Check for updates

Digital Pattern Recognition by Moments

FRANZ L. ALT

National Bureau of Standards, Washington, D. C.

1. *Introduction*

The problem which concerns us may be stated as follows: A digital computer receives information about a plane black-and-white pattern, and is to decide on the basis of this information whether the pattern is "similar", in some sense to be specified below, to a given prototype. It is usually assumed that the patterns considered are of bounded size, say each contained in a given rectangle. Thus a pattern may be defined as any subset of the points in the rectangle, namely, the subset of "black" points. The information given to the computer must be of finite length; usually the rectangle is covered by a finite number of cells, and the information to the computer amounts to signals indicating whether any given cell is white or black. Patterns used as prototypes may be, for example, a ring, the letter "A", or the like. The recognition of alphabetic and numerical characters is of particular practical importance. What constitutes "similarity" varies widely from case to case, and the methods which the machine uses in recognizing similarity must vary accordingly. To give two extreme examples, one might call two patterns similar only if they agree point for point (or rather, cell for cell); or else, one might admit as similar two patterns which are topologically equivalent (after defining a suitable topology in the space of cells), so that e.g. any simple closed curve would be called similar to a circle. Many intermediate definitions are possible.

In particular, for the recognition of printed characters one will wish to admit as similar two patterns if they differ at most in the following respects:

(A) Location

(B) Size

(C) "Stretching" and "Squeezing" in either X- or Y- direction. Mathematically, these are affine transformations preserving the X- and Y- directions, i.e., transformations of the form

$$X^* = aX + b, \qquad Y^* = cY + d.$$

The special case $a = c = 1$ corresponds to translations, i.e. to (A).

(D) While one might consider rotation through a small angle as admissible, rotations through large angles are obviously not; e.g. the characters "6" and "9" should be recognized as different. It seems preferable to omit rotation entirely and admit in its place "slanting" such as occurs in italic type by comparison with roman. Mathematically, this is characterized as an affine transformation which leaves the points of the X-axis fixed. Combinations of the transformations (A) and (D) ex-

FIG. 1. Examples of "similar" patterns

haust all possible affine transformations preserving the X- direction, i.e. all transformations of the form

$$X^* = aX + eY + b, \qquad Y^* = cY + d.$$

By virtue of these, the patterns of Figure 1 are all similar. (A) and (B) are special cases of (C).[1]

(E) A reasonable amount of "random noise".

(F) Certain other differences which appear as minor to a human reader, such as small changes of proportions, slight bending of lines, etc.

Conditions (E) and (F) are purposely left vague at this point.

We have thus described a particular concept of similarity of patterns by giving a list of transformations—namely, transformations which do not affect the similarity between patterns. One could possibly wish to specify a larger class of such transformations, corresponding to a broader concept of similarity; the ones listed are probably minimal.

How does a computer decide whether two patterns are similar under (A) to (F)? One cannot possibly try all admissible transformations to see whether one of them transforms one of two patterns into the other. What one should like to have is a *pattern function*—a number $m = F(P)$ associated with each pattern P—which is

(a) invariant under (A) to (F)—i.e. if patterns P_1 and P_2 are similar according to (A) to (F) then $F(P_1) = F(P_2)$;

(b) *characteristic* of dissimilar patterns—i.e. if patterns P_1 and P_2 are not similar according to (A) to (F) then $F(P_1) \neq F(P_2)$;

(c) easy to compute.

If such a function cannot be found, then one might try to use several functions $F_1(P)$, $F_2(P)$, $\cdots$ which satisfy (a) and (c), though not (b), in the hope that for any given pair P_1, P_2 of dissimilar patterns, at least one of the functions $F_i(P)$ would give $F_i(P_1) \neq F_i(P_2)$. The latter is the approach we propose to follow.

One is led to such a sequence of functions by the remark that a pattern may be considered as a function of two variables, rather than as a point set; namely, the function $F(X, Y)$ which is equal to 1 if the point (X, Y) is "black", and

[1] Conditions (A) and (B) have been proposed repeatedly, often in conjunction with invariance under rotation. Condition (D) does not seem to have been proposed before.

equal to 0 if it is white.[2] In general, functions of several variables can often be characterized by sequences of numbers in various ways, e.g. Taylor coefficients, Fourier coefficients, etc. A characterization which looks promising for our purpose is that by moments.[3]

2. *The Moments*

For a function of two variables $f(X, Y)$, the moments are defined as the numbers

$$M_{jk} = \iint X^j Y^k f(X, Y)\, dX\, dY, \qquad j, k = 0, 1, 2, \cdots \quad (1)$$

where the integration extends over the domain of definition of f; in our case, over the rectangle in the (X, Y) plane containing our pattern. In the special case where f is constant, equal to 1 or 0, in each cell, this may be replaced approximately by

$$M_{jk} = c\, \Sigma\, X^j Y^k \quad (2)$$

where the summation is taken over all "black" cells, c is the area of one cell, and (X, Y) are the coordinates of some point in the cell. Two patterns which are identical have, of course, identical moments. Two patterns which are not identical will differ in one or more of their moments. (The latter statement is not true for the most general kinds of patterns, but it is true for all sufficiently simple patterns,[4] including the cell patterns to which we have limited ourselves.) Furthermore, it is easy to define certain combinations of moments which are invariant under the transformations (A) to (D); it is this latter property which makes moments so attractive to us. As to the condition (E), the moments will be only slightly changed if the noise is sufficiently slight; and similarly for (F). Thus we can meet conditions (A) to (D) completely, and (E) and (F) at least approximately, by working with appropriate combinations of moments.

According to (2), the zero-order moment is equal to the black area of the pattern,

$$M_{00} = cM \quad (3)$$

where M is the number of black cells. The center of gravity of the pattern has the coordinates

$$\bar{X} = M_{10}/M_{00}, \qquad \bar{Y} = M_{01}/M_{00}. \quad (4)$$

[2] This can obviously be generalized to patterns which contain shades of gray rather than just black and white. In principle our method is applicable to this case without change. In practice it would probably fail.

[3] After submission of the original manuscript of this paper, the author's attention was drawn to similar work in progress at the A. D. Little Co., Cambridge, Mass., under the direction of V. E. Giuliano and others. The use of moments has also been advocated by Hu [7], apparently without having been tried out on specific cases.

[4] Sufficient conditions for patterns to be thus characterized by their moments may be found, e.g., in [4].

If we refer the moments to the center of gravity as coordinate origin, we get the expressions

$$\bar{M}_{jk} = c \, \Sigma \, (X - \bar{X})^j (Y - \bar{Y})^k$$

which are invariant under displacements. These may be expressed as linear combinations of M_{jk} and the moments of lower order.

$$\bar{M}_{00} = M_{00}, \qquad \bar{M}_{10} = \bar{M}_{01} = 0,$$

$$\bar{M}_{20} = c \, \Sigma \, (X - \bar{X})^2 = M_{20} - M_{10}^2/M_{00}, \qquad \bar{M}_{02} = M_{02} - M_{01}^2/M_{00}.$$

The variances

$$\sigma_X = \sqrt{\bar{M}_{20}/M_{00}}, \qquad \sigma_Y = \sqrt{\bar{M}_{02}/M_{00}} \tag{5}$$

can be used to normalize the coordinates by setting

$$x^* = (X - \bar{X})/\sigma_X, \qquad y^* = (Y - \bar{Y})/\sigma_Y. \tag{6}$$

It is also convenient to normalize the moments by dividing by M_{00}. Then the expressions

$$m_{jk}^* = c \, \Sigma \, x^{*j} y^{*k}/M_{00} \tag{7}$$

are invariant under the transformations (A), (B) and (C). Finally we set

$$\rho = \Sigma \, x^* y^*/\Sigma y^{*2},$$

$$x = \frac{x^* - \rho y^*}{\sqrt{1 - \rho^2}}, \qquad y = y^*, \tag{8}$$

and refer the moments to these coordinates. Then

$$m_{jk} = c \, \Sigma \, x^j y^k/M_{00} \tag{9}$$

are invariant under transformations (A) to (D). In particular

$$m_{00} = m_{20} = m_{02} = 1, \qquad m_{10} = m_{01} = m_{11} = 0. \tag{10}$$

The quantity ρ used in (8) is what statisticians call the regression coefficient of x^* on y^*.

If a pattern is affected by a small amount of noise, this means that a few cells are black instead of white or vice versa. If, in the summation in (9), the number of terms is large, then the few terms added or omitted because of noise will make relatively little difference. Thus (9) is "almost invariant" under (E). Similarly, transformations of type (F) involving small displacements, i.e. small changes in the values of x and y in some terms of (9), will have relatively little influence. Indeed, these changes can be considered as a special kind of noise. In general, the effect of noise will depend on its distribution; this can be characterized in various ways, among others in terms of moments. How the moments of a given pattern are affected by noise of a given kind is not a simple matter; the only thing we can say about it at this time is that if the intensity of noise is small enough, then its effect on the pattern moments is small.

It is our purpose to investigate the practical limitations on the use of moments for pattern recognition, by examining such questions as:

How many different moments must be computed to distinguish significantly different patterns, such as different alphabetic characters?

How many cells, and of what size, should be used?

How much noise can be tolerated?

What is the effect of rounding errors in the computations?

How well can characters from different type fonts be recognized as similar?

3. *Experiments*[5]

To throw some light on these questions, a number of diagrams were analyzed numerically, by computing the first few moments of each. Initially each diagram was placed within a square, about 2 inches on a side. Each square was examined by an automatic scanner which took and recorded readings on a 176 × 176 grid. Then the moments from the third to the sixth order were computed from both equations (7) and (9).

The scanner used in these experiments was built at the National Bureau of Standards and is described in [1]. It is connected to SEAC [2], which records the readings directly onto magnetic tape of the kind used with most large IBM computers. The magnetic tape is read into a 704 computer, where the computation of the moments is carried out. For each diagram the computer prints out results in the following arrangement:

$$
\begin{array}{lllllll}
M & \bar{X} & \bar{Y} & \sigma_X & \sigma_Y & \rho & \\
m_{30} & m_{21} & m_{12} & m_{03} & & & \\
m_{40} & m_{31} & m_{22} & m_{13} & m_{04} & & \\
m_{50} & m_{41} & m_{32} & m_{23} & m_{14} & m_{05} & \\
m_{60} & m_{51} & m_{42} & m_{33} & m_{24} & m_{15} & m_{06}
\end{array}
$$

The magnitudes m_{jk}^* according to (7) can be printed in a similar arrangement. It is sufficient to print moments beginning with the third order (m_{30}, m_{21}, etc.) since the moments up to the second order are standardized by (10). Optionally the computer also prints a picture of the black and white spots, and their coordinates, standardized by (6) or (8).

The following diagrams were examined:

(a) Five geometrical figures:

(a.1) A square consisting of 1681 "black" points arranged 41 × 41.

(a.2) A square of 121 points in 11 × 11 arrangement.

(a.3) A diamond, 41 points on each diagonal (diagonals parallel to the x and y axes).

(a.4) A hollow square, consisting of 160 points arranged along the circumference of a square, 41 points to a side.

(a.5) A cross of 81 points, arranged in two diagonal lines of 41 points each.

[5] The computer programs used in these experiments were planned by Sally T. Peavy and written by Robert Herboldt.

FIG. 2. A diagram ready for the scanner

(b) Twenty-six capital letters and nine numerals of a particular type font (Bodoni 375). One of these is shown in Figure 3. They are treated by the scanner in groups of 4, 8 or 9, arranged 2 by 2, 4 by 2 or 3 by 3, with enough white space between characters that in the final picture the length of each character corresponds to about 40 of the 176 grid points in the Y-direction, the width to about 30 of 176 grid points in the X-direction. One such group is shown in Figure 2. After the magnetic tape is fed into the 704, the bits corresponding to the characters are separated, and the analysis proceeds for each character separately.

The moments of each of the figures in (a) could easily have been expressed by formulas and evaluated. Since, however, a machine computing code was available for use with (b), it was easier to apply the same code to (a). However, the input to the computer for (a), rather than being obtained from the scanner, was manually punched into cards.

For each of the diagrams of (b), moments of orders 3 to 6 were computed from both equations (7) and (9).

4. *Results*

Table 1 shows the results for the five geometrical figures which were examined. Because of the symmetry of these particular figures, any moment containing an odd power of either x or y is zero. The table therefore shows only:

$$\begin{array}{llll} m_{40} & m_{22} & m_{04} & \\ m_{60} & m_{42} & m_{24} & m_{06} \end{array}$$

The moments from (7) and (9) are identical, $\rho = 0$ in all these cases. Each of the five figures possesses additional symmetries, evident from the table. Alongside these computed moments are shown the moments for the corresponding continuous case, computed by formula. Our purpose here is to show how much distortion of numerical results is caused by the use of cells of finite (and fairly large) size. The comparison of cases 1 and 2 is interesting because it indicates the effect of different grid sizes (case 2 being equivalent to the use of a coarser grid). We also get some feeling for the geometric interpretation of moments. For example, the mixed moments, m_{22}, m_{42}, m_{24}, etc. are about half the size of the pure moments m_{j0}, m_{0j}; they are relatively larger for figures clustered near the diagonals, as in 5, relatively smaller for figures clustered near the axes, as in 3. We remark that for a figure consisting of only four points, forming the

TABLE 1. Moments of Certain Geometrical Figures

Discrete Pattern				Continuous Pattern			
			1. Square ■				
	$M = 1681$				$M = \infty$		
1.80	1.00	1.80		1.800	1.000	1.800	
3.85	1.80	1.80	3.85	3.857	1.800	1.800	3.857
			2. Square ■				
	$M = 121$				$M = \infty$		
1.78	1.00	1.78		1.800	1.000	1.800	
3.73	1.78	1.78	3.73	3.857	1.800	1.800	3.857
			3. Diamond ◆				
	$M = 841$				$M = \infty$		
2.40	0.40	2.40		2.400	0.400	2.400	
7.74	0.51	0.51	7.74	7.714	0.514	0.514	7.714
			4. Hollow Square □				
	$M = 160$				$M = \infty$		
1.35	0.75	1.35		1.350	0.750	1.350	
1.93	0.90	0.90	1.93	1.929	0.900	0.900	1.929
			5. Diagonal Cross ×				
	$M = 81$				$M = \infty$		
1.78	1.78	1.78		1.800	1.800	1.800	
3.75	3.75	3.75	3.75	3.857	3.857	3.857	3.857

Note. The continuous pattern moments are rational numbers. In particular, 3.857, wherever it occurs in this table, stands for 27/7; 7.714 for 54/7; 1.929 for 27/14; 0.514 for 18/35. The other continuous pattern moments are exact as listed above. The discrete pattern moments are rounded to 2 places; most of them, when written as fractions, have unwieldy denominators.

corners of a square with sides parallel to the axes, all even moments are 1; all odd moments, i.e. m_{jk} with either j or k odd, are zero as before. For four points at the corners of a diamond, the even pure moments $m_{2j,0}$, $m_{0,2j}$ are equal to 2^{j-1}, while all mixed moments, as well as all odd pure moments, are zero.

Figure 3 is a pictorial representation of the information stored in the computer while moments are being computed. The picture is slightly distorted because of equipment limitations: it is plotted on an ordinary tabulator, which has a choice of only a discrete number of printable positions (6 per inch in the Y direction, 10 per inch in the X direction). The coordinates of the points which are stored in memory are rounded to the nearest printable value, and this produces some of the irregularities of the printed picture. The computation of the moments, however, is carried out with the original unrounded coordinates.

The figure also shows the effect of the transformation (8). This transformation is intended to remove any difference in "slant" between different specimens of patterns of the same kind; for instance, the difference between italic and roman type. It does this by reducing every pattern to a standard form which has zero slant. Some letters, however, such as Z, N, L, J, have a built-in slant in their roman form. In such a case, the machine routine will produce an artificial slantless standard form, e.g., by tilting N and L somewhat to the right, Z and J to the left. In the case of the letter A, such a built-in slant is created in some type fonts, including the one used here, by thickening the right leg of the A. The lower half of Figure 3 shows how the computer routine removes this slant. An idea of the importance of this factor can be gained from Table 2, where the values of the regression coefficient ρ used in (8) is given in the heading for each character. Values of ρ below about 0.1 may be considered insignificant. The largest value occurring is $\rho = .459$, for the letter X. As in the case of A, the apparent strong slant of X is caused only by the thickening of one of the two lines forming the X in the particular font used here.

Table 2 shows what may be considered as the main result of the present study, the moments of orders 3 to 6 for the capital letters of the alphabet and numerals 1 to 9. The interpretation of these results will be discussed in the next section. A few salient points, however, should be mentioned here.

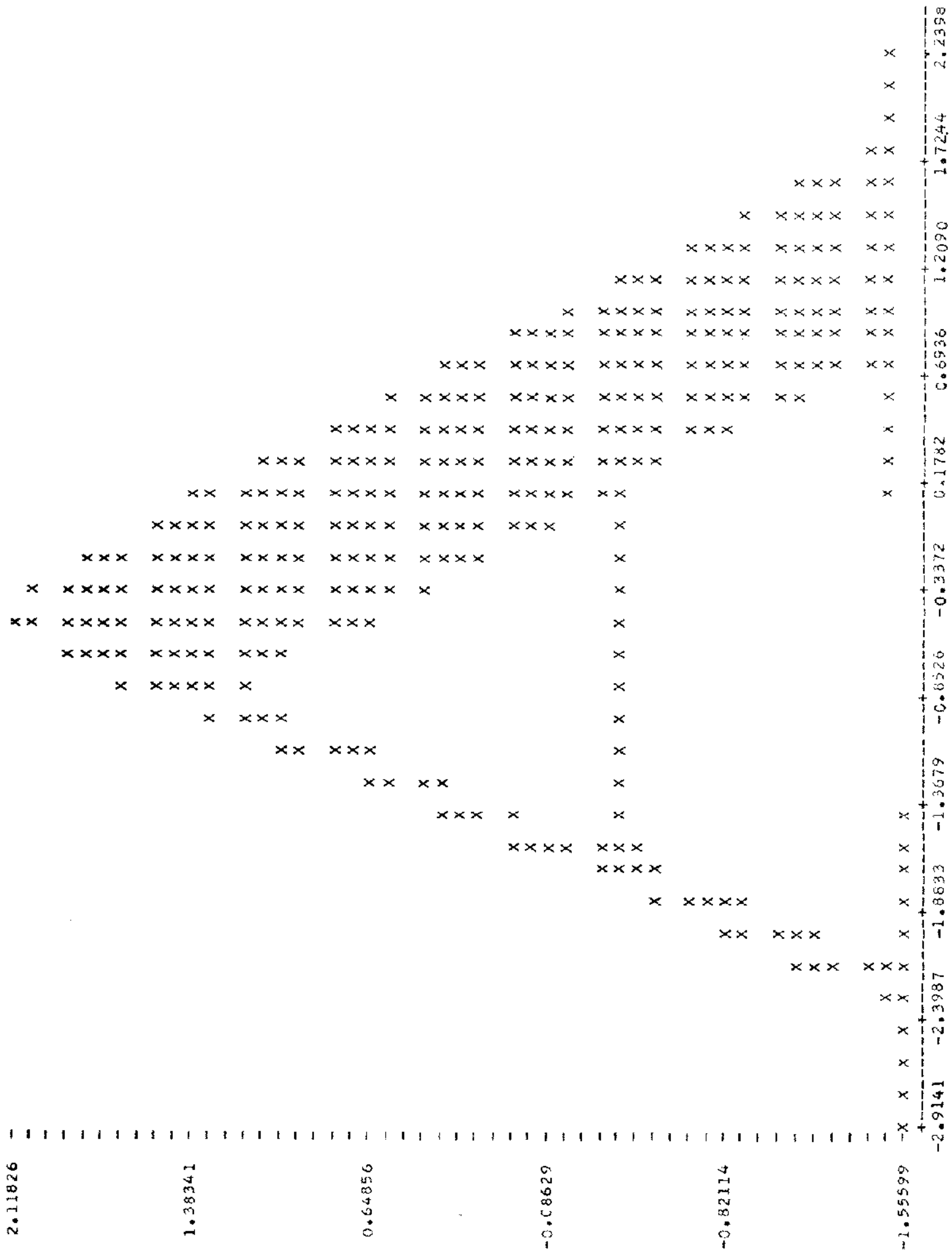
Moments involving an odd power of y, i.e. those in the second, fourth and sixth columns of the table, should theoretically be zero for any pattern which is symmetrical with respect to a horizontal axis. Indeed, an inspection of the table shows that these moments are relatively small—though not exactly zero, of course, because of noise, imperfect alignment, discretization error, etc.—for the characters B, C, D, E, H, O, and to a lesser extent for I and 8. Accidentally these moments are also relatively small for R, which does not possess this symmetry. For all other characters at least some of these moments are quite large. Usually we find at least one such moment of absolute value exceeding 1, while for the characters B, C, D, E, H, O most of these moments are below .1, and all below .26 in absolute value. (If we omit moments of order 6, which are more strongly influenced by noise, all absolute values are $\leqq .1$ except $m_{05} = .19$ for H, $m_{41} = -.13$ for E.) For the letter X, which in some type fonts has horizontal symmetry, this symmetry is destroyed by the different thickness of the two lines forming the X.

One might expect that the characters I and 1 should be almost, if not precisely, symmetric with respect to the X-axis. However, because they are so narrow, the transformation (6) has the result of magnifying the X-coordinates of all their points (because σ_X is so small) and thereby also magnifying any asymmetry present in them. This also explains the large value of ρ for the character 1 ($\rho = .234$).

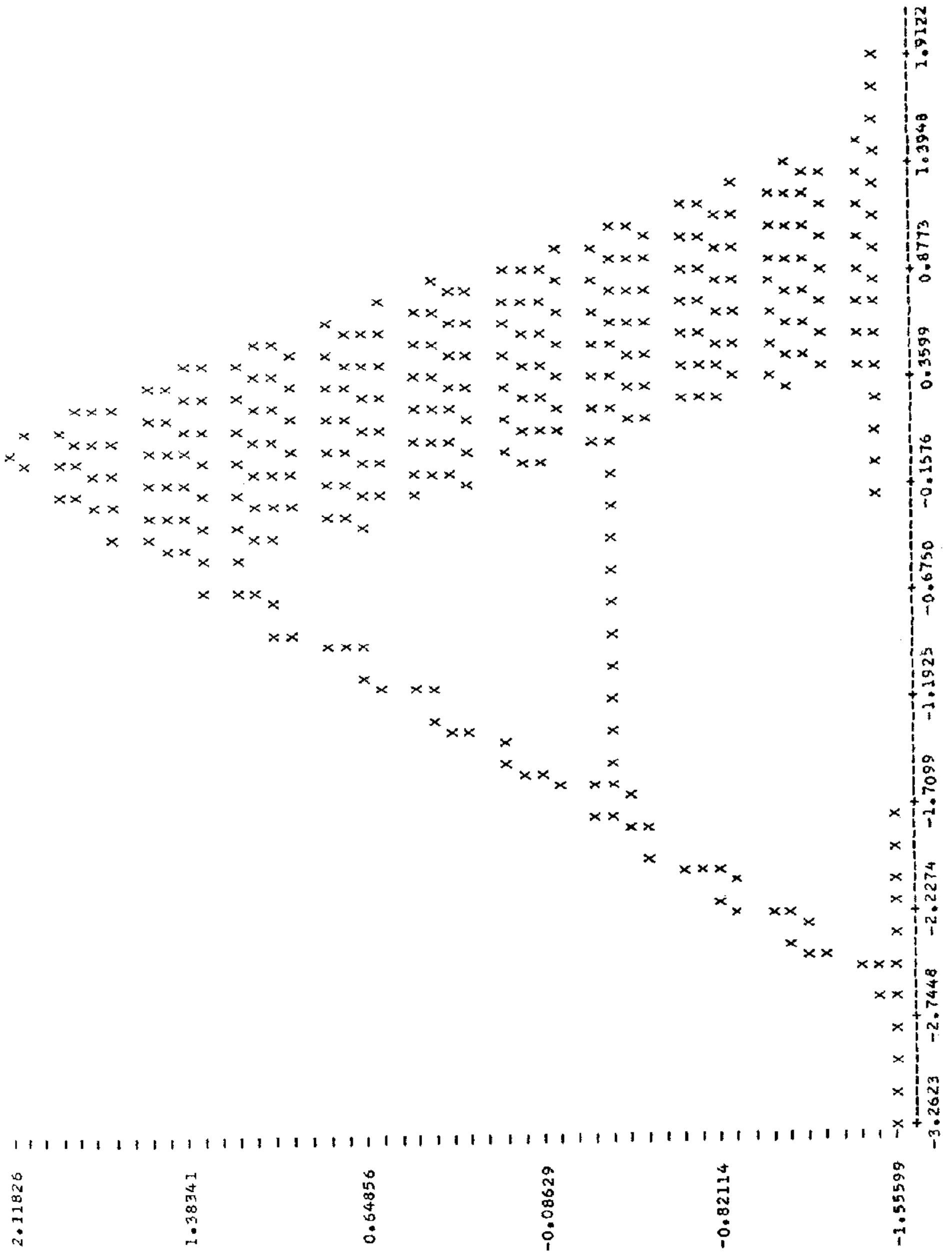
The moments involving odd powers of x should be zero for patterns symmetric with respect to a vertical axis. This is indeed seen to be the case for H, O, 8 and to a lesser extent for Q, R, T (despite the absence of symmetry in R).

In every case the observed moments are consistent with the symmetry properties of the characters used, if we assume that noise and minor deviations of characters from their ideal form may change the magnitudes of moments by about .20; a little less perhaps for the moments of lower order, a little more for those of higher order. This seems reasonable, and pending a more detailed study

FIG. 3. A character as stored in computer memory (a) before, and (b) after transformation (8).



(a)


2.11826
1.38341
0.64856
-0.08629
-0.82114
-1.55599
-3.2623
-2.7448
-2.2274
-1.7099
-1.1925
-0.6750
-0.1576
0.3599
0.8773
1.3948
1.9122

(b)

of the sensitivity of moments we shall assume that it is true not only for those moments which, because of symmetries, ought to be zero, but for all others as well.

It appears, then, that differences of about .20 between moments should be disregarded. If this is so, does our set of moments have enough resolving power to distinguish between different characters? For a preliminary answer to this question, look at two highly similar characters, such as O and Q. We find that a number of moments, including all of order 3 and 4, are quite closely similar.

TABLE 2. Moments of Printed Characters

A $\quad M = 330,\ \bar{X} = 23.22,\ \bar{Y} = 18.84$; $\sigma_x = 6.60,\ \sigma_y = 10.89,\ \rho = -0.224$

-1.18	-0.87	-0.10	0.17			
4.03	1.36	1.26	0.10	1.92		
-8.75	-4.89	-1.96	-1.59	-0.33	0.85	
25.36	11.50	6.85	2.81	2.42	0.38	4.66

B $\quad M = 530,\ \bar{X} = 19.12,\ \bar{Y} = 24.01$; $\sigma_x = 9.27,\ \sigma_y = 11.83,\ \rho = -0.077$

0.19	-0.03	-0.11	-0.01			
1.35	-0.07	0.96	0.08	1.77		
0.52	-0.09	-0.05	-0.01	-0.41	-0.08	
2.19	-0.23	1.30	0.04	1.59	0.26	3.70

C $\quad M = 331,\ \bar{X} = 15.51,\ \bar{Y} = 25.29$; $\sigma_x = 10.09,\ \sigma_y = 12.25,\ \rho = 0.010$

0.86	-0.00	0.56	0.02			
2.11	-0.01	1.21	0.00	1.75		
3.00	-0.04	1.69	0.00	1.26	0.06	
5.73	-0.10	2.96	-0.00	2.12	0.00	3.64

D $\quad M = 520,\ \bar{X} = 22.67,\ \bar{Y} = 25.82$; $\sigma_x = 11.37,\ \sigma_y = 11.37,\ \rho = 0.046$

0.23	0.05	-0.23	-0.01			
1.29	-0.09	0.83	0.00	1.87		
0.59	0.10	-0.24	0.10	-0.66	-0.04	
1.94	-0.19	0.96	-0.16	1.46	0.02	4.20

E $\quad M = 450,\ \bar{X} = 19.14,\ \bar{Y} = 25.45$; $\sigma_x = 8.20,\ \sigma_y = 12.75,\ \rho = -0.050$

0.81	-0.05	0.16	0.00			
2.27	-0.03	1.30	-0.00	1.72		
3.27	-0.13	1.26	-0.05	0.37	0.02	
7.07	-0.16	3.30	-0.02	2.31	-0.02	3.37

F $\quad M = 383,\ \bar{X} = 16.27,\ \bar{Y} = 26.98$; $\sigma_x = 6.97,\ \sigma_y = 12.03,\ \rho = 0.298$

0.74	0.65	0.03	-0.12			
2.76	0.47	1.03	0.05	1.85		
3.05	2.40	0.70	0.93	0.15	-0.62	
10.77	2.93	3.19	0.66	1.80	0.09	4.07

G $\quad M = 376,\ \bar{X} = 16.64,\ \bar{Y} = 24.99$; $\sigma_x = 10.88,\ \sigma_y = 11.53,\ \rho = -0.099$

0.43	0.06	0.37	0.25			
1.51	0.19	0.94	0.29	1.86		
1.28	0.32	0.88	0.47	0.91	0.91	
2.87	0.61	1.56	0.80	1.63	0.97	4.26

H $\quad M = 599,\ \bar{X} = 23.08,\ \bar{Y} = 25.15$; $\sigma_x = 11.06,\ \sigma_y = 12.31,\ \rho = 0.032$

0.03	0.00	-0.01	0.05			
1.22	-0.01	1.04	0.01	1.77		
0.08	0.01	0.03	0.07	-0.01	0.19	
1.73	-0.00	1.40	0.01	1.86	0.05	3.62

I $\quad M = 288,\ \bar{X} = 20.25,\ \bar{Y} = 27.00$; $\sigma_x = 2.53,\ \sigma_y = 12.12,\ \rho = -0.209$

0.16	0.04	0.02	-0.01			
3.38	0.04	1.58	0.05	1.76		
1.74	0.34	0.46	0.09	0.04	-0.03	
21.02	0.04	7.56	0.34	3.54	0.26	3.61

J $\quad M = 326,\ \bar{X} = 18.36,\ \bar{Y} = 25.88$; $\sigma_x = 3.66,\ \sigma_y = 13.68,\ \rho = 0.345$

-0.65	-0.45	-0.39	0.06			
2.89	0.52	1.35	0.13	1.71		
-4.20	-1.83	-1.75	-0.54	-0.92	0.23	
12.55	2.98	4.69	0.94	2.51	0.29	3.39

K $\quad M = 529,\ \bar{X} = 17.76,\ \bar{Y} = 29.03$; $\sigma_x = 8.81,\ \sigma_y = 11.99,\ \rho = -0.288$

0.39	-0.21	0.15	0.35			
2.21	0.52	1.41	0.20	1.97		
2.06	0.21	1.22	0.32	0.49	1.50	
7.75	3.20	4.40	1.88	3.11	0.80	4.83

L $\quad M = 363,\ \bar{X} = 21.39,\ \bar{Y} = 28.52$; $\sigma_x = 7.89,\ \sigma_y = 13.04,\ \rho = -0.281$

0.46	-0.84	0.03	0.26			
4.28	-0.19	1.09	0.13	1.71		
0.32	-4.15	-0.15	-0.87	0.01	1.00	
26.11	2.17	4.69	0.55	1.66	0.42	3.59

M $\quad M = 618,\ \bar{X} = 19.13,\ \bar{Y} = 23.34$; $\sigma_x = 10.91,\ \sigma_y = 11.37,\ \rho = -0.031$

-0.13	0.15	0.03	-0.12			
1.53	0.14	1.14	0.01	1.88		
-0.62	0.19	-0.17	0.10	0.10	-0.55	
3.08	0.65	1.97	0.39	2.26	0.06	4.32

N $\quad M = 391,\ \bar{X} = 19.57,\ \bar{Y} = 26.38$; $\sigma_x = 8.84,\ \sigma_y = 11.49,\ \rho = -0.240$

-0.33	-0.10	-0.11	-0.16			
2.74	1.18	1.66	0.36	1.90		
-1.84	-0.86	-0.98	-0.62	-0.48	-0.83	
10.46	6.19	5.71	3.41	4.03	1.49	4.51

O $\quad M = 447,\ \bar{X} = 19.65,\ \bar{Y} = 25.77$; $\sigma_x = 12.53,\ \sigma_y = 9.95,\ \rho = -0.004$

0.01	0.01	-0.03	-0.01			
1.20	0.01	0.66	-0.01	2.21		
0.20	0.01	-0.04	0.03	-0.08	-0.03	
1.60	0.01	0.63	-0.02	1.04	-0.04	6.33

P $\quad M = 429,\ \bar{X} = 22.23,\ \bar{Y} = 30.76$; $\sigma_x = 8.76,\ \sigma_y = 11.69,\ \rho = 0.321$

0.35	0.54	-0.17	-0.44			
1.99	-0.08	0.70	-0.13	2.06		
1.05	1.29	-0.29	0.49	-0.31	-1.94	
5.13	-0.30	1.37	-0.47	1.07	-0.30	5.47

Q $\quad M = 517,\ \bar{X} = 24.77,\ \bar{Y} = 37.36$; $\sigma_x = 11.50,\ \sigma_y = 13.15,\ \rho = -0.056$

-0.10	0.27	0.09	-0.31			
1.39	0.01	0.52	-0.12	2.28		
-0.29	0.40	0.04	0.22	0.39	-1.46	
2.18	-0.00	0.56	-0.07	0.81	-0.63	6.66

R $\quad M = 560,\ \bar{X} = 21.04,\ \bar{Y} = 37.97$; $\sigma_x = 8.87,\ \sigma_y = 11.95,\ \rho = -0.045$

0.13	-0.08	0.02	0.04			
1.39	0.01	1.07	-0.04	1.75		
0.37	-0.20	0.16	-0.22	-0.04	0.15	
2.39	-0.01	1.70	-0.13	1.85	-0.15	3.63

TABLE 2. MOMENTS OF PRINTED CHARACTERS (Continued)

S — $M = 328$, $\bar{X} = 15.32$, $\bar{Y} = 38.59$, $\sigma_x = 7.56$, $\sigma_y = 10.87$, $\rho = 0.034$

-0.06	-0.15	-0.11	-0.16			
1.60	0.04	1.06	0.29	2.52		
-0.19	-0.46	-0.32	-0.24	-0.14	-1.00	
2.96	0.13	1.74	0.55	1.99	0.73	8.68

T — $M = 399$, $\bar{X} = 23.97$, $\bar{Y} = 46.76$, $\sigma_x = 7.10$, $\sigma_y = 13.03$, $\rho = -0.043$

0.08	0.75	0.01	-0.36			
3.79	-0.00	1.03	-0.04	1.74		
0.56	3.16	0.14	0.60	0.07	-1.40	
16.50	0.15	3.54	-0.19	1.52	-0.18	3.80

U — $M = 349$, $\bar{X} = 17.43$, $\bar{Y} = 40.85$, $\sigma_x = 8.57$, $\sigma_y = 12.42$, $\rho = -0.084$

1.47	0.23	0.22	-0.07			
3.72	0.61	1.16	-0.09	1.87		
7.94	1.78	1.91	0.47	0.77	-0.58	
18.47	4.66	4.64	1.25	2.38	-0.70	4.83

V — $M = 351$, $\bar{X} = 16.99$, $\bar{Y} = 46.56$, $\sigma_x = 6.91$, $\sigma_y = 11.28$, $\rho = -0.243$

1.23	0.88	0.06	-0.14			
4.18	1.38	1.25	0.10	1.87		
9.16	5.01	1.92	1.55	0.20	-0.77	
26.73	11.81	6.88	2.67	2.30	0.35	4.38

W — $M = 645$, $\bar{X} = 27.01$, $\bar{Y} = 29.39$, $\sigma_x = 10.03$, $\sigma_y = 11.50$, $\rho = -0.126$

0.40	0.46	0.09	-0.16			
2.60	0.42	1.14	0.03	1.87		
3.13	2.40	0.72	0.68	0.28	-0.93	
10.95	3.60	3.51	0.80	2.09	-0.06	4.61

X — $M = 405$, $\bar{X} = 23.00$, $\bar{Y} = 28.38$, $\sigma_x = 7.38$, $\sigma_y = 12.49$, $\rho = -0.459$

0.29	0.11	-0.00	-0.01			
4.37	1.80	1.70	0.10	1.69		
3.41	1.41	0.47	0.15	-0.03	-0.05	
30.69	14.90	8.92	3.78	3.42	0.33	3.31

Y — $M = 369$, $\bar{X} = 17.72$, $\bar{Y} = 27.95$, $\sigma_x = 6.16$, $\sigma_y = 12.53$, $\rho = -0.227$

1.02	0.86	-0.04	-0.18			
4.52	1.34	1.45	0.10	1.76		
9.59	5.51	1.68	1.35	-0.04	-0.68	
32.49	12.93	7.72	2.48	2.67	0.32	3.65

Z — $M = 374$, $\bar{X} = 31.16$, $\bar{Y} = 23.62$, $\sigma_x = 7.92$, $\sigma_y = 12.14$, $\rho = 0.256$

-0.13	0.06	0.02	0.06			
2.57	-1.01	1.44	-0.06	1.62		
-0.82	0.54	-0.30	0.29	0.05	0.24	
8.46	-4.32	3.81	-1.47	2.51	-0.16	3.02

1 — $M = 200$, $\bar{X} = 22.95$, $\bar{Y} = 22.48$, $\sigma_x = 2.60$, $\sigma_y = 12.01$, $\rho = -0.234$

-0.70	-0.37	-0.11	-0.09			
6.61	-0.57	1.93	-0.05	1.73		
-7.84	-4.10	-1.68	-1.06	-0.31	-0.31	
66.65	-5.55	15.17	-1.10	4.28	-0.13	3.42

2 — $M = 316$, $\bar{X} = 19.80$, $\bar{Y} = 20.14$, $\sigma_x = 6.64$, $\sigma_y = 13.38$, $\rho = 0.213$

-0.29	-0.04	-0.04	0.02			
1.87	-0.28	0.97	-0.04	1.38		
-1.06	-0.04	-0.30	-0.15	-0.10	0.20	
4.40	-1.02	1.84	-0.30	1.19	-0.13	2.16

3 — $M = 278$, $\bar{X} = 32.82$, $\bar{Y} = 22.85$, $\sigma_x = 6.31$, $\sigma_y = 11.83$, $\rho = -0.055$

-0.65	-0.22	-0.24	-0.00			
1.94	0.14	1.10	0.09	1.61		
-2.25	-0.59	-1.09	-0.33	-0.52	-0.02	
4.78	0.70	2.35	0.40	1.64	0.20	3.00

4 — $M = 262$, $\bar{X} = 26.08$, $\bar{Y} = 22.09$, $\sigma_x = 4.67$, $\sigma_y = 10.74$, $\rho = -0.049$

-1.02	-0.28	0.16	0.17			
5.07	0.07	0.49	0.08	1.92		
-9.34	-1.78	-0.02	-0.18	0.41	0.65	
35.72	2.35	1.43	-0.08	0.72	0.37	4.54

5 — $M = 275$, $\bar{X} = 17.11$, $\bar{Y} = 23.09$, $\sigma_x = 6.58$, $\sigma_y = 12.85$, $\rho = -0.130$

-0.49	-0.32	-0.18	0.19			
1.90	0.41	0.94	0.21	1.60		
-1.88	-1.10	-0.72	-0.29	-0.34	0.44	
4.74	1.72	1.87	0.82	1.38	0.51	2.88

6 — $M = 340$, $\bar{X} = 42.94$, $\bar{Y} = 29.87$, $\sigma_x = 7.78$, $\sigma_y = 9.67$, $\rho = -0.124$

0.15	-0.18	0.11	0.28			
1.45	0.05	0.68	0.34	2.38		
0.46	-0.28	0.21	0.13	0.78	1.76	
2.48	0.08	0.86	0.54	1.61	1.98	7.70

7 — $M = 209$, $\bar{X} = 22.71$, $\bar{Y} = 36.00$, $\sigma_x = 5.27$, $\sigma_y = 14.16$, $\rho = 0.041$

-0.12	0.66	-0.01	-0.55			
2.78	-0.28	0.80	-0.07	1.61		
-0.48	1.32	0.41	-0.05	0.23	-1.82	
10.42	-1.79	2.70	-1.06	1.51	-0.66	3.61

8 — $M = 392$, $\bar{X} = 47.43$, $\bar{Y} = 33.01$, $\sigma_x = 7.89$, $\sigma_y = 11.09$, $\rho = 0.029$

0.02	-0.15	-0.01	0.11			
1.34	-0.01	0.92	0.03	1.76		
0.04	-0.37	-0.02	-0.10	-0.04	0.36	
2.15	-0.01	1.20	0.06	1.36	0.17	3.83

9 — $M = 320$, $\bar{X} = 30.09$, $\bar{Y} = 33.72$, $\sigma_x = 7.94$, $\sigma_y = 9.51$, $\rho = -0.137$

-0.14	0.18	-0.10	-0.35			
1.41	0.03	0.64	0.31	2.49		
-0.43	0.30	-0.17	-0.11	-0.78	-2.24	
2.36	0.05	0.77	0.45	1.51	1.99	8.67

At least the following, however, differ:

	m_{41}	m_{14}	m_{05}	m_{60}	m_{15}
O	0.01	−0.08	−0.03	1.60	−0.04
Q	0.40	0.39	−1.46	2.18	−0.63

Even if it should be decided, as we shall suggest later on, to omit the moments of order 6, there are still three moments of order 5 (m_{41}, m_{14}, and m_{05}) which

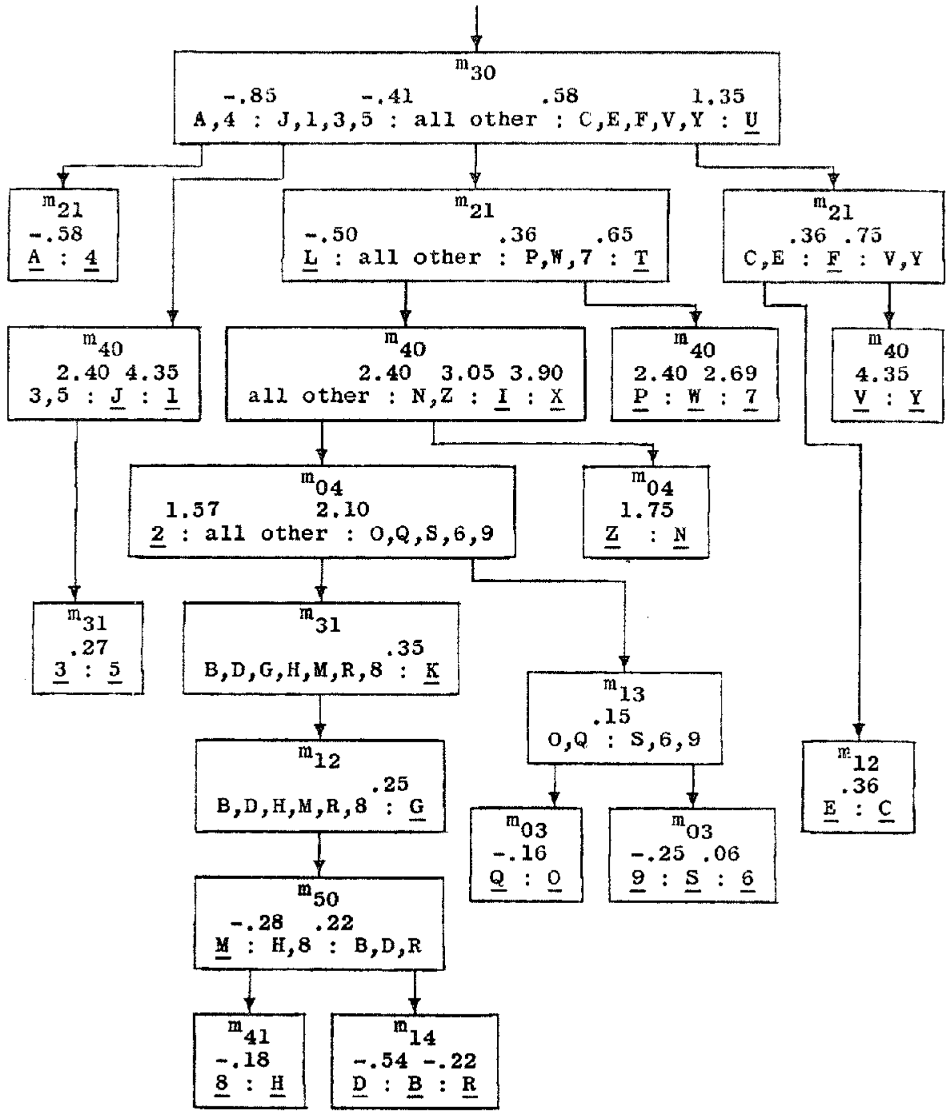


FIG. 4. Flow chart of discrimination procedure

differ greatly; and even m_{03} can be used to distinguish the characters O and Q. Similarly, the characters C and G, another pair of great similarity, differ significantly in several of their moments.

Having obtained the moments of the 35 characters of Table 2, suppose we scan an unknown pattern p and ask to which of the 35 standard patterns p is most similar. One possible algorithm for doing this is shown in Figure 4. This diagram should be interpreted as follows: Starting at the top, the first box indicates that the moment m_{30} of p is the first one to be examined. If $m_{30} \leqq -.85$, then p represents one of the characters A or 4; if $-.85 < m_{30} \leqq -.41$, p represents a J, 1, 3 or 5; etc. If $m_{30} > 1.35$, p represents the character U, and the process is terminated. If m_{30} is in one of the other classes, we proceed to the next box indicated in the flow diagram, and continue until the character has been identified. This is accomplished in a maximum of 12 comparisons, with an average[6] of 6.9. The procedure is based on the assumption that p is indeed one of the 35 characters, and not some other pattern. It is also possible to design procedures which will make it unlikely that a non-character is accepted in place of a character. To exclude such an error with complete certainty, however, is impossible, since we do not know whether there exist stray patterns which agree with one of the characters in all moments up to the sixth order.

The procedure of Figure 4 is by no means unique, nor is it optimal. It was obtained by trial and error. The problem of finding discrimination procedures in general is discussed in the next section.

5. *Methods of Discrimination*[7]

This problem of how to obtain systematic processes for classifying a pattern as belonging to a certain class is a special case of a far more general problem. The same question arises, e.g. in the recognition of spoken words by machine [3], in the classification of colors, etc. To visualize this problem, let us interpret the 22 moments $m_{jk}(p)$ of a pattern p as coordinates of a point in 22-dimensional space. (In other words, consider the function F introduced in Section 1 as being vector-valued.) Then there corresponds a point $F(p)$ in 22-space to each pattern p. Thus, there is a point $F(\mathrm{A})$ corresponding to the standard "A". If the new pattern p also represents the character A, then $F(p)$ will be close to $F(\mathrm{A})$ in some sense, though it will not necessarily be identical with $F(\mathrm{A})$. To different patterns p representing the character A correspond points which lie in some domain of 22-space, and we may consider them as defining a *probability distribution function* there. Corresponding to the 35 characters we are studying, we have 35 p.d.f.'s; given a point $F(p)$ in 22-space, our problem is to decide from which (if any) of the 35 populations p is most probably taken.

In this terminology it is easy to state how the compilation of moments has simplified our problem. It was originally a problem of classification in the space of all patterns; this has now been reduced to a Cartesian space of 22 dimensions.

This problem has been extensively studied by statisticians; see e.g. Anderson [5, pp. 142–147]. Its solution can be found if the joint probability distribution of the moments is known. The solution becomes tolerably simple if some (not very realistic) assumptions are made, e.g. that the distributions of the moments are independent and have different means but are otherwise identical. Sebestyen [6] describes a solution of a more restricted problem in which the distribution of the moments need not be known a priori.

[6] Assuming p is equally likely to be any one of the 35 characters.

[7] The author is indebted to Dr. Joan R. Rosenblatt for illuminating discussion of this section.

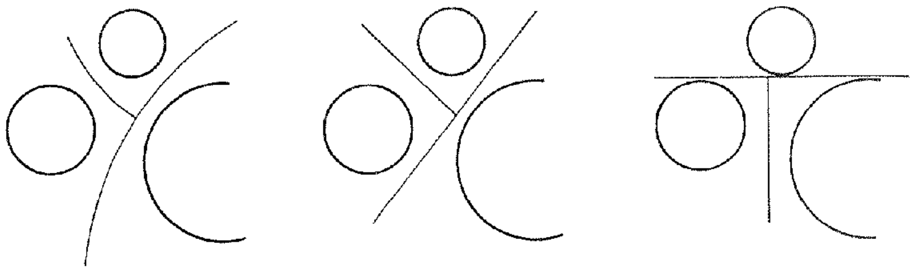


FIG. 5. Three ways of separating plane regions

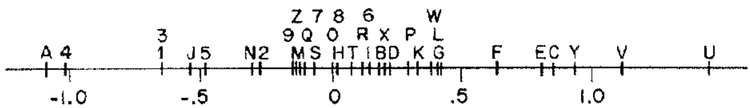


FIG. 6. Discrimination on the moment m_{30}

All the recognition procedures enumerated so far are derived from a probabilistic model of patterns. All but the simplest of them are probably prohibitively cumbersome. It is possible, however, to take an entirely different approach, which, while theoretically unsatisfactory, leads to acceptable results more quickly.

As we said before, to each pattern p corresponds a point $F(p)$ in the moment space, and the points corresponding to all possible patterns representing the same character lie in a certain region of this space.[8] If there were two regions, rather than 35, we might choose some hypersurface in the moment space which separates the two regions, and which is represented by a simple equation. Then it would merely be necessary to substitute the coordinates of $F(p)$, i.e. the moments of p, in this equation in order to decide on which side of the hypersurface $F(p)$ lies. Similarly, we can try to find a system of hypersurfaces which separate all 35 regions. We want their equations to be as simple as possible; for instance, we would first try to accomplish the separation by means of hyperplanes, since their equations are linear. Still simpler, if possible, would be hyperplanes parallel to the coordinate axes. These three possibilities are diagrammed in Figure 5 for the case of three domains in a plane, separated (a) by two algebraic curves, (b) by two straight lines, (c) by two straight lines parallel to the axes. The problem of finding such separating hyperplanes for 35 domains in 22-space is much more difficult, and to the author's knowledge no systematic procedure has been developed for this problem. A possible trial-and-error procedure runs as follows.

Select one coordinate, say m_{30}, and plot its values for all observed patterns

[8] We hope that these regions are mutually exclusive. If they are not, then either there are different patterns which agree in their first 22 moments—in which case we would have to take higher moments—or the pattern p might represent either of two characters, so that the problem of recognition has no solution. The probabilistic approach, on the other hand, applies to the case where the regions are overlapping.

which have been identified as characters. This is shown in Figure 6. We inspect this set of values of m_{30} for gaps, and place a discrimination point in each major gap. For instance, $m_{30} \gtrless -.8$ discriminates between the characters A and 4 on one side and all other characters on the other side. Similarly, $m_{30} > 1.3$ distinguishes the character U from all others, $-.8 < m_{30} < -.4$ characterizes J, 1, 3, 5, and $.5 < m_{30} < 1.3$ characterizes C, E, F, V, Y. Thus the set of 35 characters has been divided into five subsets. We next turn to another moment, say m_{21}, plot a similar diagram separately for each of the five subsets, insert discrimination points for further subdivision of subsets, etc. In this way we arrive at the procedure shown schematically in Figure 4.

The process can be mechanized fairly easily. The effectiveness of the process depends strongly on the order in which different moments are introduced. We try to use first those moments which we suspect of having strong discriminating power; and, other things being equal, use low-order moments before high-order ones.

A variant of this process is to replace Figure 6 by a plot using two moments as plane Cartesian coordinates. One can then by inspection draw some vertical and horizontal lines which decompose the set of characters into subsets. This procedure is more powerful and takes fewer steps; it is less easily mechanized; it makes it still more difficult to decide on the order in which moments are introduced.

Occasionally it may happen that one character falls on both sides of a discrimination point. This is not necessarily bad. It may sometimes be avoided

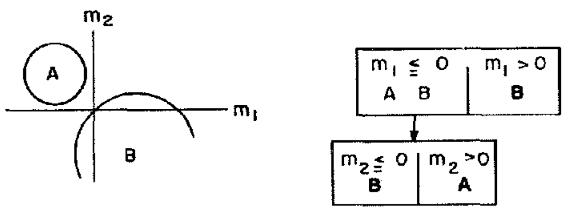


FIG. 7. One character on both sides of a discrimination point

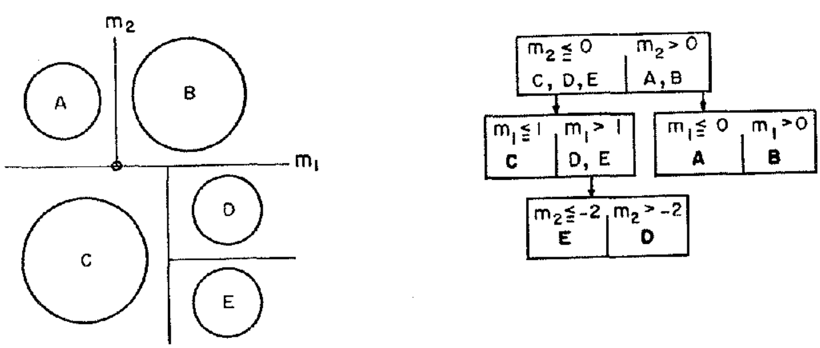


FIG. 8. Discrimination with repeated use of one moment

by different ordering of the moments, but sometimes it is unavoidable. For example, take the case of two domains A, B and two moments m_1 and m_2 shown in Figure 7. Sometimes it may be desirable to return to a coordinate which has been used before, as in the example of Figure 8.

6. *Sensitivity*

In planning the experiments, the intention was to use a rather coarse grid and see whether this would give any meaningful results at all, before going on to finer grids. This seemed desirable because the computing effort is greatly affected by grid size—doubling the fineness will almost quadruple computing time—and because it was felt that any applications of the method would be practical only if a coarse grid were adequate. On the other hand, it seemed reasonable to compute a generous number of moments to begin with and omit some later if this proved possible. Adding a line of moments at the time of the original computation causes only a small increase in computing effort, while computing such moments later, if the need for them should arise, would require a repetition of a large part of the work. In both respects, the original tentative choices—grids of about 30 x 40 points per character, moments up to the sixth order—proved to be correct.

As may be seen from Figure 4, the moments of order 6 are not needed for the discrimination of the 35 specimens we have studied, and those of order 5 occur only near the end of the process. Undoubtedly the need for the higher order moments will become more acute when we try to recognize characters taken from different type fonts as identical; or when we wish to discriminate among more classes of characters, e.g. distinguish between lower case and capital letters. Even so, there is no foreseeable need for moments of order exceeding 6. Also, there are indications that higher order moments are increasingly sensitive to noise, which makes their use undesirable.

The grid size is probably as coarse as we dare make it. One indication of this is the fact that the thinner lines of a letter are read by the scanner as having the width of one cell, as may be seen from Figure 3; if the grid were coarser there would be danger of confusing thin lines with noise. Also, the effect of noise in general, i.e. the changes in the magnitude of the moments caused by random change of a few cells, would probably be intolerably large.

A related question concerns the sensitivity of the scanning device: how much black must there be in a mixed black-and-white cell before the scanner records it as black. The present study throws no light on this question.

7. *Summary and Outlook*

What has been accomplished by the foregoing study is a demonstration of the fact that a small number of moments is adequate to characterize certain patterns and discriminate among the patterns of a certain set, such as alphabetical and numerical characters.

FIG. 9. Two Hebrew characters, d and r, differ only in the sharpness of a corner. In the Latin alphabet such differences are disregarded. Three patterns differing in corners and curvature all represent the same character, A.

In comparing this system with others which have been proposed we find both advantages and drawbacks. These other methods use either coincidence—the pattern to be read is matched up with a standard pattern, and complete agreement, except for a specified tolerance, is demanded—or they observe certain local or topological properties of the character to be read, such as corners, branch points, and closed loops. An example of what can not be done with moments is furnished by the modern Hebrew alphabet, in which e.g. the characters corresponding to D and R differ only in that the former has a sharp corner in a place where the latter is rounded. This difference would have no more effect on moments than some slight noise or change in type font. In fact, it is the kind of distinction which we wish to disregard, for in the latin alphabet it is frequently meaningless (see Figure 9). If characters are distinguished by features of this kind, one needs a procedure which is sensitive to them. The moments are sensitive to global, rather than to local, features of a pattern; that is to say, to the way in which blackness and whiteness is distributed around in a pattern. Also, the method of moments makes it especially easy to recognize as identical two patterns which differ in location, size or any of the affine invariants enumerated in Section 1.

Before the value of moments for pattern recognition, and more specifically for automatic print reading, can be judged, several questions will have to be explored, mostly by computer experiments.

The study should be extended to lower case characters.

Moments should be computed for characters from different type fonts, to see whether it is possible to find one single discrimination procedure which is valid for all fonts, or whether it is necessary to use a separate procedure for each font. The former would be a strong argument in favor of moments.

To develop such a procedure may require more powerful methods than the trial and error method using one moment after another, by which the procedure of Figure 4 was derived. Such methods should be studied; in particular, the joint distribution of moments should be determined by sampling.

The effect of changes in grid size should be investigated. A few tests should be run using as fine a grid as can be obtained from the scanner. Coarser grids can be simulated on the computer, by lumping a number of cells. In the same way one can simulate scanners of varying sensitivity: a block of cells may be assigned the value "black" if it contains, say, 10 per cent or 50 per cent black cells, etc.

The effect of noise should be studied, either by analyzing imperfect specimens of characters or by generating noise on the computer. Also, it is possible to derive theoretical expressions for the effect of noise, based on assumed noise distributions. The experiments reported here throw little light on this question, since only rather clearly defined characters have been used. There is a danger that noise in the form of black spots near the edge of the field of vision might have a large effect on moments, especially those of high order.

Some of these extended studies are now in progress, and will be reported in a forthcoming paper.

REFERENCES

1. Kirsch, R. A.; Cahn, L.; Ray, C.; and Urban, G. H. Experiments in processing pictorial information with a digital computer. Proc. Eastern Joint Computer Conf. 1957, Washington, D. C., pp. 221–229. Published by Inst. Radio Engrs. New York (1958).
2. Greenwald, S.; Haueter, R. C.; and Alexander, S. N. SEAC. *Proc. IRE 41* (1953), 1300–1313.
3. Fatehchand, R. Machine recognition of spoken words. *Advances in Computers*, Vol. 1, pp. 193–229 (Academic Press, New York, 1960).
4. Davis, P.; and Pollak, H. On the analytic continuation of mapping functions. *Trans. Amer. Math. Soc. 87* (1958), 198–225.
5. Anderson, T. W. *Introduction to Multivariate Statistical Analysis*. Wiley, New York, 1958.
6. Sebestyen, G. S. Recognition of membership in classes. *IRE Trans. IT-7* (Jan. 1961), 44–50.
7. Ming-Kuei Hu. Pattern recognition by moment invariants (letter to the editor). *Proc. IRE 49* (Sept. 1961), 1428.