On the syntactic analysis of figures

by THEODOSIOS PAVLIDIS
Princeton University
Princeton, New Jersey

INTRODUCTION

Many investigators have made use of syntactic analysis of figures for pattern recognition purposes during the recent years.[1-5] Also a general theory of pattern analysis has been formulated recently be Grenander.[6] Most of the earlier investigators concentrated their attention on the problem of **line patterns**, namely figures which could be composed by thin lines. Such is the case, for example, of the bubble chamber photographs.[4,5] This paper attempts to develop a similar analysis for the case of set patterns and in particular for polygons. Any two-dimensional figure could be reduced into this form by quantizing its levels of illumination. Then all points with a given level of illumination form a plane set. Such sets could then be approximated by polygons. This could be achieved by a computer-controlled scanner with rectilinear motion or by a number of other techniques. Sometimes the polygonal approximation is only implicit (see the section on Implementation, below). The reason for the polygonal approximation will become evident in the next section. It should be emphasized that this process and the subsequent analysis which is the subject of this paper, are very sensitive to noise, and therefore, a prefiltering of the figure may be necessary. This sensitivity is a common feature of all the techniques which deal with the analysis of figures in simpler components. A more detailed discussion of these points is given elsewhere.[7] In this paper we will try to emphasize only the basic ideas of our approach.

REPRESENTATION OF FIGURES THROUGH SIMPLER ELEMENTS

Let $\phi_1, \phi_2, \ldots \phi_n$ denote a number of directions in the x-y plane (ϕ_1 can be considered as their angle with the x-axis). Let $s_1, s_3, \ldots s_{2i-1}, s_{2n-1}$ be half planes whose boundaries are parallel to the above directions and pass through the origin. The halfplanes are assumed to lie in the side of negative y's. Let s_2, s_{2i}, s_{2n} denote similar halfplanes, but on the side of the positive y's. Note that the union of s_{2i-1} and s_{2i} is the whole plane. Figure 1 illustrates this notation. Let G be the group of parallel translations which in this case are also the similarity transformations of interest and let gs_k denote the halfplane resulting from s_k by translating its boundary by the vector g.* These halfplanes are the primitives of our analysis or signs according to the term used by Grenander.[6]

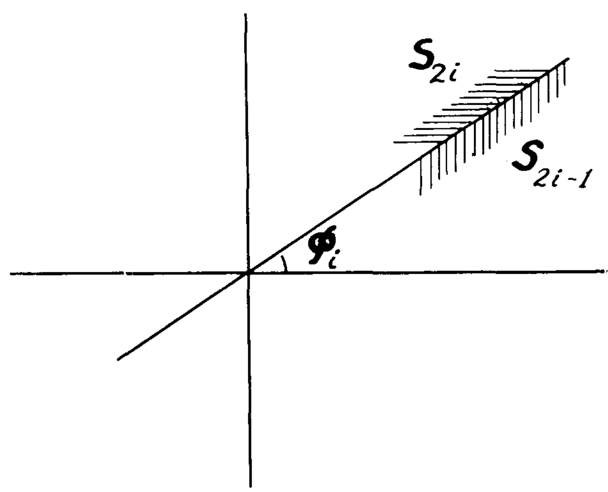

Figure 1—Illustration of the Notation Used for the Signs.

The additional assumption is made that any given figure can be approximated by a polygon with sides parallel to $\phi_1, \phi_2, \ldots \phi_n$. Obviously, the intersection of a number of halfplanes (if not empty) is a convex polygon.[6,8] It is also clear that any nonconvex polygon can be represented as a union of convex polygons

* A halfplane can be characterized by the normal on its boundary from the origin n and a second parameter denoting on which side it lies. A translation by a vector g will result in a new normal $n(1 + \frac{<n,g>}{<n,n>})$ where the expression $<n,g>$ denotes a scalar product. The derivation of this formula is trivial.

(e.g. triangles). What is not as obvious is that any (nonconvex in general) polygon can be represented as a union of convex polygons which are intersections of halfplanes determined by its sides. This will be discussed in the next section, but here we note that under these circumstances any polygon F could be written as

$$F = \bigcup_{k \epsilon K} \left(\bigcap_{a \epsilon A_k} p_a \right) \tag{1}$$

where K and A_k are index sets and p_a are signs of the form gs_m for some $g \epsilon G$ and $1 \leq m \leq 2n$. The convex polygons

$$P_k = \bigcap_{a \epsilon A_k} p_a \tag{2}$$

will be called the primary subsets of F.

One can now try to introduce a finite number of *basic components* $B_1, B_2, \ldots B_N$ which will be convex polygons and which can approximate all other such sets. This can be achieved by introducing a similarity metric in the set of convex polygons. One possible candidate for this purpose is the Hausdorff distance but other techniques could be used as well. One of these is described in the section of the Identification of Basic Components, below. More details can be found in reference[7].

As a result of either procedure, a primary component P will be substituted by a set of the form gB where $g \epsilon G$ and B is a basic component. Then Equation (1) will have the form:

$$F \cong \bigcup_{k \epsilon K} g_k \ B_{i(k)} \tag{3}$$

The similarity transformations appear in Equation (3) because the relative position of the various components is important. However, since F and gF carry the same meaning (for any g) a normalization of Equation (3) will be necessary in practice.

Besides having expressed a figure through a finite number of simpler figures, it is desirable to use only a finite number of transformations. This can be achieved as follows. Let $t(g_1,g_2)$ denote the norm of the vector $g_1 - g_2$. This is a metric defined on G. Let H_ϵ be a subgroup of G which is countable and has the property that for any $g \epsilon G$ exists on $h \epsilon H$ such that

$$t(g,h) < \epsilon \tag{4}$$

where ϵ is given. An example of $H\epsilon$ is the set of vectors with integer coordinates and $\epsilon = 0.5$.

In general, the figures are projected on a finite screen and therefore only a finite number of elements H_ϵ of H_ϵ are expected to be of interest. Then figure F can be represented by a binary matrix M whose columns correspond to the basic components $B_1, B_2, \ldots B_N$ and the rows to the elements of H'_ϵ. An element m_{ij} of M will be set equal to one of $h_i B_j$ is an approximation of a primary component of F. Otherwise, it will be set equal to zero.

Once a figure has been represented by a binary matrix (or a vector if the columns of the matrix are concatenated) then anyone of the standard classification techniques can be used.[9,10] However, the preprocessing usually results in strong clustering (possibly multimodel) of the points belonging to a certain class and it is often very easy to devise discriminant functions. Experimental results using this approach have been described elsewhere.[11]

It should be emphasized that this is not necessarily the best approach. One might use instead Equation (1) directly and substitute for each p_α the primitive $s_{j(\alpha)}$ whose $p_{|\alpha}$ is under a translation by a vector g_α which belongs to H'_ϵ. Then Equation (1) would be written as

$$F \cong \bigcup_{k \epsilon K} \left(\bigcap_{\alpha \epsilon A_k} g_\alpha s_{j(\alpha)} \right) \tag{1'}$$

In this way F is represented by two finite "alphabets" (H'_ϵ and the set of s_i's) which are joined to form "words" (the primary subsets) and a "phrase" (the figure F). This seems to be the most promising avenue for further investigation.

DETERMINATION OF THE SYNTACTIC STRUCTURE OF A GIVEN FIGURE

The first step in determining the structure described by Equation (1) is to obtain the signs present in a figure F. Let $p_1, p_2, \ldots p_n$ be the signs present in a given figure F. We proceed now to give a precise definition of the primary sets. Let Q_0 be the intersection of all the signs, i.e.

$$Q_0 = \bigcap_{i=1}^{n} p_i \ . \tag{5}$$

One can now form an increasing sequence of sets by removing from the intersection one sign at a time. If I is the index set $[i_1, i_2, \ldots i_k]$ then Q_I will denote the intersection of all the signs except $p_{i_1}, p_{i_2}, \ldots p_{i_k}$ Thus for a particular choice of indices we will have a sequence of the form

$$Q_j = \bigcap_{\substack{i=1 \\ i \neq j}}^{n} p_i$$

$$Q_{jk} = \bigcap_{\substack{i \neq j \\ i \neq k}}^{n} p_i$$

etc.

Obviously, if all the signs are removed, one obtains the whole plane which includes the figure F. Therefore, the sequence must have an element which is such that if one more sign is removed from it the next element will not be included in F. This maximal element will always exist although it may be the empty set. Now we can introduce the following definition:

Definition:

If a sequence of subsets of F of the form of Equation (6) has a nonempty maximal element, that element will be called a primary (convex) subset of F.

The following result is of importance in our analysis:

Theorem:

The union of the primary subsets of a polygon F equals F.

The proof of the theorem is rather lengthy and it is presented elsewhere.[7] Here we will show by an example how the primary sets can be obtained.

Consider the polygon ABCDFGHKN of Figure 2. The intersection of all the signs Q_0 is the triangle CPR. It is seen that this is also the intersection of the halfplanes #2,3 and 6 and therefore, only the sets Q_2, Q_3 and Q_6 will be different from Q_0. The set Q_2 is the trapezoid PDEH, the set Q_3, the trapezoid TURS and the set Q_6, the triangle MCL. Figure 3 illustrates by a tree-like diagram the generation of these sets. Each node of the tree is characterized by two numbers, one refers to the halfplane removed from the intersection and the other (in parentheses) is the numbers of halfplanes necessary to form that intersection. This will be removed successively in the next step. For example, Q_2 is the intersection of the halfplanes #3,4,6 and 7. The set Q_{23} is the rectangle TEHS which is not entirely included in the polygon and therefore is not considered any further. The same is true for the set Q_{24}. On the other hand Q_{26} is the trapezoid DEKM which is the intersection of the halfplanes #3,4,7 and 8. If either one of them is removed, the resulting set is not included in the original polygon. Hence Q_{26} is a primary subset. In the same way one can verify that Q_{27} (the trapezoid DFGP) and Q_{364} (the trapezoid ABLN) are also primary subsets. It can be readily checked that their union equals the original polygon.

This analytical procedure is well-defined and therefore the resulting primary subsets are unique.

IDENTIFICATION OF BASIC COMPONENTS

The identification of a given convex polygon with a basic component mentioned in the section on Representation of Figures Through Simpler Elements is an important step. Such an identification should conform with our intuitive notions of shape similarity.

One possible procedure is the following: Let the screen be scanned from left to right and from bottom to top. The first vertex of a polygon met in this way is used as a starting point and the boundary is scanned clockwise. During this procedure, the signs met are recorded together with their corresponding lengths so that the polygon is described by an ordered array of the form

$$\begin{matrix} s_{i_1} & s_{i_2} & \cdots & s_{i_k} \\ \ell_1 & \ell_2 & \cdots & \ell_k \end{matrix} \tag{7}$$

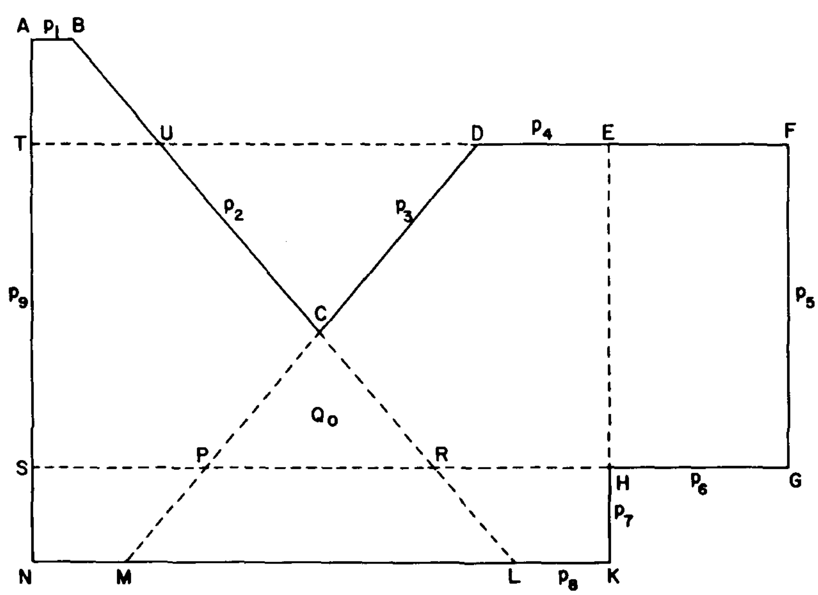

Figure 2—A Non convex Polygon with the Signs $p^1, p^2, \ldots\ldots p$ Present in Precise Notation $p_i = gs_k(i)$.

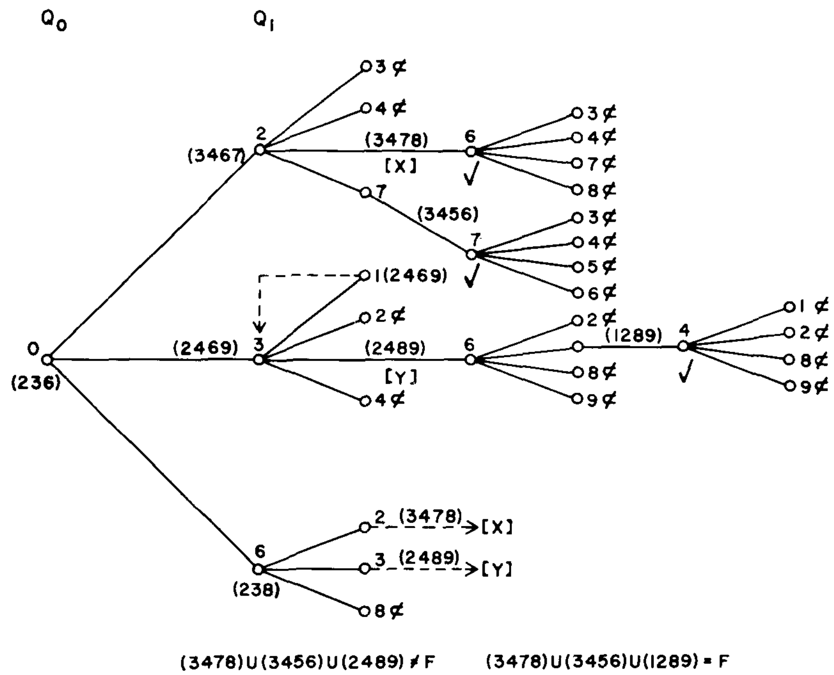

Figure 3—A Tree Indicating the Determination of the Primary Subsets (✓) of the Polygon of Figure 2. These Turn out to be the Convex Polygons ABLNA, MDEKM and PDFGP. The Symbol ⊄ Indicates a Set not Included' in the Original Figure. The Sets (3467), (2469), (238) and (2489) Correspond to DEHPD, RSTUR, MCLM and LNTUL.

Note that because of the convexity no sign will be encountered twice. Figure 4 shows an example. Note also that this representation is translation invariant. The double array of Equation (7) can be compared with similar representations of the basic components.

The latter are searched until one finds those which have the same permutation of signs $i_1, i_2, \ldots i_k$ as the sample under consideration, or, possibly, have one additional element or one fewer element. Then the sum of the squares of the differences between the corresponding lengths can be considered as a measure of how dissimilar two figures are. In the example of Figure 3 the "distance" of P_1 and P_2 is given by

$$d(P_1,P_2)=(\ell_1-\ell_1')^2+\ell_2^2+(\ell_3-\ell_2')^2+\ell_3'^2+(\ell_u-\ell_u')^2+(\ell_s-\ell_s') \quad (8)$$

The vector $\overline{AA}'$ connecting the two lowermost left-most points will give the relative translation g in Equation (3).

IMPLEMENTATION OF THE ANALYSIS

At present there is no complete implementation of the analysis described in the section on the Determination of the Syntactic Structure of a Given Figure. The process is an algorithm in the sense that it can be described through a finite set of instructions to a draftsman and one could even attempt to write a FORTRAN program which would compare sets point by point. This is possible since the two-dimensional information to be processed by a computer is always in discrete form. Another approach is to describe each halfplane by a linear inequality and then use linear programming to solve the system of various sets of them and thus obtain the sets Q_I.

Both techniques are time-consuming and therefore impractical. This limitation of course need not apply for the case of parallel processing but the only practical methods implementing the analysis on a general-purpose sequential machine are approximate. A number of them were developed by using the notion of integral projections.

An integral projection of a figure along a given direction has been defined as the length of the intersection of a figure with a line parallel to this direction taken as a function of the position of the line. It is obvious that a discontinuity in that function reveals in general the existence of a boundary of the figure parallel to that direction. This fact was used for the decomposition of figures and the method is described elsewhere together with experimental results on multi-font typewritten alphabetic characters resulting from

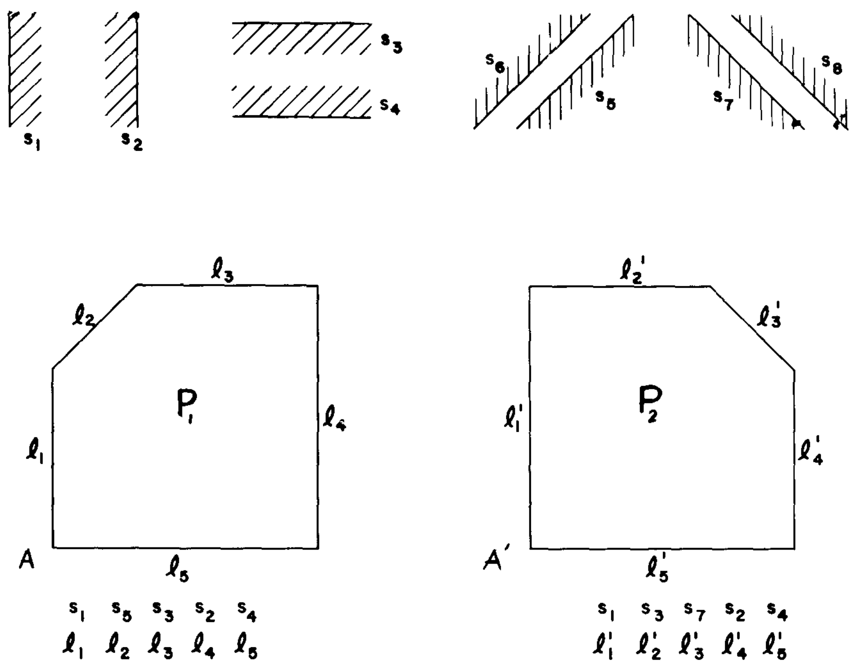

Figure 4—Representation of Convex Polygons through Signs.

its application.[11] For a group of 10 letters an error rate less than 2% was achieved on the testing set. Another method currently under investigation uses integral projections also, but in a more complex way. Only two directions are used, vertical and horizontal. In this way, the primary subsets will be rectangles. Figure 5 shows examples of original characters and their representations through rectangles. In this case the integral projections may be computed directly for the original figure. The reconstructed figure, however, is a polygon.

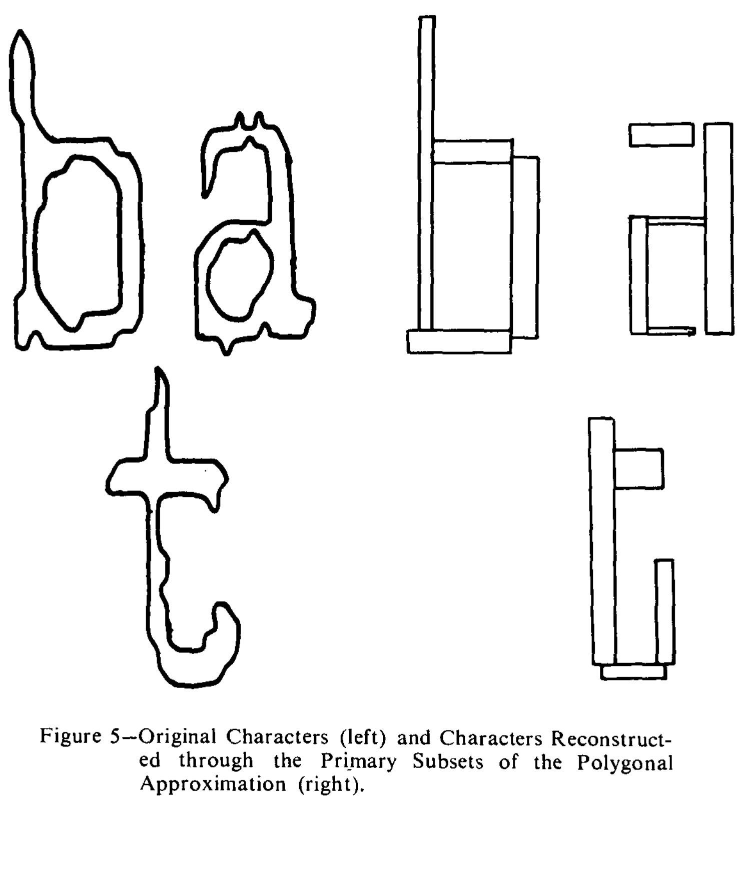

Figure 5—Original Characters (left) and Characters Reconstructed through the Primary Subsets of the Polygonal Approximation (right).

ACKNOWLEDGMENT

The program which produced the decomposition shown in Figure 5 was written by Mr. R. Gross. The original characters were stored on a tape supplied by the IBM Research Center at Yorktown Heights.

REFERENCES

1 M EDEN
Handwriting and pattern recognition
IRE Trans on Information Theory vol IT-8 1962 pp 160-166

2 H BLUM
A transformation for extracting new descriptors of shape
Symposium for the Perception of Speech and Visual Form Boston 1964

3 R S LEDLEY
High-speed automatic analysis of biochemical pictures
Science October 1964 pp 216-223

4 R NARASIMHAN
Labeling schemata and syntactic descriptions of pictures
Information and Control vol 7 1964 pp 151-179

5 A C SHAW
Ph D Thesis
Stanford University 1968

6 U GRENANDER
Foundations of pattern analysis
Tech Report No 2 Brown University 1967

7 T PAVLIDIS
Analysis of set patterns
Technical Report 25 Department of Electrical Engineering Princeton University

8 F A VALENTINE
Convex sets
McGraw-Hill 1964

9 N J NILSSON
Learning machines
McGraw-Hill New York 1955

10 C K CHOW C N LIU
Approximating discrete probability distributions with dependence tree
IEEE Trans Inf Theory May 1968 pp 462-467

11 T PAVLIDIS
Computer recognition of figures through decomposition
Information and Control (in press) A summary can be found in the Proceedings of the 5th Allerton Conference Oct 1967 pp 810-811