A character recognition system is described which incorporates a lens-and-retina input and a relatively simple computer. The system operates effectively with the image focused anywhere on the retina and may operate so as to be independent of image, size or orientation.

The characters which may be unambiguously discriminated are those which have distinct transform functions $T$. A simple algorithm is given for obtaining $T$ for both curvilinear and rectilinear figures. Examples are given and possible means of resolving the ambiguities discussed.

The Transform of a Pattern

This paper describes the conceptual mechanization of an optical scanner designed to accomplish a specific, well-defined task. After the machine is described, there is given a brief discussion of processes, events and components which form part of the psycho-physiological visual system. It appears that there are, at least superficially, a number of analogous concepts.

Computers designed to recognize patterns may take either of the forms shown in Figure 1. Some preprocessing on the inputs results in a transform (the round dot in the Figure). It is this transform which is identified or recognized by the computer. Usually, the transform is a binary number. The number may be identified as such, existing in some specific registers in the machine, or it may be physically or logically distributed, and represented by the binary 0-1 state of several elements of different circuits. But, in
either case, it is this number—this transform—on which the identification or recognition logic operates.

The two schemes shown here are, perhaps, only different diagrams of the same functional blocks. In any case, it is the intent of this paper to discuss a class of preprocessing operations and the transforms which they produce. It is taken for granted that if a binary number exists in the machine which contains sufficient information to discriminate the pattern of the input, then logical operations can be formulated to produce as output a signal identifying the pattern presented. But the form of the transform function is important because it has a strong influence on the capacity of the memory or the complexity of the logic required for recognition.

The preprocessing includes, typically, an optical-to-digital transducer system, a magnetic reader, or an acoustic-to-digital transducer system. It may, also, include certain logical operations on the digital data. And, of course, it may use an analog, rather than digital techniques.

The Chrysler Concept

Figure 2 shows the system which is to be discussed. We are concerned with producing a transform; that is, a group of digital bits which can be uniquely associated with the character presented.

For some applications, it would be nice to associate the characters in Figure 3 as being alike and to produce the same set of binary bits as the transform for each of them.

More formally, it would be nice to have the machine identify a square form, regardless of its size, its orientation, or the location of its image on the retina.

In other cases, however, these differences are as important as the form. If the machine scans the equation at the top of Figure 4, it will see 12 separate symbols, but these are made up of only 4 forms, (A, Bar, 6, and dot).

\[
A^A = [6.9 \cdot A]
\]

\[
A^A = [6 \cdot 9].
\]

Figure 4. A typical equation to be scanned, and a hypothetical transform in which disjoint sets of bits discriminate form, size, orientation and location of the image.

If we can generate a transform (shown conceptually as a binary number in Figure 4), in which certain digits are identified with shape or form, others with size, some with orientation, and still another set with location, the logical operations of recognition will be much simplified. The number of shapes which must be memorized by the machine will be substantially reduced. It is not the intent here to justify the desirability of having disjoint sets of bits, each set associated with some property of the character. There is enough literature on the subject already, both in the computer field and in the psychological analysis of human perception.
To a limited degree, the system to be described does just this. Of course, this is not the whole problem. A general purpose pattern recognizer may be confronted with Figure 5 and must recognize not only the similarities but also the differences between these sets of characters.

![Figure 5](image)

Figure 5. A typical problem for a general pattern recognizer which should discriminate both the similarities and the differences in the sets of figures.

But, even though it doesn’t do everything that might be asked of it, there seems to be a measure of utility in the system shown in Figures 6 and 7. Here a part of the preprocessing is done optically. An optical wedge is rotated so that the image is caused to rotate on the retina. The wedge driver motor also drives a commutator which puts out timing signals. The circuitry is shown in Figure 7.

![Figure 6](image)

Figure 6. The conceptual arrangement of the Chrysler Optical Processing Scanner.

![Figure 7](image)

Figure 7. Schematic circuits of the Chrysler Optical Processing Scanner.

Each of the n photo receptors in the retina is connected to a differencing circuit or "flicker filter," (Symbol F), which compares the signal now with what it was an instant ago and emits a pulse if they are different. These pulses are connected to one large bank of OR gates. The output at Point A, Figure 7, is one pulse every time any one of the receptors changes state. The number of receptors and duration of single pulses are such that the overlapping of signals from two receptors to form a single output pulse at "A" is statistically a rare event. The effect is ignored. The m timing signals direct traffic of these pulses so that, for the first increment of rotation of the wedge, the pulses are accumulated in Counter No. 1, the next increment of rotation reads pulses into Counter No. 2, and so on. If there are 180 counters, each will, after 180° of wedge rotation, have counted the flickers during 1 degree of travel. The counters are emptied every 180°.

The counter readings may be plotted as in Figure 8. The curve these points represent may properly be called a transform of the character which generated it. It is denoted as the T* transform.

Two kinds of differences exist between a real transform and an ideal one. In the left half of Figure 8A is shown the result of using an infinitely divisible retina and a finite number of counters. The curve is approximated by a step function, but no better approximation is possible without using smaller increments. In the right half of the same Figure, both a finite size and a number of retinal
elements and a finite number of counters are used. The approximation is somewhat more irregular.

In what follows the idealized transform in Figure 7B is discussed. It is assumed that there are infinitely many retinal elements and counters. The degree of approximation permitted in a real machine will depend on the job it has to do.

The T* transform in the counters may be cyclically shifted, as shown in Figure 9, until the smallest value lies in Counter No. 1. The result is denoted as the T_R* transform of the character. If after this operation, each counter is divided by the number in Counter No. 1, a T_N* transform is produced.

With this description of the system in mind, we can now begin to discuss its characteristics. The machine description above is entirely conceptual. It must be emphasized that such terms as wedges, counters, commutators, et al are only convenient terms for description; the hardware of a real scanner would be functionally the same, but physically far different.

**Discriminability and Invariance**

The transform of a circle is a horizontal line, (Figure 10). The rate at which the circle covers and uncovers the receptors as it nutates is constant.

The transform of a square is proportional to $\sin \theta + \cos \theta$, where $\theta$ is the rotation or phase angle; with a reference established by setting $\theta = 0$ at the time the pulses begin to accumulate in the first counter.

Since the transforms are different, at least these two characters can be discriminated.

Consider next the two characters in Figure 11. The transforms of these two characters are identical, except that one is out of phase with the other by an amount equal to the difference in angular orientation. But in the computer, the T_R* transforms will both have been shifted (by different amounts) until they have the same position in the counters.
Figure 10. Two simple characters which are discriminated. The $T^*$ transform of a circle is a horizontal line; for a square, it is a $(\sin \theta + \cos \theta)$.

Therefore, the $T^*_R$ transforms of these two characters are identical. If we wish to compare the transforms with a number of pre-stored bits, one set of bits in the memory will match either character.

But if, in the shifting of the transforms in the counters, we count the number of shifts performed to bring the minimum counter reading into the first counter, the result is a number which gives a measure of the orientation of the character of the retina.

In Figure 12, one more variable is introduced—size. The larger figure will "sweep out"—i.e., cover and uncover—more receptors and produce more flickers than the smaller one. And the number of flickers generated will be linearly proportional to a linear dimension of the characters. But, in generating the normalized transform, $T^*_N$, every ordinate of the $T^*_R$ transform (or every counter reading) was divided by the minimum value. The result is that Counter No. 1, which holds the minimum value, reads "1", for both transforms—and the $T^*_N$ transforms of the two characters are identical.

Figure 11. Two characters having distinct $T^*$ transforms, but identical $T^*_R$ transforms.

The information as to the size of the image, however, need not be lost. The reading in Counter No. 1 of the $T^*_R$ transform gives a measure of the size of the character.

The foregoing results are stated more formally in the Appendix.

At this point, we may summarize the properties of the scanner, as in Figure 13. The form of a character generates a group of bits, $T^*_N$, which are identical for all sizes, orientations and locations on the retina. It is truly a measure of form or shape only, unaffected by the other three properties.

<table>
<thead>
<tr>
<th>Form</th>
<th>Size</th>
<th>Orientation</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T^*_N$</td>
<td>Counter No. 1</td>
<td>$T^*_R$</td>
<td>$T^<em>$ to $T^</em>_R$</td>
</tr>
</tbody>
</table>

Figure 12. Two characters having distinct $T^*_R$ transforms, but identical $T^*_N$ transforms.

The orientation of the character is given by the count of the number of shifts to bring the minimum reading into Counter No. 1. This orientation has no meaning if we wish to compare the orientation of Character A with Character B. It is only meaningful in comparing the orientation of 2 A's, one of which can, of course, be a standard which is displayed erect and stored in the memory.

The size of a letter, measured by the entry in Counter No. 1 of the $T^*_R$ transform likewise does not compare the size of an A with a B, except, perhaps, very approximately.

In this device, then, it appears that the perception of form is independent of the other
parameters; the others are not, however, completely independent of form.

It is patently impossible for the systems described to provide a measure of the location of the image. The processing takes note of how many receptors flicker and at what time during the nutation cycle. Nowhere is there any information as to which receptors change state. This is not the only deficiency in the device.

Two characters, which are the same except for a 180° rotation, are not discriminable. This is evident from the fact that the transform is cyclic with a 180° period. For the same reason, we cannot distinguish a re-orientation of the figure through 90° from one through 210°.

A second form of ambiguity arises because if two characters or parts of characters are scanned simultaneously, the transform is the sum of the transforms of the two when scanned separately. This source of ambiguity may also be ascribed to the inability of the machine to sense the location of an image, or part of an image, on the retina.

These two ambiguities are illustrated in the next two Figures. Figure 14 shows figures which are ambiguous due to 180° rotation—polar symmetry.

\[
\begin{align*}
\text{Symbol } \oplus & \text{ denotes two characters which have equal transforms.} \\
6 \oplus 9 & \rightarrow \oplus
\end{align*}
\]

\[
[T^*]_\oplus = [T^*]_\circ + 180
\]

Figure 14. Examples of characters not discriminable (by the basic flicker-count circuits) because of polar symmetry.

Figure 15 shows figures which are ambiguous because they are each the sum of the same elements, the elements being located differently with respect to each other. An element is defined as a boundary between a black area and a white one.

A machine which cannot distinguish a 6 from a 9 is somewhat limited in its usefulness.

\[
\begin{align*}
\text{A } & = \nabla \\
\text{5} & = \quad \bigcirc
\end{align*}
\]

Figure 15. Examples of characters not discriminable (by the basic flicker-count circuits) because of failure of these circuits to provide image-location information.

The obvious question is how to add necessary circuits without losing the desirable characteristic that form, size and orientation are separately identified; and, also, without complicating the circuitry by trying to recognize which receptor is excited at a given time.

In a 360° cycle, the number of receptors turned on will always equal the number turned off. It is evident, with a little geometric reasoning, that each receptor which is affected at all must in 360° change state an even number of times. But, consider the characters in Figure 16. These two characters have identical transforms, but in one case there are two changes-of-state in each affected receptor and in the other there will be four changes-of-state in some receptors. A count of the number of receptors which change state four or more times becomes an additional

\[
\begin{align*}
\text{Figure 16. Examples of characters discriminable by counting receptors which have multiple changes-of-state.}
\end{align*}
\]
transform and provides the information which will distinguish these two characters. This scheme is, of course, in addition to, rather than instead of, what has already been described.

There are other schemes—many of them—which will assist in the discrimination of characters. Some of these schemes are better at resolving one form of ambiguity, others are better for other characters. Two of these will be discussed.

Imagine a feedback circuit associated with each receptor, such that when the receptor changes state its output is blocked for a short time interval. This has the effect of inhibiting the second of two flickers which occur too close together in a single receptor. Figure 17 illustrates the effect of this circuit on two otherwise ambiguous characters. The effect is to produce for the two characters transforms which are different. The flickers which would normally be produced by the segments shown dotted do not appear when the dotted segments follow too closely the adjacent line. It is somewhat as if fillets were added, but these fillets vary in size as the image mutates and only appear at all when the nutational phase angle is in the appropriate quadrants. It is, therefore, feasible in principle to utilize the transform as modified by this device instead of the ones previously discussed.

The next ambiguity-resolving scheme was derived from trying to make a 6 and a 9 fundamentally different by optical distortion. If the image formed on the retina results from reflection from a warped mirror, so as to produce distortion as in Figure 18, the polar symmetry disappears and distinguishable transforms results. However, the equivalent of optical distortion can also be produced as shown in Figure 19. The size and spacing of retinal elements is progressively diminished near the bottom of the retina. The effect is to increase the number of flickers produced by the lower half of the character exactly as if the image had been optically distorted as before. The transform can still be "normalized" as before to produce invariance with respect to size, location and orientation, although the computations to do so are more complex than the simple shift and divide operations described above.

Figure 17. Examples of characters which are discriminable by circuits which inhibit the second two changes-of-state occurring too close together.

Figure 18. A (hypothetical) strategem for producing discriminable images from characters not discriminable by the basic flicker-count circuits by optical distortion.

Figure 19. Schematic representation of a non-uniform retina which simulates electronically the effect of optical distortion.
This scheme suffers from the fact that the retina at the top must be fine enough to produce sufficiently smooth transforms, and finer still at the bottom. Consider the next modification in the sequence, (Figure 20).

Figure 20. Schematic representation of a non-uniform grouping of retinal elements which produces the same result as the retina of Figure 19, but without variable size or spacing of retinal elements.

Here, a uniform retinal matrix is used and transforms are generated, as illustrated in Figures 6 and 7. In addition, taps from the receptor outputs are connected to additional circuits. Near the top, groups of receptors are grouped in these circuits so that they behave as single receptors. To do this, the circuits effectively pass the flicker pulse associated with the first of the group to change state and then inhibit the remainder of the group from emitting pulses. In these circuits, therefore, the group acts as a single receptor. The effect of grouping receptors in this way is the same as if the optically distorted 6 or 9 had been focused on the retina and the characters are, therefore discriminable.

The number of variations and additions to these concepts has not been exhausted in the descriptions above. For implementation in a generalized pattern-recognizing computer, several of these schemes may be combined.

Further discussion of the ambiguities resulting from the simple circuits and of means of resolving them is not fruitful without more rigorous mathematical formulation. It is possible, however, to list a few of the transform-generating schemes which may be useful for specific applications.

1. Count the total number of changes-of-state and read them into counters by the use of timing pulses. (This is the basic count-and-sort arrangement described above.)

2. Count and sort the changes-of-state while the image moves off the edge of the retina. Alternately, after 180° of motion, switch out or inactivate those receptors just ahead of the moving image so that, effectively, an "edge of the retina" is artificially produced.

3. Count and sort separately the number of receptors which change state two, four, six or more times.

4. Count the number of receptors illuminated (or dark) at some instant.

5. Number the rows of receptors, from 1 to U, and let each row generate a number of pulses r, such that r = f(U). Together, with normalizing circuits, this scheme provides one coordinate of the location of the image.

6. Number the column, as in 5, to obtain the other coordinate for the location of the image.

7. Count and sort only the first change-of-state for each receptor in each cycle.

Imperfect Patterns

Some comment on the uncertainty in the shape or form to be recognized is appropriate. The analysis has been based upon the assumption that characters to be recognized as identical were, in fact, congruent, pure black on a pure white background.

Figure 21 shows an enlargement of typical characters from 12 point type. The irregularities on a vertical outline will effectively increase the flicker count most during that portion of a cycle when the image is moving vertically. The result is to increase the flicker count—never to reduce it.

Figure 21. A typical enlarged character, originally set in 12 point type, and illustrating characteristic irregularity of outlines.
Several 'hole filling' or contour smoothing techniques have been discussed in the literature. It appears possible that a relatively simple strategem can be effective in systems utilizing image nutation, as is described here. If the image is optically defocused, a gray outline is formed, like that of Figure 22. The lower figures represent the effect of equal blurring of the characters above. These fuzzy characters were scanned, using autocorrelation techniques depending on which receptors were stimulated above a given threshold, the thresholds or sensitivities would have to be very carefully controlled. Otherwise, the low brightness gradient would magnify, rather than diminish, the irregularity of the border.

Figure 22. A pictorial description of the influence of defocusing or blurring on a typical and a perfect letter. The blurred images are more nearly alike.

However, consider an image moving normal to such a fuzzy edge. The receptors which are at the low end of the threshold tolerance would change state late, those at the high end early. Since we are simply counting the changes-of-state and don't care at all which receptors are involved, these effects tend statistically to cancel or "average out."

During that portion of the nutation cycle when the image moves parallel with a fuzzy edge, low and high threshold receptors are likewise less likely to undergo unwanted changes-of-state. This technique is, in fact, akin to certain hole filling schemes, except that the filling is accomplished by distributing light from a point to adjacent points, rather than by distributing calculated probabilities.

The blurring of the image may also be created by other means, as by superposing a high frequency vibration on the lens-object-image system, or by passing the image through a silk screen, or equivalent grid. This latter technique may be considered to be borrowed from the photographers who "soften" their portraits to remove harsh contrast.

Nutation and Other Image Motions

Mention should be made of the reasoning which led to the use of nutation as the image motion which is most effective. It is intuitively evident (and is simply derived mathematically) that line segments of characters produce little or no flickers and hence do not input a signature to the computer while their motion is parallel with the segment. To pick up lines of character in any orientation, there should be some interval during which the motion is normal to every line. Nutation is such a motion.

It may be more formally stated that no motion generates a transform having more information content than nutation. There are, however, many motions which do as well. Any of these may be resolved into a nutation and some superposed motion. Consider, for instance, the superposition of a lateral translation and nutation, which produces a net motion, such as is shown in Figure 23, for each point in the image.

The translation adds or subtracts from the lateral (horizontal) component of nutation velocity. So long as the superposed lateral velocity $V$ is always less than the velocity of the nutation, $R\omega$, the image vector will, at some time during its cycle assume every direction from $+\infty$ to $-\infty$ slope. Such a motion produced no loss of information. So long as the superposed motion is known, the flickers it produces or inhibits do not subtract from or add to the information content of the transform. When such a superposed image movement exists, it may be found desirable to generate timing pulses of unequal duration.

Figure 23. Sketch of the cycloidal path of an image-point over a retina which results from superimposing nutation and lateral translation.
It is, therefore, nearly as easy for the scanner to "read" a moving character as a fixed one. The capability for discrimination of images without bringing them to rest in front of the lens can be a useful asset for some applications.

The technique of moving the image may be thought of as a sort of time-sharing of circuitry in which the switching is accomplished optically. Instead of connecting "edge autocorrelation" circuits to various receptors in sequence, the image is "connected" sequentially with different receptors in sequence. Thought of in this light, the optical wedge replaces the switching circuitry.

Mirrors and Fiber Optics

There are a few further strategems which may be used to reduce still further the number of logic blocks and receptors required. Not all of these can necessarily be used simultaneously. The choice of which to use depends upon the application.

In Figure 24 is shown a character being mutated over a semi-circular retinal array of receptors. A system of mirrors is so arranged that as the image leaves the retina on top, its reflection enters below. The number of receptors and the circuits associated with each is halved.

In Figure 25 is shown a different arrangement. The matrix of retinal elements is composed of the ends of optical fibers. The other end of the fibers carries light to the photo receptors proper. All of the fibers bearing the same number communicate to the same receptor. In the scheme shown, 459 fiber ends are connected to 91 photo multiplier tubes.

If the largest character to be scanned can be completely enclosed within a circle inscribed in one of the large bold-face hexagons, no loss of information will result. It is evident that with the size restriction it will never occur that one part of a character moves to turn on or illuminate a tube which is already on, or vice versa.

It is probably true that the highly organized connectivity shown, where the optical fibers are accurately grouped, is not really necessary. Within limits, the fiber ends could be connected "at random" to the photo receptors. In this case, if the multiple connections are carried too far, large sample statistics govern, and the instantaneous "sample" of elements turned on and off will always be the same as that of the population. The "population" of flickering cells have an average of zero, since every receptor turned on will be turned off during a cycle. There will, therefore, be no net change of illumination and no flickers. But if the statistical scheme used to define "random connectivity" is properly selected, random connections may be used with only small degradation of information content in the transform.

Composite Images—Words vs Letters

Some interesting properties of this scheme arise from the theorem, stated earlier, that the transform of a character is the sum of the transforms of its parts, taken as if they were separately scanned. Suppose the image focused on the retina consists of two letters, A and B. The transform will then be the sum of their separate transforms. Assume, for the moment, that there are a sufficiency of the ambiguity-resolving circuits to discriminate each letter separately. If there are no two or more characters, the sum of whose transforms is identical with that of a third one, then the transform of AB may be uniquely associated with these two letters, but not their order.

This property opens up the possibility of reading words rather than letters. Obviously, if the output is an electric typewriter, no advantage accrues to the system. If, however,
Figure 25. Another scheme for diminishing the number of photo receptors required by connecting the optical fibers (shown as circles) to photo receptors (numbered).

the output is to a language-translating computer, which must have a dictionary stored in its memory, anyway, the scanning of words rather than letters may be useful.

Without going into detail, this same theorem may be said to offer the possibility of analyzing a character into its separate parts. The analysis will, in general, not yield completely unambiguous results. However, if one leg of a character is missing, it is reasonable to consider predicting probabilistically the intended character and the rechecking by other techniques.

Scanner Characteristics Summary

The Chrysler Optical Processing System exhibits the following characteristic behavior:

1. There is generated within the logic a transform of the image in which character form, size, orientation, and location are represented by disjoint sets of bits.
2. Image movement may, if convenient, be superposed on rotation—characters need not be brought to rest to be read.
3. Character imperfections may be effectively smoothed out by blurring, without complex hole-filling circuitry or precise biasing of receptor thresholds.
4. Accurate servo-centering of images on the retina is unnecessary.
5. Words may be read instead of letters; characters may be analyzed into their elements.

From the collection of the Computer History Museum (www.computerhistory.org)
6. Various means, both optical and electronic, may be used to reduce the number of retinal elements and logic blocks required to obtain the effect associated with a fine-grained retinal matrix.

Each of the six statements above should include the phrase, "subject to limitations as discussed in the text."

The Psycho-Physiological Parallel

The description of the scanner and a survey of its capabilities, both immediate and potential, is complete. It would be amiss to omit mention of the striking, though possibly superficial, similarities which exist between this scanner and the processes, events and components of the human visual system. There are several of these psycho-physiological analogies.

The image formed on the human retina is not still, even when the observer attempts to fixate a particular point. The motion of the fixation point on the retina has been measured and looks something like that of Figure 26. The movements take the form of sharp, fast flicks, with a slower drift between flicks, rather than the smooth nutation which is convenient for the machine. Superposed on this is a high frequency tremor.

It has been reasoned that the nutation of the image is an optimum image-movement tactic. As a speculation, it would be of interest to examine how well these drifts and jerks approximate nutation. The obvious first choice for a straight-line approximation to a circular motion is a regular polygon.

In the inanimate scanner, we must accumulate the transform for whatever time is required to nutate 180° before the transform is complete. If after a transform is generated the machine memory could only be exposed to two small parts of it, what parts would yield the most information? Vertical motion yields the most information about horizontal lines and vice versa. It would be to our advantage to select the two parts of the transform so that they represent the count of changes-of-state of receptors at two instants when the image motions are perpendicular to each other. Or, conversely, after the image has moved vertically, and the changes-of-state recorded, the next motion should not be nearly-vertical since this adds less to the information already on hand than if the second movement is approximately horizontal. Suppose, therefore, that we retain the regular polygon as an approximation, but traverse the legs out of order. Furthermore, movement in either direction, A to B, or B to A, produces basically the same set of changes-of-state.

One further charge in the "regular polygon" idea is required. If the polygon is exactly regular, and of n sides, the directions of movement on the n + 1st movement will exactly coincide with the first. It would be better if we chose a figure whose sides progress angularly.

In summary—

1. Select a regular polygon of n sides (Figure 27A).
2. Adjust the angles slightly so that no two sides of any group will be exactly parallel (Figure 27B).
3. Rearrange the sides out of order so that the angle between adjacent sides is approximately 90° (Figure 27C).
4. At random, reverse the direction of the sides, taking BA instead of AB, for instance (Figure 27D).

With these considerations in mind, the movements of the image in the eye seem not too dissimilar to what we would build into our engineering artifact if mechanical convenience were not a factor. Straight-line motions which start and stop are just not as easy or as cheap to make out of metal as they are of muscle. It may be noted that the machine described fails to provide "location" information. It can, of course, be supplied by additional circuits which have not been described.

Figure 26. Typical movement of the fixation point of the human eye, consisting of a high frequency tremor, a series of jerks, and a saccadic jump to compensate for the cumulative drift due to jerk.
If image signals are generated in the human retina during both the "flick" and the drift, it is not unreasonable to suppose that the analysis of the signals generated by two different motions are analyzed by two different schemes. More specifically, can one of these motions be considered to generate signals which are analyzed in the brain without attention to which receptor is active while the other motion is associated with the location of image parts?

It would be interesting to follow the speculation a little further and see what sort of eye movements are observed if the subject is exposed to a visual field composed of heavy vertical lines and nothing else. Vertical eye movements in such a field will not produce any changes-of-state and hence produce no information. This is probably not strictly true, except in an optically perfect eye without aberration. If the principles discussed in this paper are truly analogous to human vision, and if there is feedback in the system, then a vertically-ruled field should induce more horizontal jumps in the eye scan.

The visual process of observing a moving object superposes another motion on the image motion due to scanning. In the inanimate device, information is lost if the supposed translational velocity exceeds the velocity due to nutation. In a speculative sense, it might be interesting to try to discover in a human eye what rate of movement of an object is associated with a loss of information as to its form, and to relate this to the scanning speed.

Some of the other analogies between machine and eye are certainly obvious to the reader. Both are sensitive to changes of illumination rather than illumination directly.

The machine use of a refractory period which inhibits the second of two pulses too close together has a direct analog in the refractory period of neurons. The grouping of local clusters of photo receptors to a single channel output was discussed as a means of producing the equivalent of optically distorted images. This is a reflection of the multiple connections of rods to a single neural path. Neither of these two phenomena have, to the author's knowledge, been associated previously with form perception, yet they appear, in the inanimate machine at least to be well-suited to the task.

The circuits which count the receptors which flicker more than twice requires what in the computer field is termed a flip-flop; and in the field of psychology a form of "temporal summation."

The machine requires a number of counter registers in which to store the flicker count. It is inappropriate to discuss such counters in the physiological system. In the machine, signals travel to the counters at the speed of light and are held in the counters as long as needed, as in Figure 28. In the brain, signals travel slower. We may surmise that at any instant the various counters have their counterpart in the various links in the neural

---

From the collection of the Computer History Museum (www.computerhistory.org)
pathway. That is, the several counters are analogous to the several stations along an appropriate set of neural paths which carry the signal at a given instant.

These signals may reverberate back and forth, perhaps undergoing phase shifts and other normalizing operations in the process.

Figure 7 showed the mechanical scheme by which the count of changes-of-state are switched or directed into the various counters in sequence by the timing pulses from the commutator. Figure 29 shows the same schematic in which the inhibitory or excitatory effects of one set of synapses accomplish the timing function and direct traffic for a set of receptor outputs. There is little or no direct evidence that some of the neural pulses are image signals and others are timing signals related to the scanning motion of the eye movements, but there does not appear to be any direct evidence that such a concept is impossible. The possibility also exists that a given axon, synapse or receptor acts sometimes as a timer and sometimes as a signal detector.

Conclusion

As a peroration, certain rather abstract observations must be brought forward. The neuro-visual process may be rather vaguely defined as a "mapping" of certain changes-of-state of the retinal receptors into the brain. This mapping process is simply another way of stating that a "transform" is generated in the brain as a result of retinal activity. The map or transform may be either static or dynamic; that is, it may be defined by specifying either the state of a set of neurons, or a sequence of changes-of-state. In either case, there is a powerful appeal, both intuitively and logically, in a system where the transform of two images is the sum of their individual transforms. Such a scheme is, for different reasons, appealing to the computer engineer and the psychologist.

The details of machine circuitry discussed are not likely to carry over entirely intact their usefulness into the psychological field. The machine principles may be described:

1. A moving image on a retina can provide much information about form, even with circuits which do not distinguish one receptor from another.

2. More information about form can be developed in the scanner if the direction of the scanning motion associated in time with change-of-state of a receptor is available in the form of timing signals.

3. Means can be conceived and circuits defined for form discrimination in which the transform identified is the sum of the transforms of its separate parts.

Stated in these terms, the principles on which the machine operates may be equally interesting to the psychologist.

This paper does not claim to present a new psycho-physiological theory of form perception. The description of the machine is a sequence of declarative sentences—the psychological discussion of a series of interrogative ones. The questions it asks may stir up a fruitful train of thought, but the answers are likely to go far afield from the present simple concept.

APPENDIX

Certain of the relationships between characters and the transforms they generate are repeated here in a form more suitable for mathematic analysis than oral presentation.

1. Let a character be defined as the set of points contained within a finite number
of closed boundary lines and distinguished from points not in the set. If points within the character are distinguished by being black and those outside are white, the character is positive; if the character is white on black, it is negative.

2. A positive character is solid if a closed figure can be drawn which contains all the black points and no white ones; otherwise, it is hollow.

3. A character is rectilinear if all of its boundary lines are straight. It is curvilinear if one or more of its bounds are curved.

4. A rectilinear character is convex if all of its internal angles are ≤ π; otherwise, it is concave.

5. A curvilinear character is convex if all of the internal angles between its rectilinear elements are ≤ π and if no tangent can be drawn to its curved bounds which intersects the character.

6. An element of a character is any segment of its boundary (and which, therefore, has white points on one side and black ones on the other).

7. The T* transform of an element of a character which is nutated with radius R and has an instantaneous position, as indicated by θ (Figure 30), is the rate at which the element sweeps out area, expressed as a function of R, θ, and the dimensions of the element. Specifically, by definition

$$T^* (E) = S \cdot R \frac{d\theta}{dt}$$  \hspace{1cm} (1)

where S = the projected length of E on a line parallel with the instantaneous radius R.

Letting $\theta = \omega t$ gives

$$T^* (E) = R \omega S$$  \hspace{1cm} (2)

8. If $E_1$ and $E_2$ are congruent elements which can be superposed by pure translation without rotation, then

$$T^* (E_1) = T^* (E_2)$$  \hspace{1cm} (3)

9. The transform of a character C is the sum of the transforms of its elements.

$$T^* (C) = T^* (E_1) + T^* (E_2) + \ldots + T^* (E_n)$$  \hspace{1cm} (4)

10. The transform of the combination of two elements is the sum of their separate transforms.

$$T^* (E_1 + E_2) = T^* (E_1) + T^* (E_2)$$  \hspace{1cm} (5)

Figure 30. Notation and derivation of T* transform for a straight-line element.
The transform (2) may be written

$$T^* (E) = R \omega f (\theta, d_i, \phi)$$  \hspace{1cm} (6)

where $\theta$ is the nutation angle measured from a fixed arbitrary reference

di are the dimensions of the element

$\phi$ is the angle between the reference axis $\theta = 0$ and any dimension of the character and which, therefore, specifies the angular orientation of the character.

11. If two characters are congruent, their $T^*$ transforms can be made equal by rotation of the reference axis for $\phi = 0$ through an angle $|\phi_1 - \phi_2|$. Hence, if the reference axis for $\theta = 0$ is taken at the same angular orientation with respect to $\phi$, a new transform $T^*_R$ is defined which is independent of character orientation.

12. If two characters $C_1$ and $C_2$ are congruent,

$$T^*_R (C_1) = T^*_R (C_2)$$  \hspace{1cm} (7)

13. There corresponds to every solid convex character one and only one transform $T^*_R$.

14. Every hollow character is concave.

15. For every concave character, there corresponds one and only one solid convex character which has the same transform $T^*$.

16. The family of concave characters, all of which correspond to the same solid convex character, are called an ambiguous family. An ambiguous family are denoted by primes with the same base and the unique solid convex character by the unprimed notation. Hence,

$$T^*_R (C_1^1) = T^* (C_2^1) = T^* (C)$$  \hspace{1cm} (8)

$C$ is called the unique character of the family $C_1$.

17. An algorithm for determining the unique member of a rectilinear family for which one concave member is given is as follows (Figure 31).

1. Number each line of the boundary and add arrowheads in sequence around the boundary.

2. Tabulate the angle each vector makes with the first one.

3. Retabulate the bounding vectors, ranking them in increasing order of their corresponding angles.

Figure 31. Example of generating the convex solid (unique) member of an ambiguous family given one concave member.
4. Redraw the figure, using the same vectors, head to tail, but taking them in the order in which they were ranked in Step 3.

18. The algorithm for determining the unique member of a rectilinear family for which one hollow member is given is the same as in Step 3 above. Number and rank the vectors in both the exterior and the interior boundary, as in Figure 32.

19. If two characters or elements have polar symmetry with respect to each other, their T* transforms are identical.

\[ T^* (C) \big|_{\theta} = T^* (C) \big|_{\theta + \pi} \]

20. Two curvilinear elements of a character are parallel if their chords are parallel, taking the chords as vectors having the same sense as the vectors of the curvilinear elements.

21. Given a curvilinear element \( E_1 \) described by \( y = f_1(x) \) and having a derivation \( y' = f_1'(x) \) which can be solved for \( x \) to give \( x = \phi_1(y') \) and given, also, a second parallel element characterized by \( y = f_2(x) \); \( y' = f_2'(x) \) and \( x = \phi_2(y') \), then a third element \( E_3 \) defined by the equation \( x = \phi_1(y') + \phi_2(y') \) has the property that

\[ T^* (E_3) = T^* (E_1) + T^* (E_2) \]

22. An algorithm for finding the unique member of an ambiguous family, when one curvilinear concave member is given is as follows (Figure 33).

1. Add arrowheads and number the elements as in 17 and 18. Dividing up the boundary into elements,

   (a) Each straight line segment is an element.

   (b) Each segment of a curved line is an element where the segment ends occur: (1) where a straight line meets a curve; (2) where two curved lines meet with a discontinuous first derivative; (3) at a point of tangency of a curved line with a tangent parallel to any rectilinear element having the same sense as the curve; (4) where a point is required to divide a curve into an element parallel with another and a non-parallel element.

2. Where two or more curvilinear elements are parallel with each other, eliminate them and substitute a third element which has an identical transform, using 21.

3. Rank the elements by their angles, as in 17, using the chord of an

---

Figure 32. Example of generating the unique member of an ambiguous family given one hollow member.
element to determine its angular orientation.
4. Proceed, as in 17, to draw the figure which is the unique member of the family.
23. No motion of the image of a character over a retina can provide more information about the character than nutation; nutation defines completely the unique member of the family.
24. Any motion which translates the image without rotation so that a tangent can be drawn to the hodograph of a point having any slope from $-\infty$ to $+\infty$ will provide as much information as nutation and will define exactly the unique character of a family.
25. Specifically, a linear translation, such as would be produced by moving characters in a straight line past the lens, may be superposed upon the nutation and so long as the net nutation velocity $R \omega > V$, the linear velocity, it will identify the unique member of the family.
26. The unique member of a family is completely identified by $180^\circ$ of nutation.
27. If two characters $C_1$ and $C_2$ are similar, and two corresponding sides have the ratio 1: a respectively, then
\[ a \, T_R^*(C_1) = T_R^*(C_2) \] (9)
and the corresponding transforms are proportional.
28. Define a normalized transform $T_N^*$ by
\[ T_N^{1*} = T_R^*/[T_R^*] = \theta_0 \] (10)
In which each value of the transform is divided by its value at an arbitrary nutation angle $\theta_0$.
29. Then, if two characters $C_1$ and $C_2$ are similar,
\[ T_N^*(C_1) = T_N^*(C_2) \] (11)
BIBLIOGRAPHY


