R. G. Casey

Moment Normalization of Handprinted Characters

Abstract: Handprinted characters can be made more uniform in appearance than the as-written version if an appropriate linear transformation is performed on each input pattern. The transformation can be implemented electronically by programming a flying-spot raster-scanner to scan at specified angles rather than only along specified axes. Alternatively, curve-follower normalization can be achieved by transforming the coordinate waveforms in a linear combining network. Second-order moments of the pattern are convenient properties to use in specifying the transformation. By mapping the original pattern into one having a scalar moment matrix all linear pattern variations can be removed. Comparison experiments with three sets of handprinted numerals showed that error rates were reduced by integral factors if the patterns were normalized before scanning for recognition.

Introduction

An optical character recognition (OCR) device performs, albeit in a constrained and rigidly mechanized fashion, a function normally considered to be a human cognitive activity. The OCR unit senses the spatial patterns of black and white on a document containing printed symbols, and identifies each pattern one by one as an A, a 7, a comma, or some other member of the alphabet of symbols that the machine is designed to read. The identification is recorded internally by means of a machine code.

The initial observation of a pattern is done by an optical scanner. This is a device for obtaining an electronic representation of a spatial pattern—a TV camera is an example, although other types of scanners are used more commonly in character recognition. The electronic version of the pattern may be a pair of waveforms describing the contour of the character or, more typically, a raster portraying the intensities of black and white within the pattern field.

Scanning a character for recognition also requires isolating it from the other objects in the field of view, and rejecting noise to obtain a "clear" representation. Thus, in addition to the basic optical scanning apparatus, the scanner contains logical and control circuitry for determining the portion of the document to be scanned and for filtering the electronic image. If the subsequent recognition processing is to be done digitally, the image is digitized, typically into a binary format in which a 0 bit denotes white and a 1 bit denotes black.

The author is located at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598.

It is this stored representation on which the recognition process is based. Further stages of processing consist only of comparing this image, or information obtained from it, with what the system has been designed to expect from the respective character types. Thus, scanning is a critical operation. If the pattern provided by the scanner differs sufficiently from the ideal pattern as built into the recognizer unit, the symbol being examined will not be classified correctly.

This paper concerns a scanning system especially suitable for handprinted symbols. The system objective is to obtain an electronic image that more closely matches the machine ideal than did the original printed form. To illustrate this notion we first discuss some common ways of recognizing scanned images.

- *Recognition logic*

One of the standard procedures for character recognition is to compare the states of elementary areas of the pattern to be identified with the states of corresponding regions of several stored prototype characters.[1,2] This "template-matching" concept includes linear weighting schemes in which the elements, typically binary bits denoting black and white areas of a scanned pattern, are weighted in each pattern class according to their relative importance in specifying that class.

Such "global" matching techniques are to be distinguished from systems that detect local features of a pattern, e.g., corners, indentations and concave and convex curves, and require the locations of these features to match those of the prototype in only a very coarse way. Basing recognition on a global match implies that successive samples of characters in a given category possess a high degree of spatial invariance; they should superimpose very well. It has been found that such matching procedures are useful for typewritten and machine-printed characters, which exhibit uniformity within a given font and also among selected fonts. On the other hand, because of the range of styles in unconstrained hand printing, such material has been recognized more reliably by the method of local feature detection.

Figure 1 illustrates this dichotomy. Ten typewritten 4's in each of three font styles were raster-scanned and then superimposed. A large proportion of the cell positions were "stable," i.e., they had the same state in the majority of samples. Thus the outline of a 4 is well defined in the quantized array shown in Fig. 1(a). The same experiment repeated with 100 handprinted 4's from 15 different writers resulted in the frequency distribution shown in Fig. 1(b). Only a few bit positions were reliably "black." These were not distributed broadly enough to allow the 4's to be discriminated from other handprinted numbers by template matching.

From evidence such as Fig. 1, and from previous recognition experience, it is clear that large-area templates cannot be designed to match hand printing produced under field conditions by a number of writers. The following sections of this paper describe an attempt to improve this situation by electronically scanning a printed character in such a way as to produce a transformed character that better matches the stored representation.[3] This is done by calculating several global pattern moments and transforming the character to place these moments in a standard form. The new "normalized" characters are shown to have properties of invariance that are desirable in a recognition framework. Experiments are reported that demonstrate the improvement in video quality and indicate the recognition gains that are possible with the aid of the normalization technique.

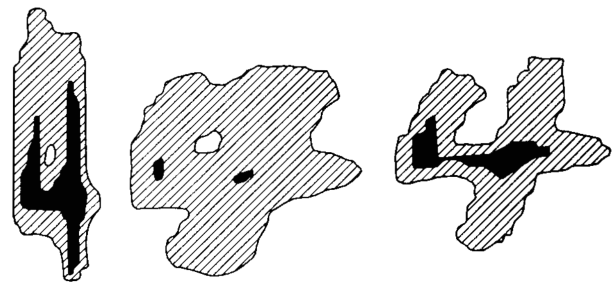


Figure 1 Superimposed characters: (a) typewritten, (b) and (c) handprinted. The shaded regions consist of points of the pattern field that were black, with a frequency within the threshold bounds. Solid black and white regions denote higher and lower frequencies of black, respectively.

• *Related work*

A system for eliminating linear variations in a pattern by transforming it so that the pattern moments are in a canonical form has been described previously by Udagawa, Toriwaki and Sugino.[4] That method differs from the one to be described in that the moments are not normalized to the area of the pattern. In addition, third- and fourth-order moments were calculated in Ref. 4 to obtain a rotationally invariant form. Only a few sample characters could be tested due to equipment restrictions (for example, no scanner was available).

Hu[5] and Alt[6] also used central moments for pattern classification and observed the basic property that diagonalizing the second-order moment matrix eliminates linear distortions. In these cases the moments were computed primarily to serve as recognition parameters.

Bakis, Herbst and Nagy[7] implemented skew corrections on handprinted characters by translating horizontal rows of bits to make the xy moment vanish. Since in their method character height is normalized while the pattern is scanned, their system differs from that described below mainly by the absence of width normalization.

Normalization technique

• *Skewed scanning*

We offer the following nonexhaustive list of characteristics in which handprinted samples vary.

1) Size (height and width)
2) Slant (and rotation)
3) Line thickness
4) Style (such as open- and closed-top 4's, looped and nonlooped 2's, etc.)
5) Ornamentation
6) Stroke proportions (relative widths of the various strokes constituting the character)
7) Stroke regularity (smoothness and well-formedness of lines and curves)

Variations in any of these characteristics degrade template-matching performance. To some extent the degree of variation is controllable by training, by using standard writing instruments, and by supervision. Under common field conditions, e.g., in making out sales checks in a department store, the amount of control is limited, as one can see by comparing hand printing produced by salesclerks and by office personnel in a particular department store (Fig. 2).

Except under the most rigid conditions, one cannot expect writers to accurately reproduce handprinted characters. In an effort to decrease the amount of varia-

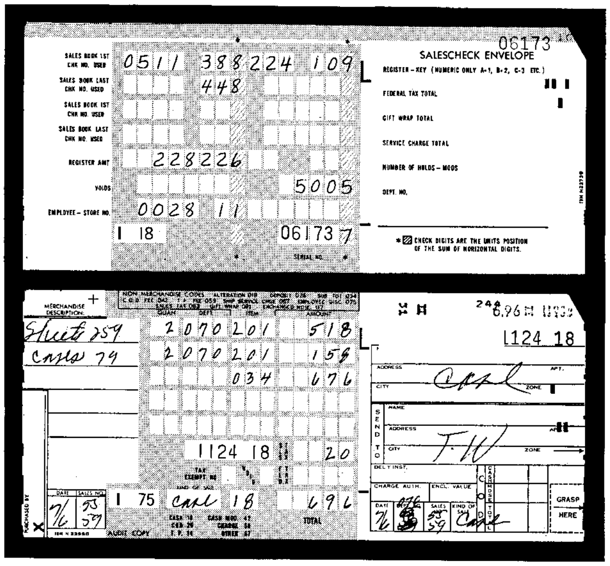
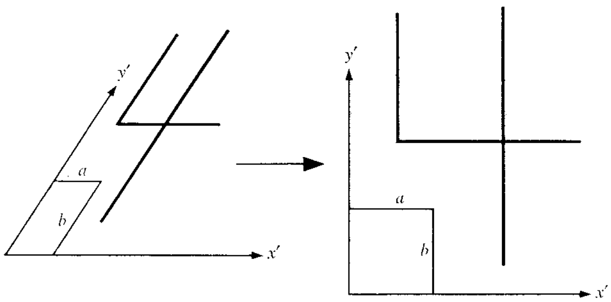


Figure 2 Sample handprinted characters. These numerals, printed by careful clerks in an unhurried environment (top) and by sales personnel (bottom) in the same department store, were scanned and used for the experiments described in the text.

Figure 3 Linear pattern transformation. The original pattern (left) is scanned obliquely in the y' direction and this scan line is stepped in the x' direction, the respective sample increments being proportional to b and a. The resulting sample values, plotted orthogonally, are shown on the right. Note that the parallelogram is transformed into a square.



tion, recognition systems incorporating normalization stages have been constructed. With the aid of feedback techniques, parameters such as the height and line thickness of characters can be adjusted to preset standards.[8,9] These operations have proved effective in enhancing the quality of scanned patterns.

If the scanner uses an electronically controlled flying spot, it is common practice to scan the input pattern in a horizontal-vertical raster and to quantize and store the result as a rectangular array of bits. However, by choosing skewed scan directions and by adjusting the sampling intervals at which the state of the pattern is observed, a transformed pattern can be obtained (Fig. 3). If the scans are linear and uniformly spaced, the transformation is linear and can be expressed mathematically as

$$\begin{bmatrix} u \\ v \end{bmatrix} = \mathbf{A} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} , \tag{1}$$

where x and y are the original (rectangular) scan coordinates, x_0 and y_0 define an arbitrary origin, u and v are the coordinates in the transformed system, and $\mathbf{A}$ is a 2×2 transformation matrix. As stored in a core memory, for example, a pattern point that previously was located at cell x, y in a rectangular scan is located at cell u, v after rescanning.

Through adjustment of the scan parameters embodied in matrix $\mathbf{A}$, this skewed scanning procedure permits regulation of the height, width, slant and rotation angle of patterns. While these are not the only characteristics that vary in hand printing (viz., the list at the beginning of this section), they are important ones, especially when the characters are identified by means of a global matching procedure.

A linear mapping is not the only type of transformation obtainable by scanning techniques. Indeed, completely arbitrary transformations are achievable. Any pattern can be changed into any other pattern as desired. Currently, however, reasonable criteria are known only for specifying the linear mappings. Other transformations corresponding to curvilinear coordinate systems or to various projective systems may also be useful and are being studied.

• *Normalized form for patterns*

In the previous section it was noted that properly skewed scan directions compensate for size, slant and rotational variations in patterns. The new scan coordinates, however, depend on the differences of a pattern from the norm. One can easily measure height and width, but slant and rotation are more ambiguous quantities because the character is not yet identified.

People, however, seem to recognize a character before assessing its slant. One procedure that the machine might follow, then, is to hypothesize the category of the input (assume that it is an A, a B, a C, etc. in turn) and try to measure its slant and rotation under these assumptions. These measurements specify an appropriate transformation. Such a procedure has been tried, but it is much more complex (a new transformation is required for each hypothesis) and probably less reliable than the method described here, which is based on the moments of the pattern. We assume, in view of the application to printed characters, that the pattern is binary, e.g., black on white.

However, the properties of the method can be generalized to arbitrary spatial patterns.

The pattern moments consist of the three quantities m_x, m_y and m_{xy}, often arranged in matrix form as follows:

$$\mathbf{M} = \begin{bmatrix} m_x & m_{xy} \\ m_{xy} & m_y \end{bmatrix}.$$

The elements m_x and m_y are the mean-square x and y deviations about orthogonal x and y axes through the centroid of the pattern. They are defined by the expressions

$$m_x = S^{-1} \int_{\mathrm{P}} x^2 \, dS \tag{2}$$

and

$$m_y = S^{-1} \int_{\mathrm{P}} y^2 \, dS, \tag{3}$$

where S is the total area of the pattern **P**. The product term m_{xy} is defined by

$$m_{xy} = S^{-1} \int_{\mathrm{P}} xy \, dS. \tag{4}$$

Suppose that a pattern P is transformed by a linear mapping procedure having matrix **A** into a new pattern P*. The transformation can be expressed in terms of new coordinates u and v that are linearly related to the original coordinates. Upon substitution of Eq. (1) into Eqs. (2) through (4) it is found that the moment matrix **M*** of the new pattern is related to the former moments by the matrix expression

$$\mathbf{M}^* = \mathbf{AMA}', \tag{5}$$

where the prime indicates the transposed matrix.

Unless the pattern is distributed along a line of zero width, any desired moment matrix can be obtained by linear transformation. That is, given **M** and **M*** one can find a matrix **A** such that (5) holds.

One particular form of the transformed moment matrix possesses especially desirable properties. This is the scalar matrix

$$\mathbf{M}^* = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}, \tag{6}$$

where k is an arbitrary constant, specified in advance. The utility of this particular form of the moment matrix is perhaps best illustrated with the aid of the following theorem, which is proved in Appendix 1.

Theorem Let P* be a pattern obtained from a given pattern P by a linear transformation of coordinates. Let a new transformation operate on P to give pattern Q, whose moment matrix has the diagonal form (6). Let another transformation be found to transform P* into a Q* having the same scalar moment matrix as Q. Then there exists a pure rotation of coordinates, a reflection, or a combination of the two that carries Q into Q*.

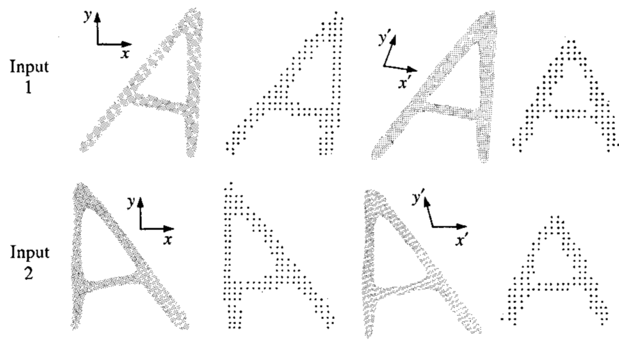


Figure 4 Removal of linear variations by moment normalization. For each input pattern the rescan procedure was computed on the basis of the moments observed on the first scanning pass.

Thus, if two patterns related by a linear transformation are mapped into new patterns having diagonal moment matrices, the transformed patterns are identical except for rotation or reflection (see Fig. 4). Reflection of a nonsymmetric character yields a backwards or an upside-down character, distortions that are not likely to be encountered in handprinted-character recognition. Reflection of a basically symmetric character, such as an A, a 0 or an 8, yields an essentially indistinguishable form after normalization. Hence, to resolve the ambiguity concerning this distortion, it is sufficient to specify that the normalizing transformation be nonreflective.

In practice it often happens that the rotational ambiguity can be handled in a logical manner. The ambiguity is present because values are assigned to only three moments, whereas the transformation is specified by four parameters. The extra degree of freedom can be disposed of by placing one additional constraint on the transformation matrix. Many such constraints can be invented; a particularly appropriate one is described below.

The x direction for scanning a line of print is usually taken as the direction of the line, and hand printing is generally slanted in such a manner that in the x direction only a scale variation is encountered. For example, even in a slanted A the cross bar is more or less horizontal. To transform a pattern to a normalized form without rotational ambiguity requires an operator that diagonalizes the moment matrix without changing the x direction. (Thus horizontal lines remain horizontal.) As shown in Appendix 2, such a transformation is unique.

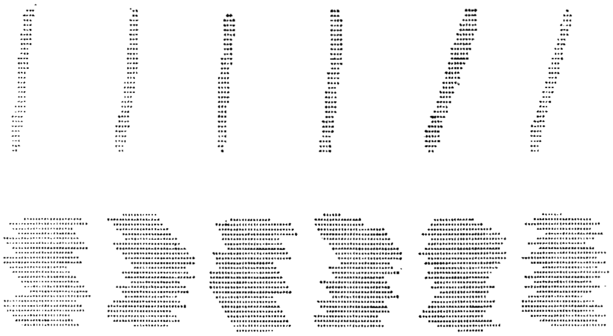
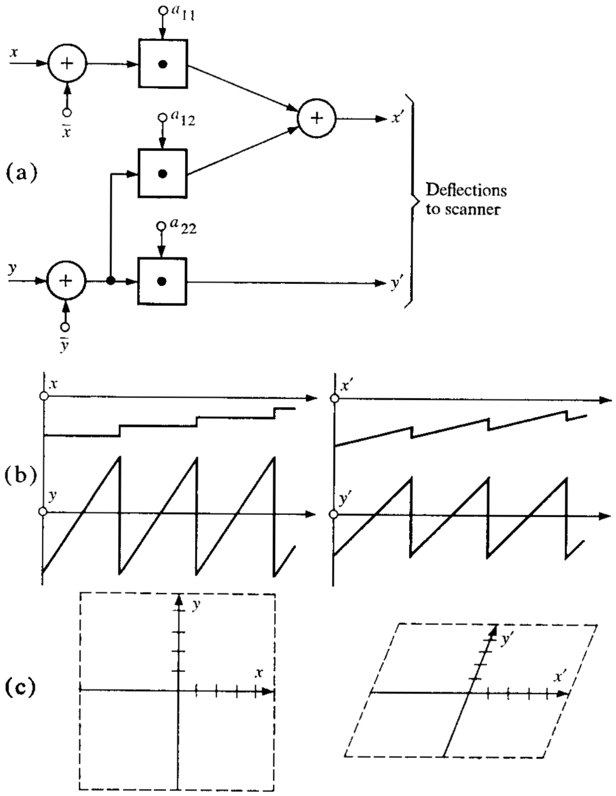


Figure 5 Transformation of handprinted 1's by moment normalization showing the widening of narrow shapes.

Figure 6 Formation of the scanner deflection waveforms: (a) system for combining waveforms, (b) transformation of the fixed input waveforms to obtain normalizing deflection waveforms, and (c) the corresponding scan programs.



• *Geometric patterns*

The only parallelogram having a moment matrix already in the normalized form is a square. Similarly, all ellipses are transformed into circles and all triangles become equilateral in the moment normalization process. When moment normalization is applied to discrete pattern representations, quantization noise and line thickness also affect the pattern obtained.

The widening of narrow shapes into quite different-looking bodies might be expected to introduce a severe recognition problem. Figure 5 illustrates several handprinted 1's as they appeared before and after normalization. For this case it has been found that the minimum eigenvalue of the moment matrix (i.e., the minimum moment of the pattern about any axis) is a measure of width that effectively separates the 1's from the other patterns. Only twelve misclassifications occurred when upper and lower threshold bounds were applied to this eigenvalue in order to sort 12,000 characters (including 3600 1's) into "1" and "not 1" categories. However, this technique was not needed to recognize 1's in the experiments described below because confusion involving the 1 class was not a significant factor. Apparently the expanded 1's (see Fig. 5) were sufficiently different from the other character types.

• *Implementation*

The pattern moments and the transformation parameters can be computed either digitally or by a straightforward analog implementation.

The coefficients of the transformation matrix serve as weights in combining the usual sawtooth y-deflection voltages with stepped x-deflection waveforms to obtain new waveforms that cause a flying spot to travel along oblique parallel paths as shown in Fig. 6. The projected cost of this circuitry constitutes only a modest portion of the overall system cost.

The most likely impediment to a practical analog design is the need for pulsed integrators to compute the moments. The speed of the system would be limited by these components to bit rates of the order of 20 kHz, or about fifty characters per second, if high accuracy (e.g., one percent) is to be maintained in the moment computation. By permitting larger errors (e.g., ten percent), or by integrating digitally with high speed adders, rates of 200 kHz or more are practicable. This latter approach would have the disadvantage of higher cost.

• *Curve-follower normalization*

The moment diagonalization procedure can also be adapted to a scanner using a curve follower.[10] In this device the CRT beam is constrained by feedback techniques to tracing a path along the contour of the input pattern. The control information is provided by the x and y contour coordinates. These x and y waveforms, properly multiplied or squared and then integrated as required by Eqs. (2), (3) and (4), define the moments of the contour. To implement the normalization procedure these moments are computed on the first scan of an input character. The transformation can be computed by either analog or digital hardware. The transformation parameters then serve as the weighting coefficients of a combining network that linearly transforms the x and y waveforms obtained in a second scan of the pattern. In this system, as opposed to the raster mode, the CRT beam follows exactly the same path (namely the contour) in both passes; however, the pattern "seen" by the system corresponds to a normalized pattern—the transformed coordinate waveforms—on the second pass. Sample contours obtained by simulating this system are shown in Fig. 7.

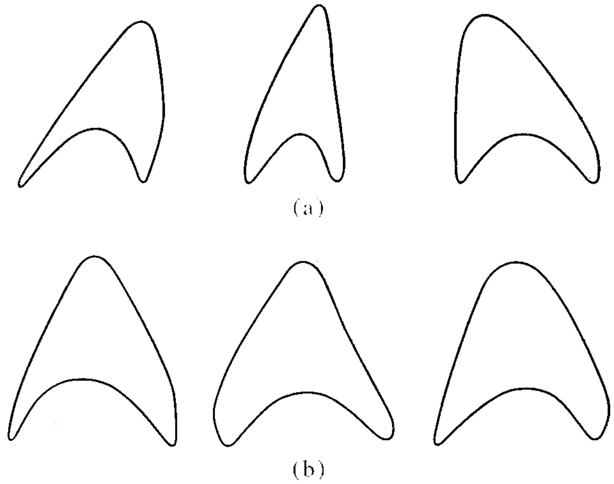
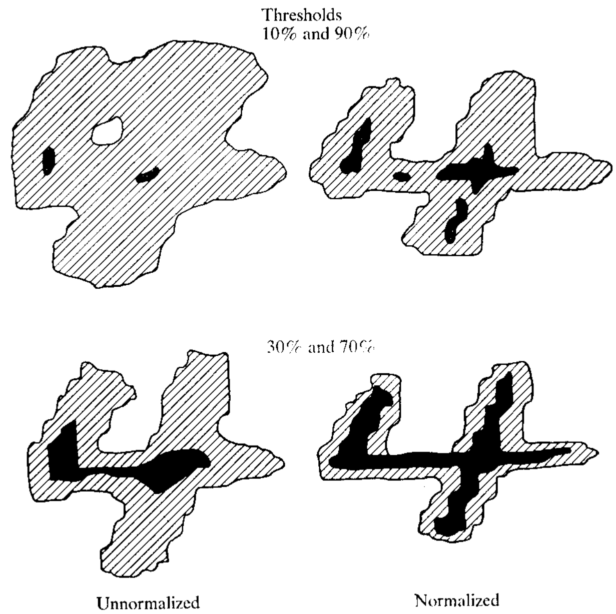


Figure 7 Sample contours: (a) input (derived from handprinted A's) and (b) normalized.

Figure 8 Superimposed patterns (shadings as in Fig. 1).



Experiments

Several experiments have been done with normalized and unnormalized handprinted characters as data. The objects of these experiments were to 1) observe whether the characters were made visually more uniform by the moment normalization technique, 2) exhibit some special features of the normalization procedure, and 3) determine whether recognition performance (particularly of correlation-based recognition systems) could be improved by normalizing the input. The primary data for these experiments were three batches of scanned handprinted numerals. A full description of these data is given in Ref. 7. Two of the data sets, referred to as Backroom 1 and Backroom 2 respectively, were similar in nature since the same writers (four office workers in a department store) were represented in each sample. The third set, the Frontroom sample, was of lower quality than the others; the writers were busy salesclerks rather than office personnel. This sample was obtained from forty writers and therefore contained more style variations.

Normalization was not done through directional control of the scanner in these experiments. Instead, the normalization was effected by a simulation program operating on orthogonally scanned versions of the characters. Whereas a scanner would have sampled the original pattern at coordinates along the computed skew-direction lines, the simulation program, instead of sending the computed sampling coordinates to a scanner, rounded off these coordinates to obtain the indices of a cell in the original video raster. The state of this cell was transferred to the appropriate cell in the normalized version of the character.

Thus the normalization amounted to a rearrangement of the bits of the initial raster. The effect of using this procedure, rather than the true rescan with a controlled flying spot, was to introduce a small amount of edge noise into the normalized pattern. If the same pattern was transformed five or six times successively (e.g., by rotating it) by simulation, a deterioration in video quality was noticeable. A single normalization of a given raster was not especially noisy.

One noteworthy feature of normalizing in this manner is that the scanner itself was eliminated from consideration in the comparison experiments. Each recognition program was run twice with the same binary patterns as inputs; in one experiment, however, the normalization routine was inserted as an intermediate stage. Thus variations in performance were a consequence of normalizing the pattern and could not be ascribed to alterations in scanner characteristics.

- *Stability of the pattern*

In Fig. 8 are shown the results of superimposing normalized and unnormalized pattern arrays. In general the normalized raster contains a greater proportion of

Table 1 Recognition experiments.

Recognition procedure	Design data	Test data	Number of errors: Unnormalized data	Number of errors: Normalized data
Templates	1000 Backroom 1	5000 Backroom 1	397[a], 669[b]	63
Clustering	1000 Backroom 1	1000 Backroom 1	97[c]	5[c]
Autocorrelation	1000 Backroom 1	3500 Backroom 1	39	16
Weighted cross-correlation	0's, 5's and 6's from Backroom 1	0's, 5's and 6's from Backroom 1	36	9
Zoned n-tuples 1	7900 Backroom 1	7900 Backroom 1	53	17
" " 2	10,000 Backroom 1	10,000 Backroom 2	100	52
" " 3	10,000 Backroom 1	10,000 Frontroom	580	307
" " 4	10,000 Frontroom	10,000 Frontroom	270	141

[a] Patterns registered at the centroid.
[b] Patterns registered with left-hand and lower boundaries at the margins of the field.
[c] Samples falling in a cluster in which a different identity predominates.

cells that are black with high frequency, implying greater stability than the raw patterns.

- *Correlation with templates*

An elementary method of designing templates is to superimpose a number of samples of the same identity and to quantize the pattern field into black, white and grey regions by applying upper and lower threshold bounds to the frequency count at each bit position.[2] The degree of agreement between a black-white input pattern and one of the ternary masks is measured by the Hamming distance (i.e., the number of mismatches) between the black and white regions of the mask and the corresponding bits of the pattern, with a constant added to account for the grey regions of the mask. A recognition scheme employing this principle was tried on several thousand normalized and unnormalized Backroom characters. The results, tabulated in Table 1, indicate improvement by a factor of six when normalized data are used.

This distance to a ternary mask is also a basic calculation in a clustering program[11] previously employed in several problem areas in pattern processing. The objective of this program is to arrange a given collection of input patterns into groups according to similarity. The program forms tentative groups, calculates a ternary mask for each group, and uses similarity to the masks to form new groupings. This procedure is reiterated until it converges.

The first 1000 normalized and unnormalized Backroom 1 patterns were respectively clustered into fifteen groups by this algorithm. With normalized inputs convergence was reached in four iterations; with the original patterns, twelve loops were required. As shown in Table 1 the similarity groups for normalized patterns contained primarily characters of a single identity. Only five samples were misclustered or fell among samples of different identity. The raw data, on the other hand, clustered poorly by this criterion.

- *Autocorrelation*

McLaughlin and Raviv[12] have recently described a scheme for implementing recognition that uses a high-order autocorrelation of the input. They show that in the decision function of interest the autocorrelation called for can be replaced by a function of the cross-correlation of the input with a template.

Decision experiments employing this technique were originally conducted with Backroom inputs. Some months later the same experiments were repeated with the normalized Backroom data. The error count was reduced from 39 to 16 by normalizing the patterns before correlating the data.

In addition, an illustration of the effect of normalization was given by this program.[11] The design program of McLaughlin and Raviv begins by finding ten patterns of greatest mutual variation among the first 100 samples of each category. In Figs. 9 and 10 are shown the selected samples for unnormalized and normalized inputs, respectively. The same 1000 samples were involved in each case although the patterns actually selected were different in the two trials. It is apparent from inspection of these two figures that much of the pattern variation within each category is removed by normalizing.

- *Weighted cross-correlation*

Chow[13] has experimented with a recognition system that correlates a pattern and a stored template in each of a number of shift positions and weights the correlations to form a score for the template. In one of the early test runs with this system all the 0's, 5's and 6's were selected from the first 1000 Backroom 1 numerals and used to design templates. (Storage limitations restricted the experiment to three classes.) These templates were then tested on the first 1000 samples from Backroom 2. When unnormalized characters were read, 36 errors were made; this number dropped to 9 errors when the design and test samples were normalized.

- *N-tuple recognition*

A detailed experimental investigation of hand printing recognition is described by Bakis, Herbst and Nagy.[7] The measurements used were local feature detectors—in effect, small templates—which were required to match a region of the pattern exactly in order to receive the value 1; otherwise the measurement value was 0. Each feature template was tried for all shift positions in a specified zone of the pattern field. In addition, several topological measurements, e.g., functions that counted the lines intersected when the pattern field was sliced in a specified manner, were added to the pool of measurements. Various subsets of this pool were tried in recognition experiments with the Backroom and Frontroom data. A Bayes decision procedure was used. Several of these experiments were repeated on the normalized patterns with results as given in Table 1.

In the first of these runs the 100 n-tuple measurements were identical for both the normalized and raw video samples. These measurements, called zoned n-tuples, consisted of six to eight input AND gates designed to detect lines, line ends and sharp bends at various orientations in the character. They can be considered "general purpose" measurements since they were not generated for optimum performance on a particular batch of data.

In n-tuple runs 2 through 4 of Table 1 the performance of these zoned n-tuples operating on normalized inputs is compared with that of a different set of 100 n-tuples corresponding to the unnormalized characters. The latter, which had yielded the best previous n-tuple performance on hand printing, had been selected from a large pool of such measurements (including the zoned n-tuples) on the basis of ability to classify the Backroom 1 samples. They yielded a fifty percent better error rate than the zoned n-tuple measurements when both were run on the same characters (see Ref. 5, p. 21). Since a rough idea of relative performance was available, it was decided to conduct the comparison on the basis of the two different measurement sets rather than to repeat the lengthy measurement selection procedure with normalized data.

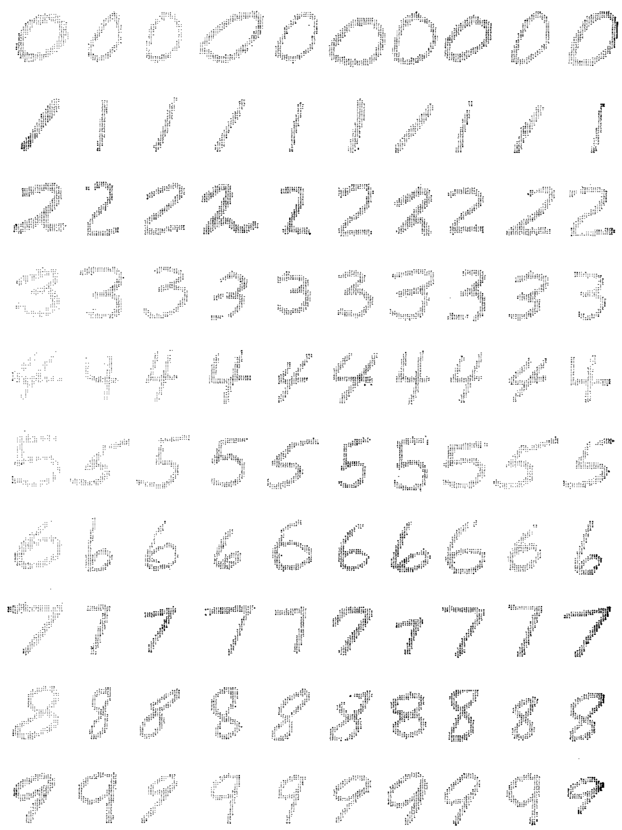
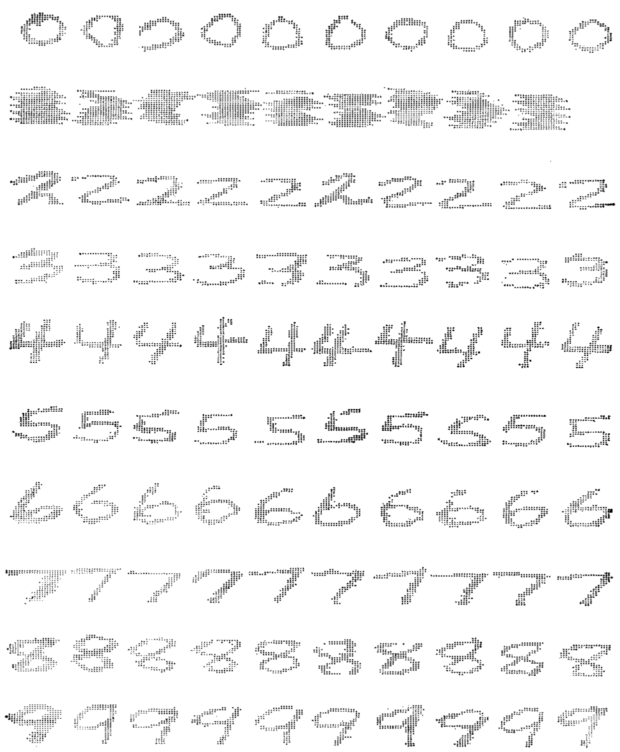


Figure 9 The 100 "most dissimilar" characters before normalization.

Figure 10 The 100 "most dissimilar" characters after normalization.



Presumably reselection would result in somewhat lower error figures in the right-hand error column of Table 1.

Conclusions

The experimental evidence gathered to date indicates that the moment normalization procedure is a valuable device for improving the recognition of handprinted characters.

The technique was originally developed with global template matching in mind, yet it has significantly improved feature detection systems as well.

Although because of different data collection procedures an exact comparison of results is not feasible, some of the recognition rates achieved with the Backroom and Frontroom normalized patterns are competitive with those attained by curve-follower techniques (see Ref. 6) on the same data. The best-performing methods (specifically the n-tuples) have previously been restricted in practice to the recognition of machine printing. Presumably, this implies that a recognition system can be built which applies the same hardware to the identification of both hand printing and machine printing. Previously, two distinct modes of operation have been implemented—curve-follower processing for handprinted symbols and raster-scan techniques for the machine-printed matter.

The fact that narrow characters are distorted in appearance by the normalizer does not appear to be a detriment. Indeed the normalizer provides size information which can be helpful in recognizing such patterns. It is interesting to note, however, that most of the recognition procedures tried had no difficulty in recognizing swollen 1's, even without width measurements from the normalizer.

The normalization approach is currently being applied to the problem of computer input of line drawings. In the process of encoding graphic information such as that contained in maps and engineering drawings, symbolic data of many different proportions and orientations are encountered. So far the recognition of these characters has proved a stumbling block to machine encoding of graphic material. After normalizing, however, these characters present a uniform appearance to the categorizer. Orientation information in this case is derived from observing the direction along which a sequence of characters occurs.

In this report only the recognition of numerals has been studied experimentally. To date, experimentation with alphabetic characters has been fragmentary, due to a lack of suitable data and the greater expense of operating with 26 classes instead of 10. However, the moment normalization technique would appear to be extensible to alphabetic input. Suppose, as is theoretically possible, that as a result of the normalization procedure two different classes become more similar, that is, more likely to be confused. This occurrence implies that scale, skew or location (the only parameters varied by the normalization routine) must be important recognition features for distinguishing between the two classes. Since these parameters are calculated, and are available in the machine, they are readily inserted into the decision process. In practice it would seem that appropriate use of the information given by the normalization step ought to guarantee against any degradation of performance.

In summary, the moment normalization procedure is conveniently implemented at the cost of two scans per character; it has been found to reduce by integral factors the error rates in recognizing hand printing; and quite possibly it enlarges the realm of recognition problems that can be handled by a single machine.

Acknowledgments

Particular thanks are due a number of colleagues for their participation in the experiments described in the text. These include C. K. Chow, N. M. Herbst, G. Nagy and J. Raviv. R. Bakis, through conversations and through his own work on pattern normalization, greatly influenced the approach reported here. P. M. Will assisted in outlining an implementation of the moment normalization procedure. Finally, it is a pleasure to acknowledge the programming assistance of G. H. Purdy.

Appendix 1: Proof of theorem

Let the moment matrices for P and P* be $\mathbf{M}$ and $\mathbf{M}^*$, respectively. Denote by $\mathbf{A}$ the transformation from P to P*. Let nonsingular matrix $\mathbf{T}$ diagonalize P, and matrix $\mathbf{W}$ diagonalize P*, to give the moment matrix $k\mathbf{I}$ in either case (k is a scalar, $\mathbf{I}$ is the identity matrix). Then the relations among the moment matrices are

$$\mathbf{TMT}' = k\mathbf{I} = \mathbf{WM}^*\mathbf{W}'$$

and

$$\mathbf{AMA}' = \mathbf{M}^*.$$

From the first relation we have

$$\mathbf{M} = k\mathbf{T}^{-1}(\mathbf{T}^{-1})'.$$

Also,

$$\begin{aligned}\mathbf{WM}^*\mathbf{W}' &= \mathbf{W}(\mathbf{AMA}')\mathbf{W}' \\ &= \mathbf{WA}[k\mathbf{T}^{-1}(\mathbf{T}^{-1})'](\mathbf{WA})' \\ &= k(\mathbf{WAT}^{-1})(\mathbf{WAT}^{-1})'.\end{aligned}$$

And thus

$$(\mathbf{WAT}^{-1})(\mathbf{WAT}^{-1})' = \mathbf{I}.$$

Therefore, $\mathbf{WAT}^{-1}$ is a combined reflection and rotation denoted by some matrix $\mathbf{D}$, and we see that

$$\mathbf{WA} = \mathbf{DT},$$

so that the normalized patterns can differ only by the rotation-reflection associated with **D**.

Appendix 2

We want to transform a video pattern represented by $f(x, y)$ into a pattern $g(x', y')$ by using the coordinate transformation

$$x = a_{11}x' + a_{12}y',$$
$$y = a_{22}y'.$$

The pattern $f(x, y)$ is assumed to vanish outside a finite region, and a_{11} and a_{22} are required to be positive numbers (to prevent inversion of coordinates). We also require that the parameters a_{11}, a_{12} and a_{22} be such that the moment matrix of $g(x', y')$ has the canonical form

$$\mathbf{M}^* = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}.$$

Using the relation

$$\mathbf{M} = \mathbf{A}\mathbf{M}^*\mathbf{A}',$$

where

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix},$$

we find the new moment matrix to be

$$\mathbf{M} = \begin{pmatrix} k(a_{11}^2 + a_{12}^2) & ka_{12}a_{22} \\ ka_{12}a_{22} & ka_{22}^2 \end{pmatrix}.$$

Thus we obtain three equations in a_{11}, a_{12} and a_{22}:

$$k(a_{11}^2 + a_{12}^2) = m_x$$
$$ka_{22}^2 = m_y$$

and

$$ka_{12}a_{22} = m_{xy}.$$

The unique simultaneous solution of these equations, subject to the conditions imposed, is

$$a_{11} = [(m_x m_y - m_{xy}^2)/km_y]^{1/2},$$
$$a_{22} = (m_y/k)^{1/2}$$

and

$$a_{12} = (m_{xy}^2/km_y)^{1/2}.$$

References

1. W. H. Highleyman, "An Analog Method for Character Recognition," *IRE Trans. Electronic Computers* **EC-10,** 502 (1961).
2. M. E. Stevens, "Automatic Character Recognition—A State-of-the-Art Report," *NBS Tech. Note 112,* May 1961.
3. It is also possible to apply the technique presented here to transform the recognition logic to agree better with the pattern to be recognized. In certain cases, e.g., decision tree logic based on the binary pattern matrix, this may be the more efficient way to proceed (C. Naylor, private communication).
4. K. Udagawa, J. Toriwaki and K. Sugino, "Normalization and Recognition of Two-Dimensional Patterns with Linear Distortion by Moments," *Electronics and Communications in Japan* **47,** 34 (1964).
5. M. K. Hu, "Pattern Recognition by Moment Invariants," *Proc. I.R.E.* **49,** 1428 (1961).
6. F. Alt, "Digital Pattern Recognition by Moments," *Optical Character Recognition,* edited by G. L. Fischer et al., Spartan Books, Inc., Washington, D. C. 1962, p. 153.
7. R. Bakis, N. Herbst and G. Nagy, "An Experimental Study of Machine Recognition of Handprinted Numerals," *IEEE Trans. System Science and Cybernetics* **SSC-4,** 119 (1968).
8. P. Essinger and E. C. Greanias, "Normalizing Multilevel Quantizer," U. S. Patent No. 3,248,699 (April 26, 1966).
9. C. N. Liu and G. L. Shelton, "An Experimental Investigation of a Mixed Font Print Recognition System," *IEEE Trans. Electronic Computers* **EC-15,** 916 (1966).
10. E. C. Greanias, P. F. Meagher, R. J. Norman and P. Essinger, "The Recognition of Handwritten Numerals by Contour Analysis," *IBM J. Res. Develop.* **7,** 14 (1963).
11. R. Casey and G. Nagy, "Recognition of Printed Chinese Characters," *IEEE Trans. Electronic Computers* **EC-15,** 91 (1966).
12. J. L. McLaughlin and J. Raviv, "Nth Order Autocorrelations in Pattern Recognition," *Information and Control* **12,** 121 (1968).
13. C. K. Chow, publication pending.

Received April 3, 1968; revised January 29, 1970