey's brain.

As far as the use of the grey scale information is concerned, the work described in this section is of a minor nature and considerably more research is required to make this readily available information more useful.

III. Macroanalysis and Macrooperations

A. Parallel Processing

It is generally felt that the parallel processing of data will result in more efficient methods of pattern recognition but few actual applications are reported in the literature. However, many of the operations in Section II could be programmed to become parallel in nature. Ledley [62] has presented an interesting discussion outlining the various possibilities associated with a large scale parallel processing computer. A computer specifically designed to handle spatial problems like the pattern recognition of planar images is detailed by Unger [122], [124]. Of major interest is a hardware implementation of these concepts, the ILLIAC computer at the University of Illinois, which is described by Wallack [127] and McCormick [71] and seems to be very promising.

Glucksman [29] discusses a simulation study of a parallel processing pattern classifier for alphabetic characters. The features upon which the recognition is based are obtained by a series of basic operations which are parallel in nature and resemble a propagation process. Because of this aspect, a high rate of classification is predicted for a proposed hardware implementation. A similar operation is the "medial axis transformation" outlined in Section III-*C* which could readily be determined efficiently using parallel methods, either optical or digital. Philbrick [93] investigates the computational aspects of this problem.

Finally the work of Hawkins [40] concerning the processing of pictures using image intensifier tubes looks both practical and interesting and probably represents the only existing application of parallel processing. Techniques are described for accomplishing parallel operations of addition, subtraction, multiple sums, spatial filtering, and threshold logic with high resolution in periods of microseconds.

B. Texture

Texture refers to the patterns of information or arrangement of the structure found in a picture. This attribute is relatively simple for a human to describe at a glance but presents enormous difficulties for a computer program. Consequently, its incorporation in feature extraction processing has been limited.

Sebestyen [103] has dealt with the subject of texture in the context of reconnaissance photo interpretation using computer techniques. He suggests four features or attributes which may be gleaned from these types of considerations:

1) The dynamic range of the contrast based on some normalized scale.
2) The average density of objects or parts of objects in a scene.
3) The spacing between shapes or details.
4) The sensitivity of the textural computations to effects of rotation and translation.

Holmes *et al.* [42] also discuss this problem and suggest that patterns may differ from their background only from the point of view of texture or 2-dimensional frequency content. Sometimes these present such low-contrast detail as to tend toward a random distribution [31]. Taking into account that objects in aerial photographs possess a relatively uniform density, Holmes [43] has made use of a

Kolmogorov–Smirnov filter which depends on the statistical consistency within a given object image to achieve identification in the photo interpretation problem.

Although the concept of texture has not yet been applied to biomedical pattern recognition problems it would seem to be a profitable approach. The difficulty, of course, lies in the ability to write fast algorithms for gauging the texture.

C. Shape Descriptors

Intuitively it is preferable to describe an object using gross properties rather than local or neighborhood descriptors. At the first glance at a picture, a human is likely to use general terms in relating the features found in a given scene. However, although the specific global properties of shape have eluded any precise theory, it is undoubtedly one of the most important aspects of pattern recognition. Basically it is desirable to describe the structure of an object independent of orientation, translation, or even perturbation. Interesting discussions on the recognition of shape can be found in Uhr ([121], pt. III) and in the papers of Blum [4], [6]. The latter reference discusses the relationship of these theories to some computer models of this process.

Kirsch *et al.* [56] describes several simple counting routines which could be used to extract general information about the global properties of a picture. In addition, included in a proposed man–machine system of library subroutines, are programs for superimposing, smearing, and magnifying scenes or parts of scenes.

A different approach to this problem has been suggested by Kazmierczak [53] who has devised a method for recognizing characters by representing them by a characteristic potential distribution of a 2-dimensional field of flow. The shape criteria are then extracted from this distribution. Similarly, Stevens [111] has proposed a technique based on contour projection which essentially involves the concept of an expanding array of sensors. The motivation for this idea comes from a study of shape recognition in the octupus. "... the principle used is that of fired-cell coincidence counting with respect to the projection of excitation from given contour-indicating cells against the cells which are also members of contours." The work of Harmon [38] mentioned in Section II-*D* is based on a similar philosophy.

Related to the above methods and embodying their virtues is an interesting shape descriptor referred to as a medial axis transformation (MAT). This concept was developed by Blum [4]–[6] and explained and discussed by Kotelly [59] and Philbrick [93]. Blum [6] describes the generating model which is used to define the MAT: "Consider a continuous isotropic plane (an idealization of an active granular material or a rudimentary neural net) that has the following properties at each point: 1) excitation—each point can have a value of 0 or 1, 2) propagation—each excited point excites an adjacent point with a delay proportional to the distance, and 3) refractory or dead time—once fixed, an excited point is not affected by a second firing for some arbitrary interval of time. A visual stimulus from which the contours or edges have been extracted impinges on such a plane at some fixed time and excites the plane at those points. This excitation spreads uniformly in all directions but in such a way that the waves generated do not flow through each other." The MAT is then defined as the locus of the corners in the wavefront. The (propagating) contours have been likened to the front of a grassfire ignited on the pattern boundary and the MAT is then the locus of points where the fire is extinguished. Several examples of MATs are shown in Fig. 19. Note that, if the MAT turns out to be a straight line, then the shape of the object under consideration is symmetrical. The gross properties of the transformation are obviously related to the macroscopic and structural properties of the pattern. That this is clearly so is demonstrated by Blum [6] using a sketch-like representation of a human and then distorts it as shown in Fig. 20: the basic properties of the MAT remain unchanged.

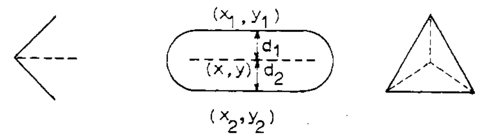

Fig. 19. Typical medial axis transforms (MAT).

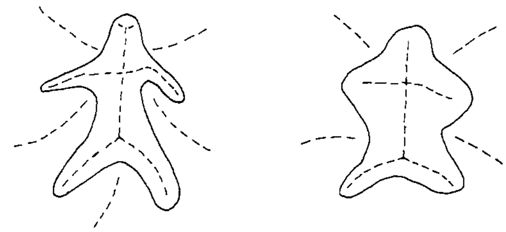

Fig. 20. The MAT for both a simple sketch and a distorted version of a man.

Kotelly [59] describes a mathematical model for this theory and relates it to Huyghen's principle. He also suggests the interesting property that, if one does consider the process as one of wavefront generation, then objects could be classified utilizing a graphical plot of the total length of arc of the wavefront at a given time versus the independent variable time. Examples are given to support this hypothesis but it is doubtful that sufficient discrimination between shapes could be obtained in this way. A simple method for generating this type of representation is presented by Rosenfeld and Pfaltz [99].

Philbrick [93] states that this "... is a method of transforming the input pattern into a description which is directly related to shape and from which both local and gross shape properties may be extracted." In order to define the operation mathematically, consider two points (x_1, y_1) and (x_2, y_2) both of which lie on the boundary of a pattern as shown in Fig. 19. The respective distances from a point (x, y) to (x_1, y_1) and (x_2, y_2) are d_1 and d_2. The transform

$$F(x, y) = d$$

where

$$d = \sqrt{(x - x_1)^2 + (y - y_1)^2} = \sqrt{(x - x_2)^2 + (y - y_2)^2},$$

such that:

1) (x_1, y_1), (x_2, y_2) are on the boundary pattern and are distinct points
2) $d_1 = d_2 = d$
3) d_1 and d_2 are minimized over all allowable values of (x_1, y_1) and (x_2, y_2).

Otherwise $F(x, y) = 0$. A computer method for generating the MATs is given.

A similar approach to describing shape has been discussed by Kasvand [51], [52]. He extracts a skeleton of a given cell under study by searching for "regions of interest." In general, the center of interest is defined as the "extremum over a neighborhood in the processed picture of some functional involving the features of interest in a particular problem." These types of skeletons are also considered in detail in the literature [99], [92], [100]. The reference by Pfaltz *et al.* [92] contains some very interesting algorithms and also introduces the concept of "elongation" which is related to "skeletonization."

An alternate description of shape can be achieved using the method of chain encoding of the contour as was described in Section II-*F*. Freeman and Garder [28] have done this in their work on the computer solution of apictorial jigsaw puzzles. Pfaltz and Rosenfeld [91] make a comparison of this technique with the skeleton representation and demonstrate that the latter requires slightly more computer storage than boundary encoding. However, the skeleton is superior if it is frequently necessary to ascertain whether points lie interior or exterior to a given region as is the case with region shading. Another advantage, when it is uneconomical to analytically describe the set, is the ease with which the intersection of two given sets can be determined.

Both of the above approaches have considerable promise as descriptors and require more research to make them useful for difficult pattern recognition applications.

D. Articular Analysis

Consider the comment by David [12]: "Research over the past few years has shown that linguistic structure is a key not only to recognition of speech but to interpretation of printed characters, cursive handwriting, hand-sent Morse code and indeed to any human communication medium." The importance of contextual analysis and linguistics from the point of view of pattern recognition is discussed by Kirsch *et al.* [56], Dreyfus [17], and Munson [79]. This approach is further stressed by Narasimhan [85]:

> One was fortunate in a certain basic sense, that one's first introduction to picture processing was through bubble chamber pictures and not alphanumeric characters. For it became clear from the very beginning that an appeal to most of the existing pattern recognition models would not only not work, but the models themselves were basically inadequate. These models are based on the categorization of images as belonging to one or another of a finite set of prototypes. Bubble chamber picture scanning defines a context for visual data processing and pattern recognition in which the concepts "prototypes" and "images" become virtually meaningless.

This new approach to the problems of pattern recognition is articular in nature. In a significant paper by Lipkin *et al.* [67], the statement is made that "It is characterized by the reproducible investigation of structural properties." Discussing the often unsatisfactory results achieved by the more conventional approaches he refers to the fact that "it is . . . clear that the failure of a machine procedure to identify the proper object for analysis is attributable to the fact that the machine has been given no information about the *structure* of the images." McCormick [71] and Narasimhan [83], [85] stress the importance of a syntactic model of a visual description as the basis for the interpretation of pictures. The structure of an image is describable by means of a set of hierarchic operations and labels. ". . . the procedure for labelling divides into a series of well-defined steps or levels. At each level, the labelled outputs from the lower levels serve as inputs to the current level of labelling" [85]. The obvious similarity with language is stressed when Narasimhan states further that ". . . given a picture belonging to a well-defined class, the proposal is to generate the type of descriptive statements as indicated earlier on the basis of the assignment of a hierarchic system of labels to the points making up the picture." This type of philosophy leads directly to the requirement for some generative grammar for a given set of images.

Minsky [76], Ledley and Wilson [63] and Narasimhan [85] discuss the various possible hierarchies. At the lowest level is a property detector or feature extractor. All of the methods already discussed may be invoked at this particular microscopic stage where the basic properties of an image are determined. At the next level a name is assigned to certain defined classes of property lists. Note that the latter will usually represent a quantitative measure of the various attributes associated with an object or part of an object. Minsky [76] also refers to "computed names" which are "constructed for classes by processes which depend on the class definitions." At the next higher level, statements are made about ordered relationships between objects, classes, and properties. The final step is the "identification of the pattern through an analysis of the hierarchical syntactical definitions involving parts and their relationships" [62]. Kirsch [57] and Ledley [62] draw analogies between the recognition of images as described and the hierachical recognition process involved in languages:

Images: parts $\leftrightarrow$ structures $\leftrightarrow$ pictures.
Language: characters $\leftrightarrow$ sentences $\leftrightarrow$ paragraphs.

Lipkin *et al.* [67] discuss the problem of determining when a given pattern recognition algorithm is performing adequately as well as how the behavior of the computer can be improved. He suggests a man–machine communication system via a visual display generated by a computer. The latter could be programmed to display synthesized images which would be representative generalizations of the class under investigation. These synthetic images would in essence divulge to the person the machine's state of knowledge. Accordingly, alterations in certain algorithms or class definitions could be communicated to the computer. In fact these interchanges would be useful in providing insight into the basic structure of the problem. The authors suggest the development of iconic microgrammars ("subsets of English grammatical rules, roughly speaking, which analyze and synthesize ("determine") a coherent set of English sentences") to suit different pattern recognition tasks. This interesting paper also relates these ideas to pattern recognition problems associated with biological images. The subjects discussed in this paper are too vast to review in detail. A comprehensive survey of the literature emphasizing the linguistic approach to pattern recognition is given by Feder [23]. Simmons [107] also presents a survey study of the use of English as a medium of analysis and communication with a computer.

Several applications of this type of approach appear in the literature and these vary considerably in their degree of sophistication and development. A comprehensive discussion of these methods is beyond the scope of this paper and readers are referred to the references. Hierarchical sequences of operations involving structural properties are invoked by Grimsdale *et al.* [34] and Spinrad [109] for the purpose of character recognition. Eden [20]–[22] and Mermelstein [74] consider the problem of recognizing handwritten material by first characterizing a record according to a set of primitive symbols after which the latter are incorporated into a contextual analysis to recognize the letters. Presumably their work could be extended to the recognition of words. Pfaltz *et al.* [92] is concerned with a map analysis system (MANS) in which the primitives are represented by the type of skeletons discussed in Section III-*C* but only the two relationships, "contains" and "is contained in" are used. Problems of data structure and processing are also discussed by the authors. Hankley and Tou [37] apply contextual analysis to the interpretation and classification of single fingerprints and this is perhaps one of the more interesting and advanced applications. A topological coding system is used as a basis for generating code sentences describing a given fingerprint and the use of learning techniques in this context is also discussed. The work of Sutherland [114] on SKETCHPAD, a graphical display system, although not directly related to pattern recognition, is an important application of hierarchical data structures and procedures for generating display files. Ledley [64] and Ledley *et al.* [65] use syntactical analysis for the pattern recognition of chromosomes and have developed two languages called BUGSYS and SYNTAXSYS for this purpose. The paper represents an important contribution and should provide the impetus for further research in this area.

Clearly, the techniques discussed here can allow for a large degree of man–machine interaction, primarily using computer graphics. The BUBBLE SCAN system discussed by Narasimhan [85] is an example of such a setup. It is felt that this concept of pattern recognition shows the most promise for solving the difficult applications problems, and advances should be expected in this direction.

IV. Conclusion

Many of the problems and philosophies associated with image processing have been discussed. The main emphasis was on the concept of feature extraction in the context of pattern recognition. Of particular interest are computer algorithms which can be used in actual applications involving real data. The techniques are divided into two categories, microanalysis and macroanalysis, but it is difficult to rigidly include any topic under a particular heading.

Bibliography

[1] A. G. Arkadev and E. M. Braverman, *Computers and Pattern Recognition*. Washington, D. C.: Thompson, 1967.
[2] W. W. Bledsoe and I. Browning, "Pattern recognition and reading by machine," *1959 Proc. Eastern Joint Computer Conf.* (Boston, Mass.), 1959, pp. 225–233.
[3] H. D. Block, N. J. Nilsson, and R. O. Duda, "Determination and detection of features in patterns," in *Computer and Information Sciences*. Washington, D.C.: Spartan, 1964, pp. 75–110.
[4] H. Blum, "An associative machine for dealing with the visual field and some of its biological implications," AF Cambridge Research Laboratories, Bedford, Mass., Electronics Research Directorate, Rept. AFCRL 62–62, February 1962.
[5] ——, "A machine for performing visual recognition by use of antenna-propagation concepts," *Western Electronic Show and Convention*, Los Angeles, August 1962.
[6] ——, "A transformation for extracting new descriptors of shape," in *Models for the Perception of Speech and Visual Form*. Boston, Mass.: M.I.T. Press, 1967, pp. 362–380.
[7] A. E. Brain and J. H. Munson, "Graphical-data-processing research study and experimental investigation," Stanford Research Institute, Menlo Park, Calif., Final Rept. 22, April 1966.
[8] R. Casey and G. Nagy, "Recognition of printed Chinese characters," *IEEE Trans. Electronic Computers*, vol. EC-15, pp. 91–101, February 1966.
[9] M. M. Chodrow, W. A. Bivona, and G. M. Walsh, "A study of hand printed character recognition techniques," Information and Processing Branch, Rome Air Development Center, Research and Technology Division, Griffiss AFB, Rome, N. Y., Final Report, Tech. Rept. RADC-TR-65-444, February 1966.
[10] S. H. Chung, "Electrophysiology of the visual system," Notes of M.I.T. Special Summer Program Course 6.59s on Pattern Recognition, M.I.T., Cambridge, Mass., July 5–15, 1966.
[11] D. O. Clayden, M. B. Clowes, and J. R. Parks, "Letter recognition and the segmentation of running text," *Information and Control*, vol. 9, pp. 246–264, 1966.
[12] E. E. David, A review of G. S. Sebestyen's book *Decision-Making Processes in Pattern Recognition* (N. Y.: Macmillan Co., 1962), *IEEE Trans. Electronic Computers*, vol. EC-12, pp. 334–335, June 1963.
[13] S. Deutsch, "Pattern recognition in the mammalian nervous system," Microwave Research Institute, Polytechnic Institute of Brooklyn, N. Y., Res. Rept. PIBMRI-1306-65, February 1966.
[14] ——, "Conjectures on mammalian neuron networks for visual pattern recognition," *IEEE Trans. Systems Science and Cybernetics*, vol. SSC-2, pp. 81–85, December 1966.
[15] G. P. Dinneen, "Programming pattern recognition," *1955 Proc. Western Joint Computer Conf.*, Los Angeles, vol. 7, pp. 94–100.
[16] W. Doyle, "Recognition of sloppy, hand-printed characters," *1960 Proc. Spring Joint Computer Conf.*, pp. 133–142.
[17] H. L. Dreyfus, "Alchemy and artificial intelligence," The RAND Corporation, Santa Monica, Calif., RAND Rept. P-3244, December 1965.
[18] R. O. Duda and J. H. Munson, "Graphical-data-processing research study and experimental investigation," Stanford Research Institute, Menlo Park, Calif., Tech. Rept. ECOM-01901-24, August 1966.
[19] R. O. Duda, P. E. Hart, and J. H. Munson, "Graphical-data-processing research study and experimental investigation," Stanford Research Institute, Menlo Park, Calif., Tech. Rept. ECOM-

01901-26, March 1967.
[20] M. Eden, "On the formalization of handwriting," in *Structure of Language and Its Mathematical Aspect*. Providence, R. I.: American Mathematical Society, 1961, pp. 83–88.
[21] M. Eden, "Handwriting and pattern recognition," *IRE Trans. Information Theory*, vol. IT-8, pp. 160–162, February 1962.
[22] M. Eden, "Handwriting generation and recognition," in *Recognizing Patterns*, P. A. Kolers and M. Eden: Eds. Cambridge, Mass.: M.I.T. Press, 1968, pp. 137–154.
[23] J. Feder, "A linguistic approach to pattern analysis. A literature survey," Dept. of Elect. Eng., New York University, N. Y., Tech. Rept. 400-133, February 1966.
[24] G. E. Forsen, "Processing visual data with an automaton eye," in *Pictorial Pattern Recognition*, G. C. Cheng, R. S. Ledley, D. K. Pollock, and A. Rosenfeld, Eds., Washington, D. C.: Thompson, 1968, pp. 471–502.
[25] H. Freeman, "On the encoding of arbitrary geometric configurations," *IEEE Trans. Electronic Computers*, vol. EC-10, pp. 260–268, June 1961.
[26] ——, "Techniques for the digital computer analysis of chain-encoded arbitrary plane curves," *1961 Proc. National Electronics Conf.* (Chicago, Ill.), vol. 17, pp. 421–432.
[27] ——, "On the digital computer classification of geometric line patterns," *1962 Proc. National Electronics Conf.* (Chicago, Ill.), vol. 18, pp. 312–324.
[28] H. Freeman and L. Garder, "Apictorial jigsaw puzzles: the computer solution of a problem in pattern recognition," *IEEE Trans. Electronic Computers*, vol. EC-13, pp. 118–127, April 1964.
[29] H. A. Glucksman, "A parapropagation pattern classifier," *IEEE Trans. Electronic Computers*, vol. 14, pp. 434–443, June 1965.
[30] V. M. Glushkov, "Automata theory and its applications," *1965 Proc. Internatl. Federation for Information Processing Congress* (New York), vol. 1, pp. 1–8.
[31] D. N. Graham, "Image transmission by two-dimensional contour coding," *Proc. IEEE*, vol. 55, pp. 336–346, March 1967.
[32] E. C. Greanias, P. F. Meagher, R. J. Norman, and P. Essinger, "The recognition of handwritten numerals by contour analysis," *IBM J. Res. and Develop.*, vol. 7, no. 1, pp. 14–21, January 1963.
[33] P. H. Greene, "An approach to computers that perceive, learn, and reason," *1959 Proc. Western Joint Computer Conf.*, San Francisco, pp. 181–186.
[34] R. L. Grimsdale, F. H. Summer, C. J. Tunis, and T. Kilburn, "A system for the automatic recognition of patterns," *Proc. IEE* (London), vol. 106, pt. B, no. 26, pp. 210–221, March 1959.
[35] G. F. Groner, "Real-time recognition of handprinted text," The RAND Corporation, Santa Monica, Calif., RAND Memorandum RM-5016-ARPA, October 1966.
[36] A. Guzman-Arenas, "Some aspects of pattern recognition by computer," M.Sc. thesis, M.I.T., Cambridge, Mass., February 1967.
[37] W. J. Hankley and J. T. Tou, "Automatic fingerprint interpretation and classification via contextual analysis and topological coding," in *Pictorial Pattern Recognition*, G. C. Cheng, R. S. Ledley, D.K. Pollock, and A. Rosenfeld, Eds. Washington, D. C.: Thompson, 1968, pp. 411–456.
[38] L. D. Harmon, "Line drawing pattern recognizer," *Electronics*, vol. 33, pp. 39–43, September 2, 1960.
[39] J. K. Hawkins, "Photographic techniques for extracting image shapes," *Phot. Sci. and Engrg.*, vol. 8, no. 6, pp. 329–335, November–December 1964.
[40] ——, "Parallel electro-optical picture processing," in *Pictorial Pattern Recognition*, G. C. Cheng, R. S. Ledley, D. K. Pollock, and A. Rosenfeld, Eds. Washington, D. C.: Thompson, 1968, pp. 373–385.
[41] W. H. Highleyman, "An analog method for character recognition," *IRE Trans. Electronic Computers*, vol. EC-10, pp. 501–512, September 1961.
[42] W. S. Holmes, H. R. Leland, and G. E. Richmond, "Design of a photointerpretation automaton," *1962 Fall Joint Computer Conf., AFIPS Proc.*, vol. 22, Philadelphia, Pa., pp. 27–35.
[43] W. S. Holmes, "Automatic photointerpretation and target location," *Proc. IEEE*, vol. 54, pp. 1679–1686, December 1966.
[44] L. P. Horwitz and G. L. Shelton, "Pattern recognition using autocorrelation," *Proc. IRE*, vol. 49, pt. 1, pp. 175–185, January 1961.
[45] M. K. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Information Theory*, vol. IT-8, pp. 179–187, February 1962.
[46] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction, and functional architecture of the cat's visual cortex," *J. Physiol.*, vol. 160, pp. 106–154, 1962.
[47] D. H. Hubel, "The visual cortex of the brain," *Sci. Am.*, vol. 209, pp. 54–62, November 1963.
[48] B. Julesz, "Visual pattern discrimination," *IRE Trans. Information Theory*, vol. IT-8, pp. 84–92, February 1962.
[49] M. Kabrisky, *A Proposed Model for Visual Information Processing in the Human Brain*. Urbana, Illinois: University of Illinois Press, 1966.
[50] L. A. Kamentsky and C. N. Liu, "Computed-automated design of multifont print recognition logic," *IBM J. Res. and Develop.*, vol. 7, no. 1, pp. 2–13, 1963.
[51] T. Kasvand, "Computer simulation of pattern recognition by using cell-assemblies and contour projection," *1965 Proc. IFAC Tokyo Symp. Sys. Engrg. Control Sys. Design*, pp. 203–210.
[52] ——, "Recognition of monochromatic two-dimensional objects," report available from the author at the National Research Council of Canada, Ottawa, Ontario, 1968.
[53] H. Kazmierczak, "The potential field as an aid to character recognition," *1959 Proc. Internatl. Conf. Information Processing*. Paris: UNESCO, June 15–20, 1959, pp. 244–247.
[54] H. Kazmierczak and K. Steinbuch, "Adaptive systems in pattern recognition," *IEEE Trans. Electronic Computers*, vol. EC-12, pp. 822–835, December 1963.
[55] J. L. Kelly, Jr. and O. G. Selfridge, "Sophistication in computers: a disagreement," *IRE Trans. Information Theory*, vol. IT-8, pp. 78–80, February 1962.
[56] R. A. Kirsch, L. Cahn, C. Ray, and G. H. Urban, "Experiments in processing pictorial information with a digital computer," *1957 Proc. Eastern Joint Computer Conf.*, vol. 12, pp. 221–229.
[57] R. A. Kirsch, "Computer interpretation of English text and picture patterns," *IEEE Trans. Electronic Computers*, vol. EC-13, pp. 363–376, August 1964.
[58] P. A. Kolers, "Some psychological aspects of pattern recognition," Notes of M.I.T. Special Summer Program Course 6.59s on Pattern Recognition, M.I.T., Cambridge, Mass., July 5–15, 1966.
[59] J. C. Kotelly, "A mathematical model of Blum's theory of pattern recognition," AF Cambridge Research Laboratories, Bedford, Mass., Rept. AFCRL 63-164, April 1963.
[60] F. N. Krull and J. E. Foote, "A line scanning system controlled from an on-line console," *1964 Fall Joint Computer Conf., AFIPS Proc.*, vol. 26 (San Francisco, Calif.), pp. 397–410.
[61] D. G. Lebedev and D. S. Lebedev, "Quantizing the images by separation and quantization of contours," *Engrg. Cybernetics*, no. 1, pp. 77–81, January–February 1965.
[62] R. S. Ledley, "Thousand-gate-computer simulation of a billion-gate computer," in *Computer and Information Sciences*, J. T. Tou and R. H. Wilcox, Eds. Washington, D. C.: Spartan, 1964, pp. 457–480.
[63] R. S. Ledley and J. B. Wilson, "Concept analysis by syntax processing," *Proc. Am. Documentation Inst.*, vol. 1, October 1964.
[64] R. S. Ledley, *Use of Computers in Biology and Medicine*. New York: McGraw-Hill, 1965, pp. 339–384.
[65] R. S. Ledley, L. S. Rotolo, J. B. Wilson, M. Belson, T. J. Golab, and J. Jacobsen, "Pattern recognition studies in the biomedical sciences," *1966 Proc. Spring Joint Computer Conf.* (Boston, Mass.), pp. 411–430, 1966.
[66] R. S. Ledley, F. H. Ruddle, J. B. Wilson, M. Belson, and J. Albanan, "The case of the touching and overlapping chromosomes," in *Pictorial Pattern Recognition*, G. C. Cheng, R. S. Ledley, D. K. Pollock, and H. Rosenfeld, Eds. Washington, D. C.: Thompson, 1968, pp. 87–97.
[67] L. E. Lipkin, W. C. Watt, and R. A. Kirsch, "The analysis, synthesis, and description of biological images," *Ann. N. Y. Acad. Sci.*, vol. 128, pp. 984–1012, 1965.
[68] S. Lorch, "Computer image processing of biological specimens," *1967 Proc. DECUS Spring Symp. on Display Applications* (Rutgers University, New Brunswick, N. J.), pp. 59–61, April 14–15, 1967.
[69] F. N. Marzocco, "Computer recognition of handwritten first names," *IEEE Trans. Electronic Computers*, vol. EC-14, pp. 210–217, April 1965.
[70] S. J. Mason and J. K. Clemens, "Character recognition in an experimental reading machine for the blind," in *Recognizing Patterns*, P. A. Kolers and M. Eden, Eds. Cambridge, Mass.: M.I.T. Press, 1968, pp. 155–167.
[71] B. H. McCormick, "The Illinois pattern recognition computer, ILLIAC III," *IEEE Trans. Electronic Computers*, vol. EC-12, pp. 791–813, December 1963.

[72] M. L. Mendelsohn, W. A. Kolman, B. Perry, and J. M. S. Prewitt, "Computer analysis of cell images," *Postgrad. Med.*, vol. 38, no. 5, pp. 567–573, November 1965.
[73] P. Mermelstein and M. Eden, "A system for automatic recognition of handwritten words," *1964 Fall Joint Computer Conf., AFIPS Proc.*, vol. 26, pp. 333–342.
[74] ——, "Experiments on computer recognition of connected handwritten words," *Information and Control*, vol. 7, pp. 255–270, 1964.
[75] M. J. Minneman, "Pattern recognition employing topology, cross correlation and decision theory," Republic Aviation Corp., Farmingdale, New York, Rept. RD-TR-64-5096, July 31, 1965.
[76] M. M. Minsky, "Steps toward artificial intelligence," in *Computers and Thought*, E. A. Feigenbaum and J. Feldman, Eds. N. Y.: McGraw-Hill, 1963, pp. 406–450.
[77] I. B. Muchnik, "Local characteristic formation algorithms for visual patterns," *Automation and Remote Control*, vol. 27, no. 10, pp. 1737–1747, October 1966.
[78] J. H. Munson, "The recognition of hand-printed text," IEEE Pattern Recognition Workshop, Puerto Rico, October 1966.
[79] ——, "Some views on pattern-recognition methodology," Internatl. Conf. Methodologies of Pattern Recognition, University of Hawaii, Honolulu, January 24–26, 1968.
[80] M. Nadler, "An analog-digital character recognition system," *IEEE Trans. Electronic Computers*, vol. EC-12, pp. 814–821, December 1963.
[81] M. Nadler, Review of Marzocco's paper "Computer recognition of handwritten first names," *IEEE Trans. Electronic Computers*, vol. EC-14, p. 746, October 1965.
[82] G. Nagy, "State of the art in pattern recognition," *Proc. IEEE*, vol. 56, pp. 836–862, May 1968.
[83] R. Narasimhan, "Labeling schemata and syntactic descriptions of pictures," *Information and Control*, vol. 7, pp. 151–179, 1964.
[84] R. Narasimhan and J. P. Fornanago, "Some further experiments in the parallel processing of pictures," *IEEE Trans. Electronic Computers*, vol. EC-13, pp. 748–750, December 1964.
[85] R. Narasimhan, "Syntax-directed interpretation of classes of pictures," *Commun. ACM*, vol. 9, no. 3, pp. 166–173, March 1966.
[86] U. Neisser, *Cognitive Psychology*. New York: Appleton-Century-Crofts, 1967.
[87] K. L. Nielsen, *Methods in Numerical Analysis*. New York: Macmillan, 1956.
[88] N. J. Nilsson, *Learning Machines*. New York: McGraw-Hill, 1965.
[89] ——, "Adaptive pattern recognition: a survey," *1966 Bionics Symposium*, Dayton, Ohio, May 1966.
[90] A. P. Petrov, "Perception possibilities," *Engrg. Cybernetics*, no. 6, pp. 67–75, November–December 1964.
[91] J. L. Pfaltz and A. Rosenfeld, "Computer representation of planar regions by their skeletons," *Commun. ACM*, vol. 10, no. 2, pp. 119–125, February 1967.
[92] J. L. Pfaltz, J. N. Snively, Jr., and A. Rosenfeld, "Local and global picture processing by computer," in *Pictorial Pattern Recognition*, G. C. Cheng, R. S. Ledley, D. K. Pollock and A. Rosenfeld, Eds. Washington, D. C.: Thompson, 1968, pp. 353–371.
[93] O. Philbrick, "A study of shape recognition using the medial axis transformation," AF Cambridge Research Laboratories, Bedford, Mass., Rept. AFCRL-66-759, November 1966.
[94] J. M. S. Prewitt and M. L. Mendelsohn, "The analysis of cell images," in *Advances of Biomedical Computer Applications, Ann. N. Y. Acad. Sci.*, vol. 128, article 3, pp. 1035–1053, January 31, 1966.
[95] J. M. S. Prewitt, B. H. Mayall, and M. L. Mendelsohn, "Pictorial data processing methods in microscopy," *Proc. Soc. Phot. Instr. Engrs.*, Boston, Mass., June 1966.
[96] M. L. Reisch, "Image processing of the histology of human lungs," M. Eng. thesis, Dept. Elec. Engrg., McGill University, Montreal, Canada, January 1969.
[97] W. M. Rintala and C. C. Hsu, "A feature-detection program for patterns with overlapping cells," *IEEE Trans. Systems Science and Cybernetics*, vol. SSC-4, pp. 16–23, March 1968.
[98] L. G. Roberts, "Machine perception of three-dimensional solids," in *Optical and Electro-Optical Information Processing*, Cambridge, Mass.: M.I.T. Press, 1965, pp. 159–197.
[99] A. Rosenfeld and J. L. Pfaltz, "Sequential operations in digital picture processing," *J. ACM*, vol. 13, no. 4, pp. 471–494, October 1966.
[100] D. Rutovitz, "Data structures for operations on digital images," in *Pictorial Pattern Recognition*, G. S. Cheng, R. S. Ledley, D. K. Pollock, and A. Rosenfeld, Eds. Washington, D. C.: Thompson, 1968, pp. 105–133.
[101] W. F. Schreiber, "Picture coding," *Proc. IEEE*, vol. 55, pp. 320–330, March 1967.
[102] G. S. Sebestyen, *Decision-Making Processes in Pattern Recognition*. New York: Macmillian, 1962.
[103] ——, "Machine aided reconnaissance photointerpretation," *8th SPIE Tech. Symp.*, Los Angeles, Calif., August 1963.
[104] O. G. Selfridge, "Pattern recognition and modern computers," *Proc. Western Joint Computer Conf.*, pp. 91–93, March 1955.
[105] O. G. Selfridge and U. Neisser, "Pattern recognition by machine," in *Computers and Thought*, E. A. Feigenbaum and J. Feldman, Eds. New York: McGraw-Hill, 1963, pp. 237–250.
[106] H. Sherman, "A quasi-topological method for the recognition of line patterns," *Proc. Internatl. Conf. Information Processing*, Paris: UNESCO, pp. 232–238, June 15–20, 1959.
[107] R. F. Simmons, "Answering English questions by computer: a survey," *Commun. ACM*, vol. 8, no. 1, pp. 53–70, January 1965.
[108] F. W. Smith and M. H. Wright, "Automatic ship photo interpretation by the method of moments," Sylvania Electronics Systems, Mountain View, Calif., Rept. SESW-M 1169, November 1967.
[109] R. J. Spinrad, "Machine recognition of hand printing," *Information and Control*, vol. 8, pp. 124–142, 1965.
[110] W. Sprick and K. Ganzhorn, "An analogous method for pattern recognition by following the boundary," *Proc. Internatl. Conf. Information.* Paris: UNESCO, pp. 238–244, June 15–20, 1959.
[111] M. E. Stevens, "Abstract shape recognition by machine," *1961 Proc. Eastern Joint Computer Conf.*, Washington, D. C., pp. 332–351.
[112] R. H. Stock and J. J. Deener, "A real-time input preprocessor for a pattern recognition computer," *1st Ann. IEEE Computer Conf.* (Chicago, Ill., September 6–8, 1967).
[113] I. H. Sublette, "Character recognition by digital feature detection," *RCA Rev.*, vol. 23, pp. 60–79, March 1962.
[114] I. E. Sutherland, "Sketchpad: A man-machine graphical communication system," Lincoln Laboratory, M.I.T., Cambridge, Mass., Tech. Rept. 296, January 1963.
[115] W. Teitelman, "Real time recognition of hand-drawn characters," *1964 Fall Joint Computer Conf., AFIPS Proc.*, vol. 26, pp. 559–575.
[116] S. Tomita and S. Nouguchi, "Recognition of handwritten Katakana characters," *Electronics and Communications in Japan, J. Inst. Elec. Comm. Engrs. of Japan*, vol. 50, no. 4, pp. 174–183, April 1967.
[117] F. M. Tonge, "A view of artificial intelligence," *1966 Proc. 21st ACM National Conf.* (Los Angeles, Calif.), pp. 379–382, August 30–September 1, 1966.
[118] J. S. Tsai and C. L. Sheng, "A preprocessing method to identify pattern classes with a deterministic pattern function," Dept. Elec. Engrg., University of Ottawa, Ontario, Tech. Rept. 67-12, November 1967.
[119] L. Uhr and C. Vossler, "A pattern-recognition program that generates, evaluates, and adjusts its own operators," *1961 Proc. Western Joint Computer Conf., AFIPS Proc.*, vol. 20, pp. 555–569.
[120] L. Uhr, "Pattern recognition," in *Electronic Information Handling*, A. Kent and O. E. Taulbee, Eds. Washington, D. C.: Spartan, 1965, pp. 51–72.
[121] ——, *Pattern Recognition*, L. Uhr, Ed. New York: Wiley, 1966.
[122] S. H. Unger, "A computer oriented toward spatial problems," *Proc. IRE*, vol. 46, pp. 1744–1750, October 1958.
[123] ——, "Pattern detection and recognition," *Proc. IRE*, vol. 47, pp. 1737–1752, October 1959.
[124] ——, "Pattern recognition using two-dimensional, bilateral, iterative, combinational switching circuits," *1962 Symp. Mathematical Theory of Automata*, Brooklyn, N. Y.: Polytechnic Press of the Polytechnic Institute of Brooklyn, Microwave Res. Inst. Symp. Ser., vol. 12, 1962, pp. 577–591.
[125] A. Vander Lugt, "Signal detection by complex spatial filtering," *IEEE Trans. Information Theory*, vol. IT-10, pp. 139–145, April 1964.
[126] V. N. Vapnik, A. Y. Lerner, and A. Y. Chervonenkis, "Learning systems for pattern recognition using generalized portraits," *Engrg. Cybernetics*, no. 1, pp. 63–77, January–February 1965.
[127] P. J. Wallack, "Introduction to Illiac IV," Rept., Dept. of Computer Science, University of Illinois, Urbana, Ill., April 30, 1968.
[128] J. S. Wholey, "The coding of pictorial data," *IRE Trans. Information Theory*, vol. IT-7, pp. 99–104, April 1961.
[129] S. Yamada and J. P. Fornanago, "Experimental results for local filtering of digitized pictures," Dept. of Computer Sci., University of Illinois, Urbana, Ill., Rept. 184, June 15, 1965.