

E. C. Greanias
P. F. Meagher
R. J. Norman
P. Essinger

The Recognition of Handwritten Numerals by Contour Analysis

Abstract: A character recognition system has been developed for the recognition of handwritten numerals. This system uses a logically controlled cathode ray tube scanner to generate basic measurements that characterize significant features of the numeral shapes. A contour-follower procedure is used to control the scanner. In addition, special scanner subroutines initiated by feedback from the recognition logic are utilized. Character shape data are generated in a sequential form, which can be analyzed for recognition with an easily realizable logic.

An experimental model has been built that recognized 99.3% of numerals written by 45 subjects after 30 minutes of training. The error rate for these characters was 0.11%. The rejected character rate was 0.59%.

Introduction

Over the past ten years, optical character recognition has developed as a method for entering printed characters into information processing systems. This paper describes a practical method for the recognition of numerals that have been recorded by hand with an ordinary pen or pencil.

Several authors^{1,2} have noted that the general pattern recognition system can be described in two principal parts:

1. A first part (sometimes called the *receptor*) that produces a set of quantities which characterize the pattern in terms of a prescribed set of measurements. These measurements can be continuous or discrete, or both.
2. A second part (sometimes called the *categorizer*) that applies a decision criterion to the pattern measurements for a decision as to which, if any, of the allowable classes the pattern belongs.

The effectiveness of a recognition system will depend upon how well the significant differences between pattern classes are characterized by the prescribed set of measurements,³ and the accuracy with which

the decision criterion categorizes the pattern measurements. Pattern recognition literature contains a number of papers that discuss optimum decision functions⁴ and practical methods for the selection of the decision criteria. The recognition system that is described in this paper employs a simple, direct method for generating significant measurements that characterize the fundamental shape differences among handwritten numerals. The categorizer that has been adopted employs detailed logical statements of the acceptable combinations of these measurements for each character to decide the identity of the unknown pattern. A weighted sum of probabilities decision function might have been used, but "detailed logical statements" were chosen instead to simplify the description of the handwritten numeral shapes that are acceptable for recognition. Furthermore, the implementation of the detailed logical statements for handwritten numerals was sufficiently economical to convince the authors that cost optimization of the categorizer is a secondary consideration for this system.

To provide a practical solution for the handwritten numeral recognition problem, it was required that preparing documents for the machine reader should

not constrain the writer so severely as to inhibit any other functions that he had to perform at the same time. To accomplish this, considerable tolerance was provided for:

- *Character registration.* If necessary, characters can be located anywhere within an area several inches square.
- *Character size.* The height of characters can vary over a 4-to-1 range.
- *Shape.* The tolerance for shape is defined in terms of the shapes that occur in unconstrained handwritten numerals. The initial recognition criteria were selected to cope with more than 90% of the character shapes found in 3000 samples of unconstrained numbers.⁵
- *Line quality.* Lines from medium hard pencils and dark-ink ball point pens in reasonably good condition are acceptable.
- *Character rotation or slant.* Normal characters can be rotated $\pm 20^\circ$.

This paper will outline the basic operation of the recognition system that was developed to solve this problem and some of the procedures that were used to define the recognition criterion.

Basic operation

A block diagram of the basic operation of this pattern recognition system is shown in Fig. 1. The characters are scanned on the stationary document by a flying-spot scanner. Signals from the scanner corresponding to the document reflectivity are usually the direct output of the receptor. In this system, they are not used directly, but feed a beam-control logic which in turn generates the signals that are analyzed. The beam-control logic is also controlled by feedback from the categorizer.

The characters are scanned optically along a variable path that depends upon the information obtained as scanning proceeds. The actual character on the document serves as the primary store of raw character information. The scanner is the means for retrieving the information as required from this "memory". Specific regions of the character are scanned selectively (and repetitively, when necessary) according to their significance in character identification. The scanner is also made to seek significant features in new areas of the character when prior scanning does not provide conclusive results.

Scanning starts with a search pattern that "looks" for a character on the document. This search pattern may take any form that is compatible with the document but would typically be a raster scan. When a character is detected by means of a "black" photo signal, the search pattern is stopped and signals are generated in the beam-control logic to cause the beam to follow the edge of the character with an exploratory circular motion, as shown by the photograph in Fig. 2. The signals that are passed on to the categorizer for

analysis are essentially the beam deflection signals, with the circular pattern filtered out to provide a smoothed outline of the character⁶ as illustrated by the photograph in Fig. 3. After recognition has been accomplished, the search operation is resumed in the region beyond the character.

Machine-reading capabilities

As a consequence of this mode of operation, several important machine-reading capabilities can be achieved in a straightforward and economical way.

• Independence of character size and location

To accomplish recognition, the contours of each character are traced two or more times. During the first tracing, the extreme positive and negative excursions of the character's outer contour in x and y are stored. On subsequent passes around the character, the beam is located relative to the character extremities and therefore is not dependent on the character location or the point of initial contact. The stored peaks of the character contours provide a measure of character size

Figure 1 Block diagram of character recognition system utilizing contour analysis.

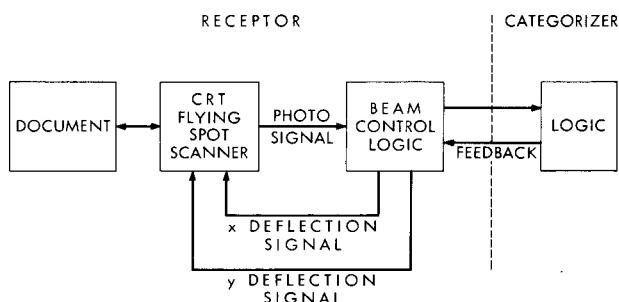
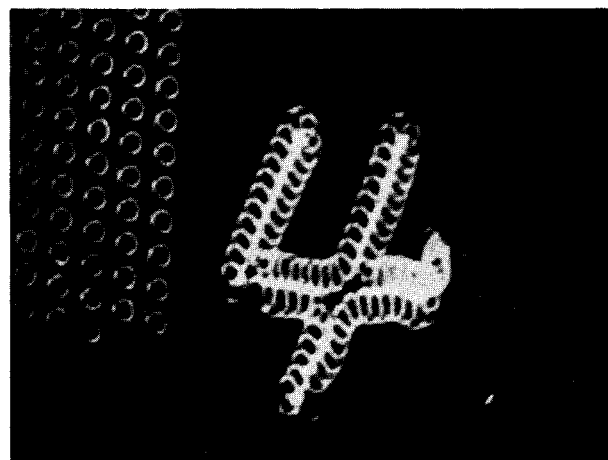


Figure 2 Photograph of deflection signals, showing search and follow modes.



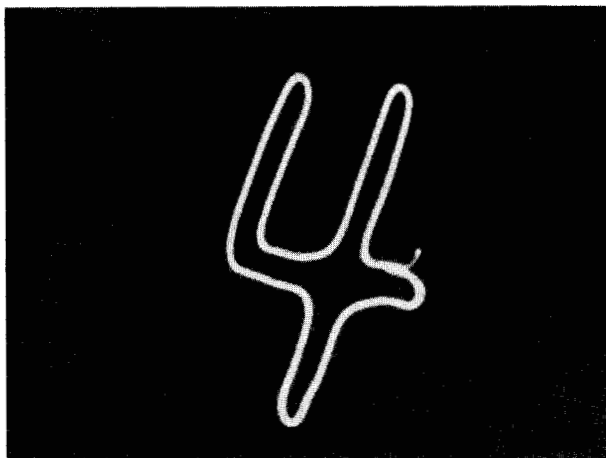


Figure 3 Photograph of filtered deflection signals.

which is used to adjust (i.e., to normalize) the vertical and horizontal distance measures of the system. The character size measurement is also used to adjust the scan resolution (circle size) on subsequent passes around the character. This ensures that all important shape features will be detected, and that the time to follow differently sized characters of the same shape will be approximately the same. Figure 4 contains photographs of actual scanning traces that occurred during the first and second pass around the character **m**. The improved resolution on the second pass makes the shape of the **m** apparent.⁷

• *Tolerance for line quality*

It is evident that successful recognition of a character in this system depends upon how well the line edge is followed. For this reason it is worth while to employ some fairly sophisticated methods to determine the location of lines having marginal quality. The follower method used causes the beam to move in and out of the line periodically. By using the recent histories of the "black" and "white" signals, the "local" black-white threshold that is maintained to discriminate the reflected light signals can be adjusted to automatically accommodate large variations in the character line and paper background reflectivities. Furthermore, the occurrence of "black" signals can be predicted on the basis of the direction of the line edge that is being followed. When a discontinuity occurs, a special scanning subroutine is used to check the validity of signals that do not occur as predicted. For example, in the case of weak lines, the failure of "black" to occur when predicted initiates a routine which causes the beam to explore an adjacent region of the line with a modified black-white threshold.

• *Tolerance for shape*

The analysis of the filtered beam deflection signals is based on two properties of the line edges that they represent:

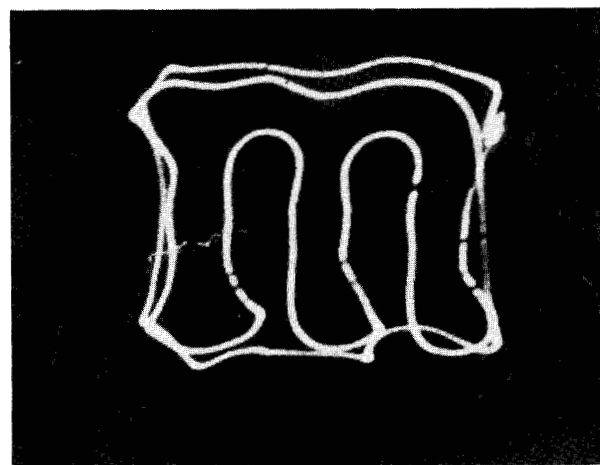
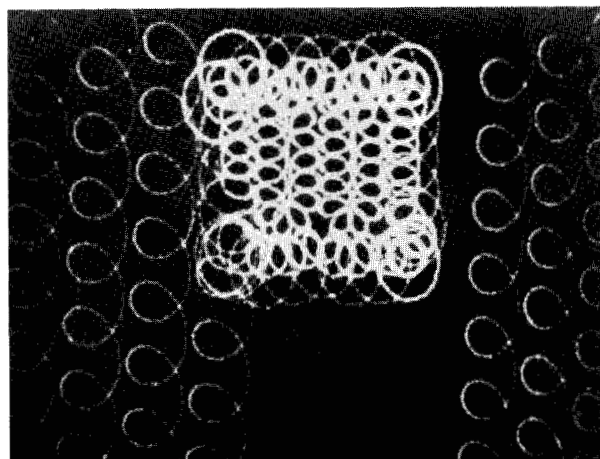
- a) The relative location of the line edge within the character.
- b) The approximate direction of the line edge.

By considering both properties simultaneously, the significance of most line edge signals can be detected without resorting to precise measurements of location and direction. In some instances, such as certain 8's and 0's, identity cannot be established by coarse measurements on the outer contour of the character. For these cases, the scanner is redirected to obtain more information from the interior of the character, which also can be analyzed with simple, coarse measurements, to resolve the conflict. Since accurate recognition can be accomplished without resorting to precise measurements of line locations or directions, this system provides large tolerance for shape variations.

The experimental reader

The experimental model of this recognition system

Figure 4 Photograph illustrating the effects of circle size adjustments following the character "m".



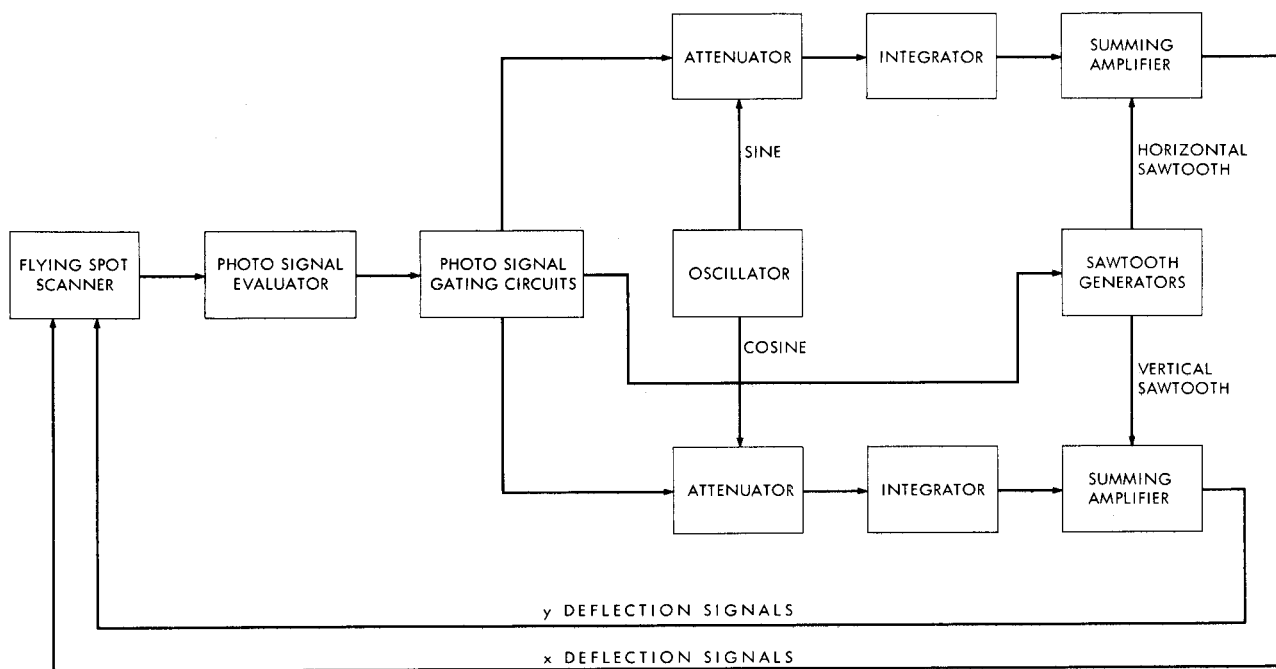


Figure 5 Block diagram of beam control logic.

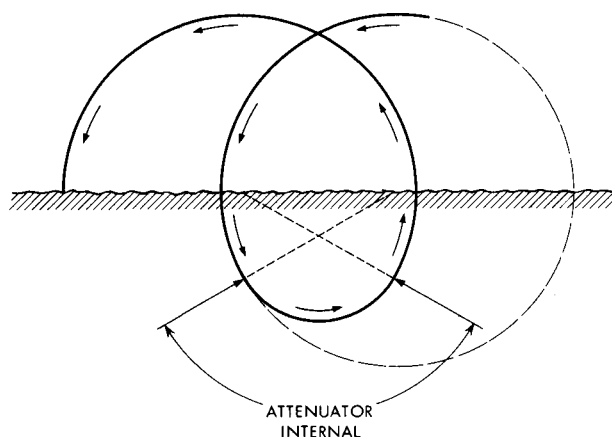


Figure 6 Beam motion during line following.

that has been constructed and tested employs a cathode ray tube flying-spot scanner designed to read numerals in a $2'' \times 3''$ area on an IBM card. The scanning procedure includes the search mode, which is used to find the characters, the follow mode, which is used to explore the outer contour of the characters, and an internal scan mode, which is used to explore the inner contours of selected characters. The basic principle of the beam-control logic is illustrated schematically in Fig. 5. The x and y deflection signals are generated by summing the exploratory circle component and sawtooth sweep signal for each axis. When a character is detected and the search mode is terminated, the outputs

of the sawtooth generators are frozen. Circular deflection components are still fed to the "summers," but to achieve edge follower action the radius of the circular motion is attenuated for a fixed interval after the beam has entered the line. This attenuation interval is timed to occur when the beam is moving essentially parallel to the line edge but opposite to the direction of advance along the edge. Follower action is accomplished by integrating the x and y circle components that provide the circular motion. Because of the action of the integrators, the beam cannot move abruptly when the radius is suddenly attenuated. Instead, at the instant of attenuation the beam does not move, but the center of the circle is shifted in the direction of advance. When the attenuation interval is finished, the center of the circle is also shifted in the direction of advance. This in effect generates a cycloidal motion (see Fig. 6). The duration and magnitude of the attenuation determines the advance per circle.

To generate the waveforms for analysis, the outputs of the integrators are filtered by a low-pass network that cuts off at approximately one-third of the exploratory circle frequency. On the first pass around the character, the filtered signals are fed into capacitor peak-storage circuits. The outputs of the positive and negative peak storage circuits for each deflection axis are connected through a resistor divider network. In this model the vertical deflection range is divided into four equal increments and the horizontal deflection range into three equal increments by comparing the filtered deflection signals on subsequent passes with voltage thresholds defined by the resistor divider

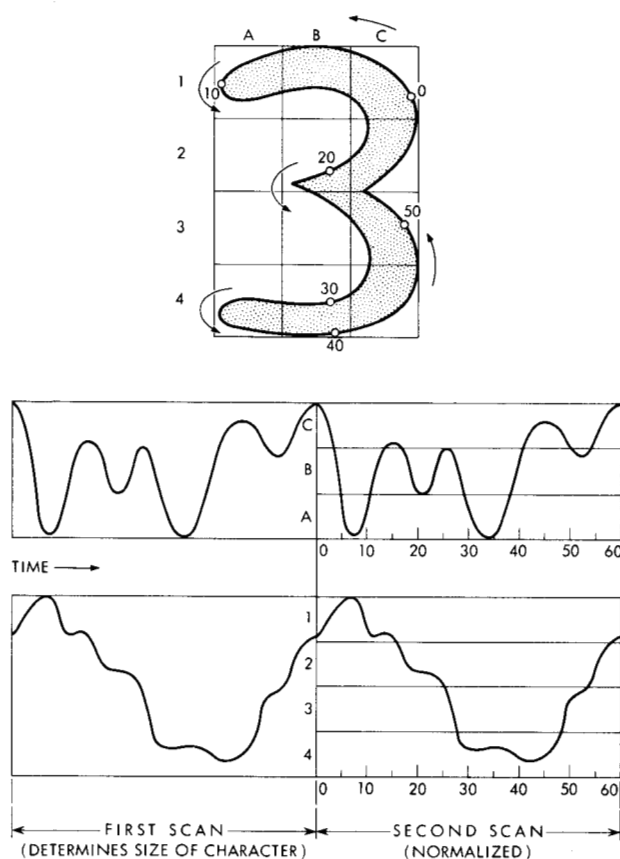


Figure 7 Normalized scan matrix.

networks. These thresholds define 12 possible zones in a 3×4 array that the deflection beam can enter as it travels around the character. Signals that identify the location of the beam within the 12 zones are continuously generated in the model. Figure 7 illustrates this operation.

Line directions are determined in this experimental model by first generating the time derivative of the filtered horizontal and vertical deflection signals. A comparison is then made of the instantaneous values of these derivatives to determine in which of eight 45° sectors (i.e., N, NE, E, SE, S, SW, W, NW) the beam is directed. This is accomplished by implementing the following relationships:

$$A \quad dx/dt < (\tan 22.5^\circ) dy/dt$$

$$B \quad -dx/dt > (\tan 22.5^\circ) dy/dt$$

$$C \quad dy/dt < -(\tan 22.5^\circ) dx/dt$$

$$D \quad -dy/dt > -(\tan 22.5^\circ) dx/dt$$

The conditions for the instantaneous line directions are:

$$N = A \cdot \bar{B}$$

$$NE = B \cdot \bar{C}$$

$$E = C \cdot \bar{D}$$

$$SE = D \cdot \bar{A}$$

$$S = \bar{A} \cdot B$$

$$SW = \bar{B} \cdot C$$

$$W = \bar{C} \cdot D$$

$$NW = \bar{D} \cdot A$$

When the scanning beam passes the point of original contact with the character, a signal is generated by a circuit which monitors the condition of the follower integrators and the direction of beam motion. This pulse signifies the start of the second pass around the character and turns on all of the subsequent recognition operations.

Recognition is based upon selected sequences of line directions and locations that are detected after the first pass around the character.¹⁰ To provide adequate shape tolerance, the filtered beam deflection signals from 3000 uncontrolled characters were recorded and studied by simulation on an IBM 7090 computer. The line direction and location signals of each character were sampled 60 times around the periphery to produce 120 code digits (one digit for direction and one digit for position, per sample) as the input for the simulation program.

First, significant features were identified for each character type. Then shape feature tests were defined, and recognition logics were assembled that satisfied 2752 of the characters in the sample. These recognition criteria were selected in accordance with the logic designers' subjective judgments regarding "normal" character shapes. The shape features that were emphasized for each character are summarized in Table 1.

Figure 8 illustrates a typical recognition criterion for the numeral 5. The features that are required are identified by the small arrows adjacent to the character outline. For example, stroke labeled I is defined by the occurrence of an east (E) beam motion anywhere in the second (2) row of the array or east (E) or southeast (SE) motion in the third (3) row of the array. When this simple condition is satisfied for an adequate short interval, a latch is set to store the event and enable the detection of the second required stroke (II) by a corresponding logical circuit. Similar implementation is provided for strokes (III), (IV), (V) and (VI). The output that is obtained when stroke (III) is satisfied is used in the recognition statement of other characters and is defined as a shape feature test. Strokes (V) and (VI) call for extended occurrence of the indicated horizontal motion so longer durations of the signals are required to set the corresponding latches.

The characters 2, 3, 4 and 7 are also recognized by similar recognition logics but, when the possible occurrence of 0, 1, 6, 8, or 9 is detected, further scanning of the internal regions of the numeral is required to insure that errors do not occur.^{12,13} To accomplish this, the scanner is forced to break through the line and

Table 1 Shape features identified for each character.

Significant Feature	Location	Use*
long N, NE or NW strokes	right side	0 1 7
long S, SW or SE strokes	left side	0 1 7
horizontal line end	top left	2 3 7
horizontal line end	top right	5 6
horizontal line end	bottom left	3 5
horizontal line end	bottom right	2
loop	bottom	2
west bay	bottom left	3 5
north bay	top center	4
northeast bay	top right	9 8
short arc	top left	9
NE stroke	top left	4
horizontal line end	right center	4
east bay	top right	6 5
notches	left	8
notches	right	8
southwest bay	bottom left	4 9
short arc	bottom right	6 5
large lake	center	0
small lake	top center	8 9
small lake	bottom center	8 6

* Note: Many of these features are also used as inhibit conditions in other characters.

follow the internal contour. See Fig. 9. The information derived from the inner contour scan is used to define shape feature tests similar to those derived from the outer contours and final recognition takes place when appropriate inside shape features are detected.

Results

The recognition logics for this model were implemented with conventional circuits using approximately 800 transistors. An additional 200 transistors were required for the circuits that generated the line direction and location signals. The cathode ray tube deflection circuits and beam-control logic required approximately 400 transistors. The reading capabilities of this model were determined in an experimental study at Tufts University and are described in the Letter to the Editor by M. N. Crook and D. S. Kellogg.¹⁴ Approximately 200 untrained subjects wrote 20,000 numerals, of which 92% were correctly recognized. This is comparable to the results of the original simulation. After 45 writers were given approximately 30 minutes of initial training to write "normal" characters, they wrote 23,500 numerals, of which 99.3% were recognized correctly. The error rate for these characters was 0.11%, and the reject rate was 0.59%.

The performance of this experimental reader cannot be described adequately in terms of error statistics

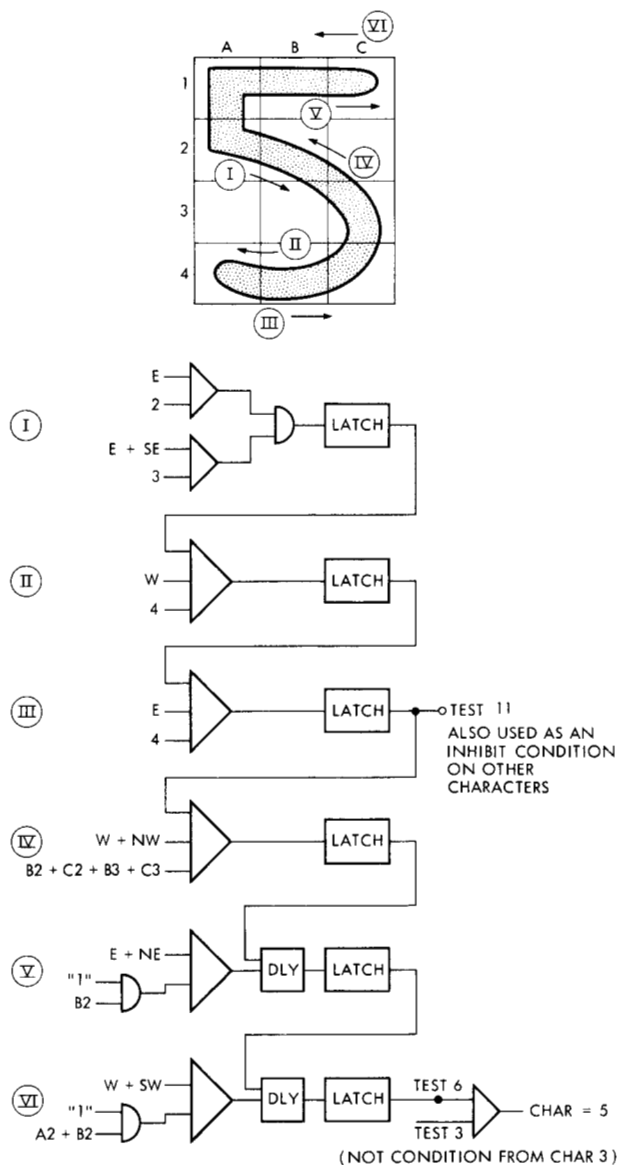


Figure 8 Recognition criterion for the numeral 5.

alone. The variations in character shape and line quality with which the reader can cope are not apparent in these figures. Unfortunately, there are no suitable quantitative measures of character shape and quality for use in a precise description of the machine reader's tolerances. However, an indication can be obtained from a description of the most frequent causes of failure to read. These are:

- *Linked or touching characters.* When adjacent numerals touch or are linked together, at least one will be rejected.
- *Broken characters.* Large gaps ($> 0.015''$, beyond the limits of the subroutine that is provided to cope with

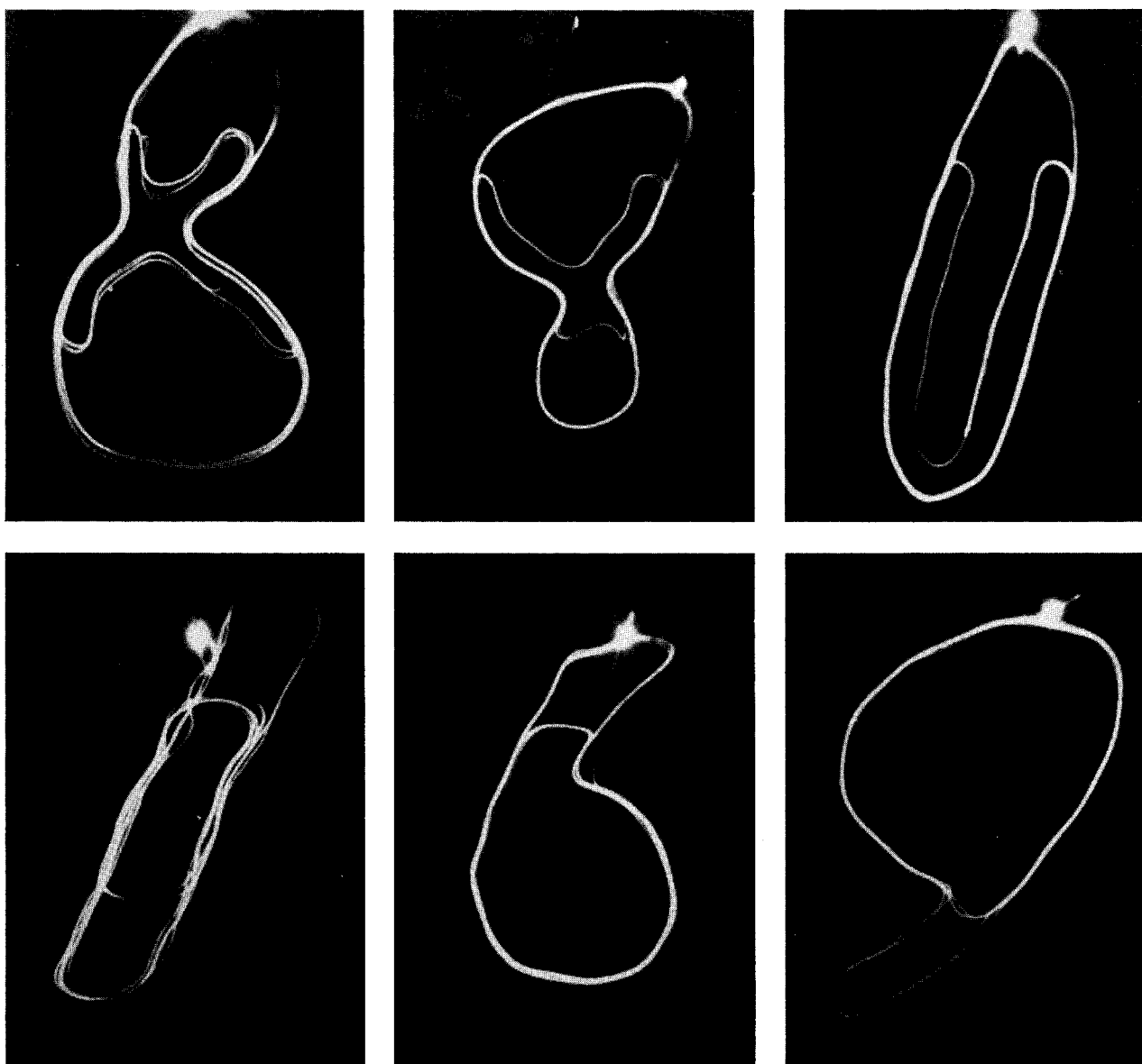


Figure 9 Cathode ray tube display showing two internal contours of the numeral 8 and one contour each for numerals 0, 1, 6 and 9.

discontinuities) due to lines that fail to join or to weak impressions can cause rejects or errors. This occurred most frequently in 5's with horizontal top bars that did not join the rest of the character.

- *Characters with extraneous line segments.* These lines occurred accidentally or as part of flourishes. Some examples were 7's with extraneous horizontal lines at the bottom that could cause ambiguities between 7's and 2's, and 1's with extraneous lines at the top that could cause ambiguities between 7's and 1's.

- *Characters with exaggerated proportions.* This can occur in any of the numerals. For example, extremely

long stems on 9's or 7's would make the top appear insignificant and cause possible ambiguities with the 1.

- *Characters similar to other character shapes.* The "closed top" type of 4 was frequently missed when the horizontal bar failed to cross the vertical stem because the resultant shape was very similar to the 9. (When the subjects were trained to write for this reader, they were encouraged to use the open-top 4.)

Conclusion

It should be noted that the recognition system described here does not employ a single omnipotent technique to read handwritten numerals. The indicated

tolerances for character shape, size and quality of this reader were obtained by providing specific capabilities to handle specific problems. Although some features of this reader are similar to ideas published elsewhere^{8, 9, 15, 16} the authors believe that by providing the combination of functions that have been described, a new high in pattern recognition flexibility has been achieved. Some noteworthy features of this reader are:

- a) The special scanning subroutines that cause the beam to re-explore poor lines with a modified black-white threshold to accommodate weak regions of the line.
- b) The special scanning subroutines that cause the beam to seek specific shape features in selected regions

of the character for the resolution of ambiguities that still exist after normal scanning.

- c) The simultaneous use of line-direction and position-relative-to-extremities information for recognition.

The development of more powerful readers might be achieved by improving the performance of these functions or adding new functions to the reader.

Acknowledgment

The authors wish to acknowledge the contributions of D. Alexander and P. Hurley in the design and construction of the experimental model, and of H. Penafiel in the programming and design by simulation of the recognition logic.

References and footnotes

1. D. M. Green and T. Marill, *IRE Trans. on Electronic Computers*, EC-9, 472-477 (1960).
2. W. H. Highleyman, *Proc. IRE*, 50, 1501-1514 (1962).
3. E. C. Greanias and Y. M. Hill, *Proc. IRE*, 7, 119-126 (1957).
4. C. K. Chow, *IRE Trans. on Electronic Computers*, EC-6, 247-254 (1957).
5. Subjective judgment was used in the selection of recognition criteria. In general, the numerals that were similar to most of the other samples in this population were considered "normal", and the recognition criteria were designed to accept them. The tolerance that was provided to accept all "normal" characters made it possible to recognize many numerals with shape combinations that did not occur in the "normal" group.
6. Recognition based on the analysis of analog waveforms that correspond to the shape of the character was also accomplished in the early work of W. Sprick (Ref. 8). Other references and suggestions regarding curve following recognition techniques are summarized in M. E. Stevens' N.B.S. report (Ref. 9).
7. Although this paper covers handwritten numerals, a lower case printed m is shown to provide a more effective illustration of the value of circle size control.
8. W. Sprick. U.S. Patents 2,738,499 and 2,838,602.
9. M. E. Stevens, National Bureau of Standards Technical Note 112, May 1961, pp. 53-55.
10. A similar sequential recognition procedure was described by Greanias et al. (Ref. 11).
11. E. C. Greanias, C. J. Hoppel, M. Kloomok, and J. S. Osborne, *IBM Journal*, 1, 8-18 (1957).
12. The recognition of the closed-top 4 was based on the occurrence of a pointed top or an adequate crossing of vertical stem by the horizontal line.
13. Internal scanning is tried on the 1 because with a thick line it could be mistaken for a 0. When internal scanning is tried on the 1, the beam always finds an "outer" contour.
14. M. N. Crook and D. S. Kellogg, *IBM Journal*, this issue, p. 76.
15. *Communications Theory*, edited by Willis Jackson, Butterworth's Scientific Publications, 1953, pp. 323-325.
16. M. Nadler, "An Analog-Digital Character Recognition System," presented at *IFIP Congress 62*, Munich, September 1, 1962.

Received September 19, 1962