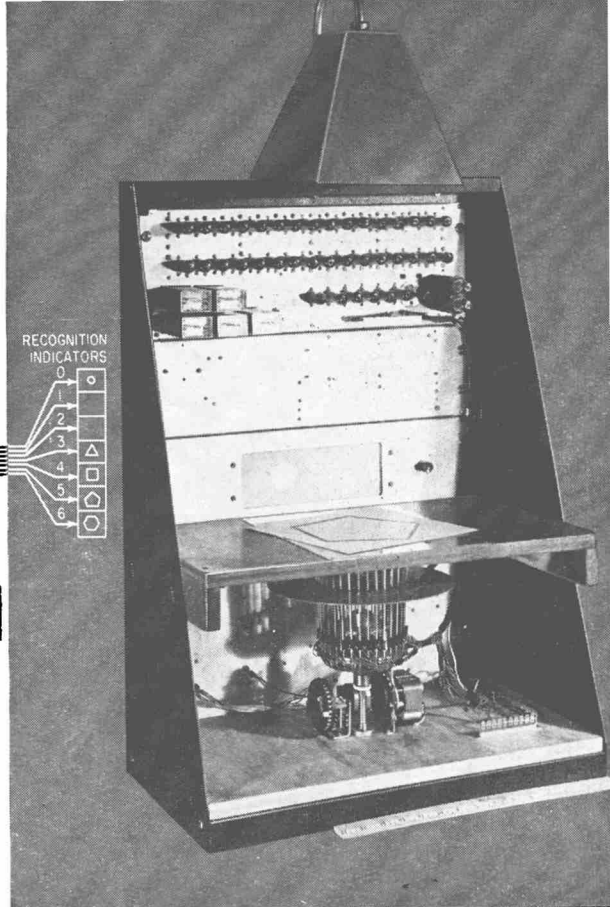


FIG. 1—Generalized block diagram of an n -gon recognition system which can make use of the dilating circular scan. At right, pattern-recognition machine recognizes a pentagon line drawing and displays it on screen



Line-Drawing Pattern Recognizer

Use of a dilating circular scan resolves some of the problems encountered in the development of a non-specialized reading machine. Technique is applicable to automatic detection of letters and numbers over a variety of styles

By LEON D. HARMON
Bell Telephone Laboratories, Inc.,
Murray Hill, N. J.

MACHINE DETECTION of two-dimensional visual patterns has been generally restricted to carefully constructed and positioned alphanumeric characters. Although a number of systems and proposals for obtaining limited recognition exist¹⁻⁴, and although considerable effort is currently being put into finding more general and flexible systems⁵⁻¹¹, many difficult problems remain to be solved.

Ideally, pattern recognition should be independent of character size, rotation, position, style and

noise. A really useful reading machine, for instance, will be one that can recognize the symbol 5 whether it is continuous or in two parts, and whether it has been printed by a machine or scrawled by a child.

Such general or Gestalt recognition is approached by the machine that is described here. It recognizes line drawings of circles, triangles, squares, pentagons and hexagons independent of their rotation and, within limits, of their size, precision of drawing, or positioning. In addition, it distinguishes and counts separate objects up to six with limited independence of size, shape and object position. Furthermore, the technique is ap-

plicable to automatic detection of letters and numbers over a wide variety of styles.

This Gestalt recognition is achieved by the use of a dilating circular scan. This scanning method yields similar transformations for geometrically similar figures — object-size changes translate into time-of-arrival changes while object rotation preserves topological relationships.

Consider an array of picture elements arranged in c concentric rings, having r elements in each ring. If each ring is inspected sequentially, say from center outward, the resultant c sequences of r signals can be shown to have the

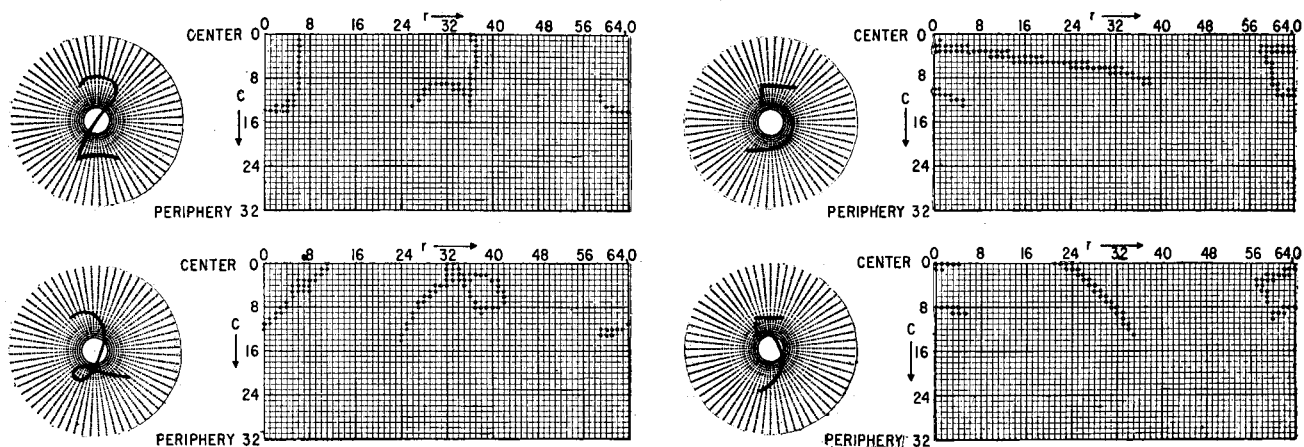


FIG. 2—Numerical examples. Here, size changes appear as vertical (time) shifts while rotation causes lateral displacements. Many figures have unique invariants when interrogated with this circular scanning

desired transformation properties. A logically equivalent implicit scan is possible with unidirectional propagation along radii and with unit delays inserted between adjacent elements on each radius.

Suppose that an equilateral triangle is centered on such an array. The first intersections of the scanning interrogation ring with the figure will be three simultaneous hits on the mid-points of the triangle's sides. As the scan continues its outward sweep, points on each side of the initial intersections are progressively touched. The last parts of the triangle which are inspected are the vertices, those points at the greatest radial distances from the center.

Thus the outward puckering scan produces three radially equidistant epicenters of activity. An accumulating register of r places, representing the r elements of each ring in turn, will display three single active cells initially, equally spaced along the length of the register. There will then be a spread of activated cells in each direction from each of the initial locations. Finally the apices of the equilateral triangle will cause the last three cells of the register to be activated. Again these are equally spaced, representing the 120-deg radial symmetry of the vertices.

A similar but smaller triangle would have generated similar signals earlier in the sweep cycle; conversely for a larger presentation. Rotational changes are represented as lateral shifts in the contents of the register. Thus the detection of three simultaneous, equidistant, and uniformly expanding waves of ac-

tivity is sufficient to recognize an equilateral triangle independent of size or rotation. (Bounds on size exist, of course. The lower limit is fixed by the resolution of the array of elements while the upper bound is determined by the field seen by the array.)

An isosceles triangle will generate easily detectable time and position asymmetries. Squares and rectangles are differentiable in a similar manner; their unique common property in this system is, of course, a count of four. A convex polygon of any number of sides can be detected, given sufficient resolution; circles are represented by a simultaneous filling of all places in the inspection register.

Let us consider now how a system can be developed to make use of such scanning. Suppose for example that the input transducer consists of 32 concentric rings of 64 photocells each. Let a sweep cycle comprise the simultaneous gating of all 64 cells in each ring, one ring at a time, into 64 corresponding amplifier channels. Thus for an outward puckering scan, the innermost ring is first interrogated, then the next, and so on, until the gating-out of the contents of the outermost ring completes a sweep cycle. Figure 1 shows one system for using this information.

The beginning of each sweep initiates a storage operation in each of the 64 function generator locations. A linear accumulation (e.g., current) continues in each location until stopped by the first black element encountered by the expanding sweep.

The value of each function at

the end of sweep may be read out as a voltage level which is proportional to the radial distance of the corresponding image point from the center of the array.

At the end of a sweep, a comparator makes peak detections on the 64 generated functions. This operation is performed by local differencing over several adjacent locations. A count of such peaks is sufficient to identify an n -gon, while a measurement of inter-peak distances (given by the addresses of the radii on which the peaks occur) establishes criteria for separation of isosceles and equilateral or of square and rectangle. The alphabet of line drawings recognizable by this system is not restricted to a few polygons.

Many figures have unique invariants if they are interrogated with an expanding circular scan whose origin is near the center of gravity of the figure.

A modest number of more complex comparator operations produces much more sophisticated recognition. Detection of a few types of specially defined groups of signals distributed in time over the 64 channels leads to identification of numerals, for example. A series of experiments on many examples of hand-printed digits has shown that very modest extensions of logic considerably enlarge the classes of recognizable images. Figure 2 shows several examples of numerals which have been recognized in these experiments.

Each of the four numerals was drawn freehand and roughly centered on a layout of the photo-receptors. The rectangular grids

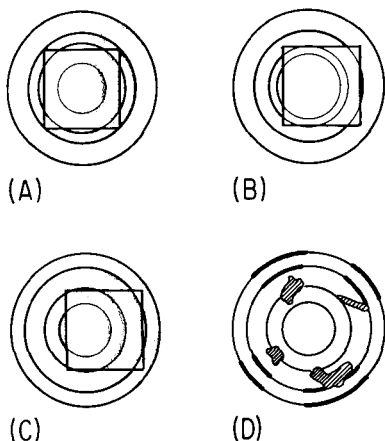


FIG. 3—How the sweep encounters patterns (A to C) and objects (D)

are obtained by unfolding the circular array after making a cut from the center to the top of that array. Thus points along the top of the rectangle represent points on the inner circle of the photocell matrix while the bottom line contains the points forming the periphery of the circular array. The left and right vertical edges of the rectangle were coincident in the original layout, being the two edges of the cut.

This simple transformation makes it considerably easier to follow the scanning events and to discover useful detection criteria. The expanding sweep circle is seen to be represented by a horizontal line 64 cells wide moving down in time through 32 positions. Any given horizontal line may be thought of as the contents of a 64 place register at any epoch. Each activated

receptor location on the circular array is shown as a dot in the rectangular representation. Size changes appear as vertical (time) shifts while rotational transformations cause lateral displacements of the pattern. Since the left and right edges of the rectangle are actually contiguous in the original array, lateral displacements simply wrap around the rectangular grid, disappearing off one vertical edge and reappearing at the other.

A set of tests for the examples shown in Fig. 2 will now be given. These tests are for illustrating principles and should not be considered the most complete or effective rules which can be found. We shall invoke two simplifications: the numerals will be centered and there will be no rotation.

An ascending function will mean a mostly continuous, mostly monotonic decrease in values of c as r increases. The first 2 shown would be said to have an ascending function from $r = 0$ to $r = 6$; conversely the first 5 has two descending functions between $r = 0$ and $r = 6$. In a similar manner we may describe horizontal segments, concave or convex functions, etc. We test for the following: (1) An ascending function for $0 \leq r \leq 8$. This describes the curved section of the upper right portion of most 2's. (2) An ascending function for $24 \leq r \leq 32$. This describes that portion of the tail of a 2 lying to the right of center. (3) A partially horizontal or ascending function for

$32 \leq r \leq 40$. The presence of at least a short segment which is level or rising during this interval indicates the left end of the tail or the bottom of the down-stroke for most 2's and is independent of whether or not a loop is present. (4) A two-valued function for $0 \leq r \leq 8$ and $60 \leq r \leq 64$ establishes the presence of the top bar and top of the curved portion of a 5. (5) A descending function for $16 \leq r \leq 40$ describes the tail of a typical 5.

The limits of r need not be taken as absolute. If our tests are based on relative r values, then the results of the tests can be independent of rotation of the sample numerals.

Rules based on the five tests listed above are sufficient to obtain unique recognition of the samples given and a variety of other examples of these two numerals. The first three tests concern properties of the number 2 which are relatively invariant over a population of about twenty samples which have been checked. Similar rules have been successfully applied to all of the ten digits for a number of handprinted samples. The extension to letters, while more complicated, does not, in principle, appear difficult.

To demonstrate principles, a simplified model has been built to recognize line drawings of circles, triangles, squares, pentagons, and hexagons. The device detects and counts the segments intersected during one expanding sweep cycle. It is the largest count which determines the correct number for identification. In general, this maximum count may occur only briefly during part of a cycle, or it may be generated by several different segments encountered sequentially during the cycle. For any non-concentric placement of the figure with respect to the sweep center, one or more sides will be encountered earlier than the others. The relationships are depicted for the case of a square in Fig. 3.

In 3A, the innermost circle represents the minimum sweep diameter where no part of the figure has been encountered. The next circle out, corresponding to a time later in the sweep cycle, indicates the four segments cut from a well-centered figure. The next circle out has just encountered the vertices. Finally the sweep has passed out of the

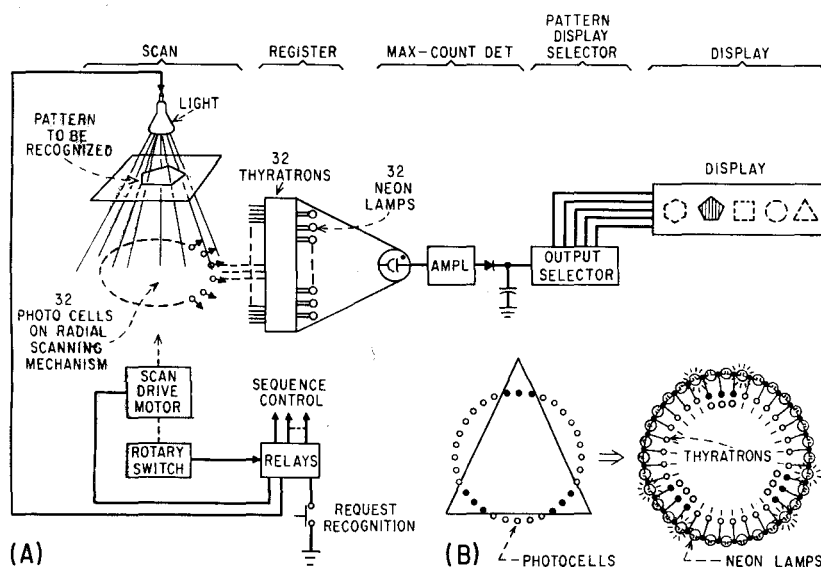


FIG. 4—The n -gon recognition system as constructed (A), and an illustration of how n sides produce $2n$ discontinuities (B)

square; all portions of the inspection circle have encountered the figure.

A slightly non-concentric placement, as in 3B, results in first one side of the square intersected, then three, perhaps four for a moment, then one again. This is in contrast to the case 3A where if any sides were contacted at any instant, all four were. In 3C where the de-centering is excessive, there are never more than three sides intersected at any radius of the expanding sweep. It is required then to detect, store, and sum the number of separate sides encountered during one sweep cycle. This number can never be more than n for any position of an n -gon, and except for accurate centering of regular figures it is generally less than n at any instant.

Let the expanding circular sweep be quantized into a sufficient number of elements, and suppose that those elements which have not passed through the figure are given the value of ZERO; those which have encountered the figure are called ONE. The system then must count the transitions between strings of ONES and ZEROS around the circle. This count thus indicates the number of sides in the figure.

We can distinguish separate objects as follows. For any closed figure lying within the swept area and enclosing the center of sweep, all ONES will eventually be generated; this is not the case for separate objects as in Fig. 3D. Here there will be four discrete strings of ONES and ZEROS but not all ZEROS will have been changed to ONES by the end of sweep. This simple fact is used to provide discrimination of separate objects from continuous (or nearly continuous) line drawings. The count does not depend on the size or shape of the objects, but only on continuity of form, that is, on unbroken strings of ONES.

The block diagram for this system is shown in Fig. 4A. In this version, 32 rather than 64 quantized positions are used in each ring. Also, successive positions of the expanding sweep provide continuous signals rather than 32 discrete signals described earlier in the general model.

The requisite scanning and resultant signals which provide input for the machine are obtained from

32 small photocells which are mechanically puckered across the input plane. As successive portions of the line drawing are encountered, a 32-thyratron register records this information. At the completion of a sweep cycle, this stored data is examined for ONE-ZERO strings (a series of ONES followed by a series of ZEROS).

Conventional means for determining these binary sequences in a register involve shifting the contents of the register or otherwise sequentially commutating the digits. The present system avoids this complication by a simple parallel operation. For a count of the numbers of strings of ONES and ZEROS, the essential information is contained in the transitions between strings, that is, the ONE-ZERO locations in the binary number. For a closed ring of digits, as in the register representing the state of the 32 photocells, the numbers of strings of ONES and ZEROS are equal. Consequently there will be $2n$ transitions for n strings of ONES. Since the number of strings of ONES (or ZEROS) corresponds to the number of sides of the inspected figure, a triangle is represented by 6 transitions, a square by 8, etc.

Parallel inspection of the thyratron register is obtained by connecting a small neon lamp (NE-2) between adjacent thyratron plates. As photocells pass beneath portions of the figure being scanned, they fire their respective thyratrons, which thus change state from ZEROS to ONES. Those thyratrons yet unfired, because their photocells have not been darkened, remain as ZEROS. At the junctions of ONES and ZEROS, the neon lamps light, since there is sufficient potential difference between the plates of fired and unfired thyratrons.

Figure 4B shows that n sides produce $2n$ discontinuities and thereby $2n$ lighted neons.

The lighted lamps are counted in the maximum-count detector circuit (see Fig. 4A) which sums the light output of these lamps and produces a voltage proportional to the number of sides of the figure. This part of the system consists of an enclosure with the 32 neon lamps arrayed in it together with a photocell. The signal from this cell is amplified, peak detected, and stored in a capacitor.

Since the peak detector indicates only the maximum number of neons lit at any instant, there will be fewer than the correct number of sides indicated in some cases. Consider Fig. 3B and 3C. Unless the object and scan centers are coincident, there may not be n intersections at any one time; rather the n sides of the object may be encountered successively during the sweep. One simple solution to this problem is to save sides, that is, to prevent the n strings of ONES from flowing into each other as the vertices of the figure are encountered. This is accomplished by inhibiting that thyratron controlled by the last photocell passing through a vertex. It is locked up by adjacent thyratrons and consequently the neon lamps on either side remain lighted for the duration of the scan.

In the final step, a level detector classifies the stored signal and lights up the corresponding display. Since the amplitude of the signal is proportional to the number of sides in the figure inspected, the level detector is set to report a hexagon for the highest signal, a pentagon for the next lower, and so on through square, triangle, circle, and no-figure.

This detector circuit comprises five thyratrons, each of which has a preset firing threshold corresponding to the signal it is expected to classify. The stored, peak-detected signal is applied in common to all five stages. A common holdoff bias is then removed from one thyratron at a time, starting with the hexagon (highest level) detector and proceeding monotonically to the circle (lowest level) detector. When that stage is encountered where the signal is sufficient to fire the thyratron, recognition occurs.

This machine can also count up to 6 small objects placed at the inspection station. Only if a closed figure is inspected will all photocells be shadowed, whereas if discrete objects are presented, only a few cells, those passing beneath the objects, will have responded by the end of a sweep cycle. The distinction is easily made by monitoring the total plate current of the 32 thyratrons.

A working model of this system, built to demonstrate principles, is shown in the photo. A schematic of

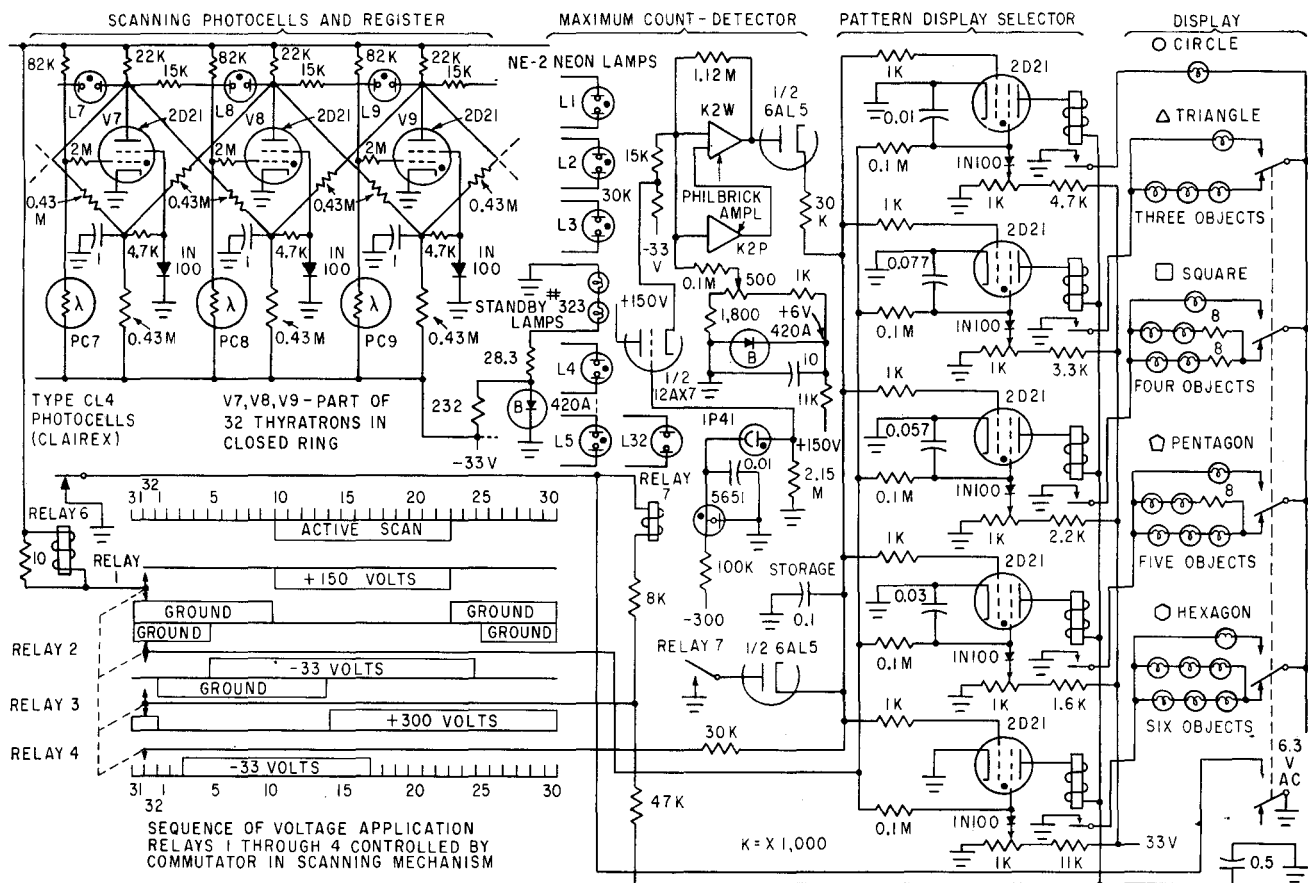


FIG. 5—Schematic of the n -gon recognition system based on the principle of dilating circular scan

the system is shown in Fig. 5.

Scanning photocells are actuated by a motor-driven mechanism when identification of a figure is requested. The same motor drives a rotary tap switch which programs the sequences and durations of voltage applications to different parts of the circuit. The rotary tap switch has two sections of 32 contacts each, one section being used to close control relays and the other section to open them. These relays enable and disable different sections of the circuit by switching appropriate voltages. The sliding contactor arrangement in Fig. 5 represents the operation of these relays.

In the register (Fig. 5), the ONE signals derived from the lines of the figure by the photocells are used to fire thyratrons, thus registering these events for subsequent use. At the start of a cycle, the 32 thyratrons are disabled (no plate voltage) so that the scanning mechanism can come up to speed. When it does, relay 1 applies voltage to the plate bus and the active scan cycle begins. As a photocell passes under a line of the figure, it generates a positive pip, large enough to over-

come the bias on the thyatron control grid, causing firing.

The 32 thyratrons form a closed ring, tied plate-to-plate by small neon lamps such as L_7 , L_8 and L_9 . When one thyatron conducts, its plate potential falls, while that of its neighbors remains high, and so the two lamps on either side will light up. If a neighbor is fired, the lamp between the fired thyratrons, having little or no potential difference across it, goes out. However, another lamp goes on between the newly fired thyatron and its unfired neighbor. As more thyratrons are fired, lamps are progressively turned on and off, the lit one at any moment locating the boundary of fired and unfired stages. As the cycle proceeds, two lamps will light at the first points of contact with each side of the figure, then separate and run out toward the vertices. The photocell in the lamp housing will see $2n$ lit lamps corresponding to all sides scanned, regardless of the temporal sequence. Thus an analog voltage proportional to the number of sides is generated.

In the final step of the recogni-

tion cycle, this electrical analog of an n -gon is classified and then identified by a lighted display.

I acknowledge the assistance of C. F. Matke and S. E. Michaels.

REFERENCES

- (1) E. C. Greanias, C. J. Hoppel, M. Klooomok and J. S. Osborne, Design of Logic for Recognition of Printed Characters by Simulation, *IBM Journal*, Jan. 1957.
- (2) D. H. Shepard, P. F. Bargh and C. C. Heasley, Jr., A Reliable Character Sensing System for Documents Prepared on Conventional Business Devices, *IRE Wescon Convention Record*, Part 4, 1957.
- (3) Solarton Electronics Corp., Ltd., Surrey, England.
- (4) K. R. Eldredge, F. J. Kamphoefner and P. H. Wendt, Automatic Input for Business Data Processing System, *Proc Eastern Joint Computer Conference*, Dec. 1956.
- (5) W. W. Bledsoe and I. Browning, Pattern Recognition and Reading by Machine, *Proc Eastern Joint Computer Conference*, Dec. 1959.
- (6) J. S. Bomba, Alpha-numeric Character Recognition Using Local Operations, *Proc Eastern Joint Computer Conference*, Dec. 1959.
- (7) R. L. Grimsdale, F. H. Sumner, C. J. Tunis and T. Kilburn, A System for the Automatic Recognition of Patterns, *Proc IRE*, Dec. 1958.
- (8) H. Sherman, A Quasi-Topological Method for Recognition of Line Patterns, *Proc Int Conf on Information Processing*, Paris, June 15-20, 1959.
- (9) S. H. Unger, Pattern Detection and Recognition, *Proc IRE*, 47, No. 10, Oct. 1959.
- (10) W. Sprick, Character Reader—U. S. Patent 2,838,602, June 10, 1958.
- (11) L. D. Harmon, Line Drawing Pattern Recognizer, *Proc Western Joint Computer Conf*, May, 1960.