18 August 1972, Volume 177, Number 4049 SCIENCE

Visual Pattern Analysis in Machines and Animals

The same principles may underlie the operation of sensory neurons, computer programs, and perception.

H. B. Barlow, R. Narasimhan, and A. Rosenfeld

The study of systems that are capable of processing pictorial information has been of concern to three distinct categories of research workers: the neurophysiologist who is interested in mechanisms at the level of single cells and assemblies of cells in the visual pathway, the psychologist who is interested in explaining how animals recognize and discriminate between visual patterns, and the computer scientist who is concerned with the design of systems that can analyze, synthesize, and recognize various classes of pictures. During the past decade, these three groups, working more or less independently, have accumulated large amounts of information, but the progress made by each group may not have been fully appreciated because it is difficult to retrieve and utilize information in areas with which one is not familiar. In this article we review some of the more important work on biological and computer visual systems and point out possible directions future research may take. Interaction between psychologists, physiologists, and computer scientists would seem to be necessary before much further significant progress can be made in our understanding of visual perception in animals and man, and before adequate automatons can be developed.

Automatic systems for the recognition of pictorial patterns used to aim primarily at the classification of patterns into prespecified categories by means of some variant of the following method. A set of measurements, either local, global, or both, were made on the pattern to be classified. These values were compared with those obtained from ideal prototypes of the categories, and, by appropriately weighting and combining the results of these comparisons, the pattern being classified was assigned to one or other of the categories. This unstructured classificatory approach was found to be incapable of analyzing complex objects or scenes (configurations of objects). For such situations, new methods, sometimes referred to as "descriptive" or "syntactic" (*1*) because of certain analogies with formal theories of language, become necessary.

The computer methods now used have two new features: (i) They are hierarchical, and (ii) relational structures are analyzed at each of the various levels within the hierarchy. As in some earlier methods, a computer is programmed to begin by detecting local features of a given scene; then, rather than comparing the features with an ideal prototype, it searches for relationships among features that indicate the presence of particular subpatterns. Specific relationships of subpatterns, in turn, indicate that more global patterns are present, and so on, until the desired structural description of the scene has been obtained.

The hierarchical type of scene analysis appears to be consistent with what is currently known about biological visual systems. The vertebrate visual pathway is separable anatomically into distinct levels, and the retinal and lower cortical levels have been found to contain cells responding to highly specific features similar in many respects to those used in computer programs. Later in this article we discuss some of the ideas that have been used in setting up computer programs capable of performing scene analysis and indicate the relevance to psychology of some of this work.

Physiological Feature Analysis

The brain is connected to the skin, the eye, the ear, and other sense organs by a very large number of parallel channels—the sensory nerve fibers. By detecting impulses in them, and thus intercepting the messages on their way to the brain, it was shown that each fiber could be excited by stimuli applied to a restricted part of the sensory surface, and that the fibers differ among each other in the types of stimuli to which they respond (*2*). Thus the sensory environment is mapped onto the sensory fibers, and their activity represents all the information that is available to the animal about the spatial distribution and quality of the stimuli impinging on it.

When the eye was investigated neurophysiologically, it became evident that something more than mapping occurs in the neural layers of the retina, where the image is processed and messages from the eye to the brain are initiated. In the frog, the properties of a stimulus that is selected in the laboratory to cause a vigorous response from a certain type of optic nerve fiber very much resemble the properties of a stimulus that would, under natural conditions, elicit the snapping response by which the intact animal captures its food. Thus the frog's retina contains part, at least, of the complex mechanism for recognizing edible objects in the environment

---

Dr. Barlow is professor of physiology at the University of California, Berkeley. Dr. Narasimhan is professor and head of the computer group at the Tata Institute of Fundamental Research, Bombay. Dr. Rosenfeld is professor at the Computer Science Center, University of Maryland, College Park.

Table 1. Trigger features of neurons at different anatomical locations in various vertebrate species, with sources of reference.

Anatomical location	Trigger feature
	Goldfish
Retina (*9*)	Local redness or greenness Directed movement
	Frog
Retina (*3*)	Convex edge Sustained edge Changing contrast Dimming Dark
Optic tectum (*10*)	Newness Sameness Binocularity
	Pigeon
Retina (*11*)	Directed movement Oriented edges
	Ground squirrel
Retina (*14*)	Local brightening or dimming Local blueness or greenness Directed movement
Lateral geniculate body (*14*)	Color coded units
Optic tectum (*14*)	Directional units Oriented slits or bars Complex units
	Rabbit
Retina (*12*)	Local brightening or dimming Directed movement
	Rabbit (continued)
	Fast or slow movement Edge detectors Oriented slits or bars Uniformity detectors
Lateral geniculate (*12*)	Greater directional selectivity
Tectum (*13*)	Habituating units
	Cat
Retina, lateral geniculate, cortex (*4–7*)	For main types see Fig. 1
	Infrequent types
Retina (*15*)	Directed movement Uniformity detectors
Lateral geniculate (*15*)	Local blueness or greenness Binocular, directional, and orientational units
Optic tectum (*16*)	Directed movement Complex units
	Monkey
Retina (*17*)	Local brightening or dimming Local redness, greenness, or blueness
Lateral geniculate (*18*)	Various forms of color coding
Cortex (*19*)	Similar to cat; some color coded
Inferotemporal cortex (*20*)	Very complex; possible hand detector

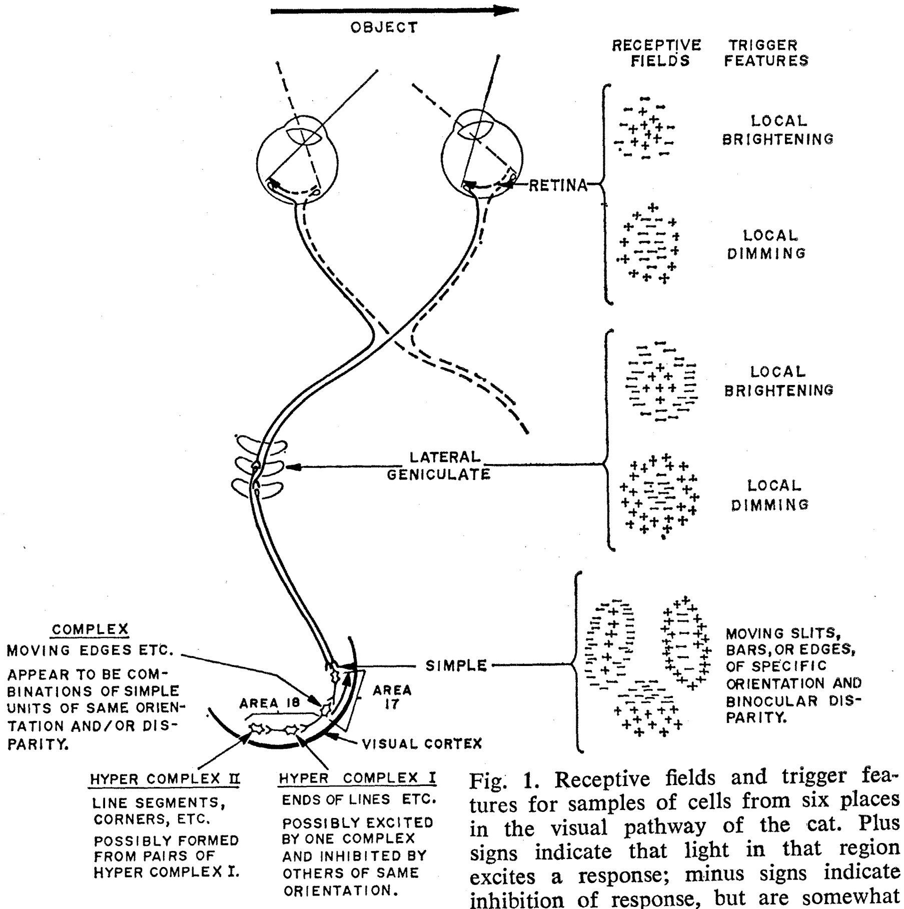


Fig. 1. Receptive fields and trigger features for samples of cells from six places in the visual pathway of the cat. Plus signs indicate that light in that region excites a response; minus signs indicate inhibition of response, but are somewhat ambiguous. They may mean that light in that region prevents responses that would otherwise occur as a result of excitation elsewhere in the same receptive field. This diagram is also simplified in other ways, for movement and disparity (see text) are often specific requirements, and there are some rarer types of units (see Table 1). The visual cortex is divided anatomically into regions, areas 17 and 18. [Adapted from Kuffler (*4*) and Hubel and Wiesel (*5*)]

(*3*), and it does not relay a simple copy of the visual image to the brain.

The type of information processing that is done can be inferred from the type of stimulus that is required to activate a sensory neuron. Many species have been studied in this way, but the cat is a good one to consider in more detail because a succession of feature-filtering operations can now be described. Figure 1 shows diagrammatically the types of receptive fields of samples of cells in the retina, lateral geniculate body, and visual cortex. Those found in the cortex are thought to belong to successive elements of a hierarchy; the simple type is followed by complex, low-order hypercomplex, then high-order hypercomplex. If this is correct one can sample neurons at six successive levels of processing (*4–6*).

Most of the cortical neurons can be activated through both eyes, and the image of objects in the real world may be so positioned on each retina that the two pathways reinforce each other and thus cause a more vigorous response, or they may be so positioned that the response is diminished by mutual inhibition. The positioning must be very precise in both retinas for reinforcement to occur, so that the stimulus must lie at a particular distance from the cat, as well as in the appropriate direction in the visual field. Furthermore, the precise relative positioning is found to vary for different cells, so that the binocular parallax of objects at different distances leads to excitation of different neurons with different disparity specificity (*7*). Information about depth must surely be extracted from the paired retinal images in this way, though it can only be a first step toward the estimation of the distances of objects in the visual field.

Looking at the progression of trigger features at the successive stages of processing, one sees that a more and more specific spatiotemporal pattern of visual excitation is required, but, on the other hand, the requirements become less specific in other respects. In the retina, the stimulus light must be in exactly the right place and must be either lighter or darker than the background (depending upon whether it is an on- or off-center neuron), but the response is not greatly influenced by varying the luminance over a wide range, provided that the contrast is unchanged. The positioning must also be precise to excite a simple cortical cell with a bar, slit, or edge; but, for the higher order cells, positioning is less critical provided that the appropriate

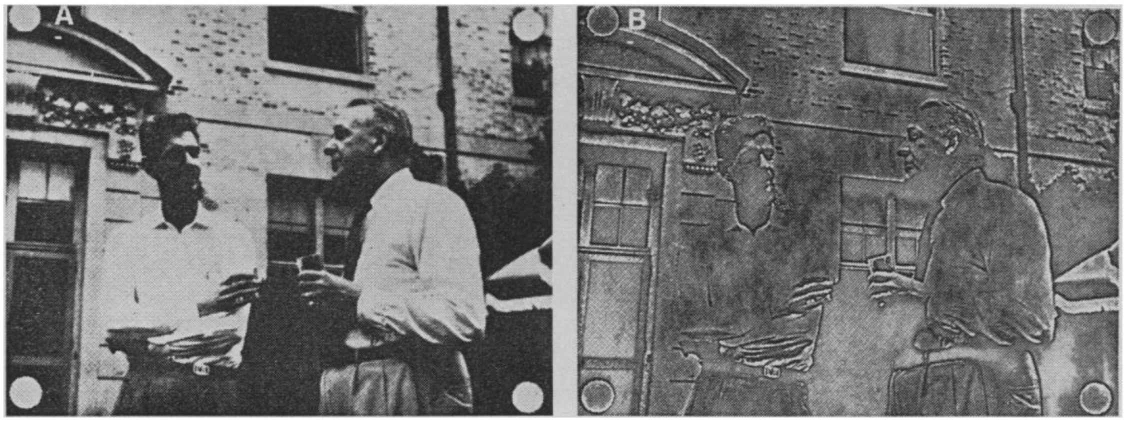


Fig. 2. Example of unsharp masking. (Left) Original scene. (Right) The same scene after superimposing a defocused negative. The person on the left is F. Attneave, who first pointed out the relation between Gestalt principles and redundancy reduction (*34*). [Courtesy of M.I.T. Press (*26*)]

specific characteristic of the pattern is present. For the complex cells, the specific characteristic is orientation; for hypercomplex I, it is orientation and the termination of a line at one or both of its ends; and for one type of hypercomplex II, a line must have a certain length and orientation, but its position is immaterial over a large range. Thus it appears that the cells at successive stages of processing develop increasing specificity for some aspects of the trigger features, while at the same time reduced specificity, or increasing generality, is allowed for other aspects. We do not, however, understand at all clearly the exact functional role of these operations, and it is on questions like this that computer scientists may make valuable suggestions.

In Table 1, we list some of the various feature detectors that have been described in vertebrate visual systems. The trigger features are described by a simple name or phrase, but these should be viewed with caution for three reasons. First, the appropriateness of some of the names is disputed even in the case of retinal units; for instance, "convexity" detectors may simply be a class of edge detectors with very active surround suppression (*8*). Second, it is difficult both to establish and to confirm claims about the specific response requirements of a class of units which is only rarely found, so that the opportunity to repeat the observations rarely arises. Third, the verbal phrases used for trigger features fail to describe many of their properties and add teleological overtones to physiological evidence. It is essential to refer to the original reports referred to in Table 1 in order to understand what was actually observed (*9–20*).

Some work has been done on the mechanisms whereby the specificity of a feature detector is achieved. While the detectors of simple brightening and dimming may work by linear addition of excitatory and inhibitory influences from different parts of their receptive fields (*21*), there is evidence that inhibition plays a rather different role where the trigger feature is more complex. Instead of achieving specificity by detecting the logical conjunction of pattern features, the system often seems to work by an inhibitory "veto" (*22*) that prevents responses to stimuli which do not have the required specific feature. Notice the remarkable complexity that can be achieved at an early stage in the visual pathway. For example, the rabbit has orientationally selective units in the retina; in the cat, such a degree of specificity is not reached until the cortex, but this may be related to the requirement for the preservation of exact positional information for use by the binocular stereopsis mechanism.

Local Feature Analysis by Computer

Redundancy reduction is of basic importance in the design of information processing systems that can perceive and interact with the external world. Many of the properties of sensory pathways in animals, such as adaptation, lateral inhibition, and selective sensitivity to direction of motion, can also be understood as measures to reduce the redundancy of representation of the incoming signals (*23*). The central idea is to emphasize singularities or discontinuities for higher level processing by taking advantage of regularities in the

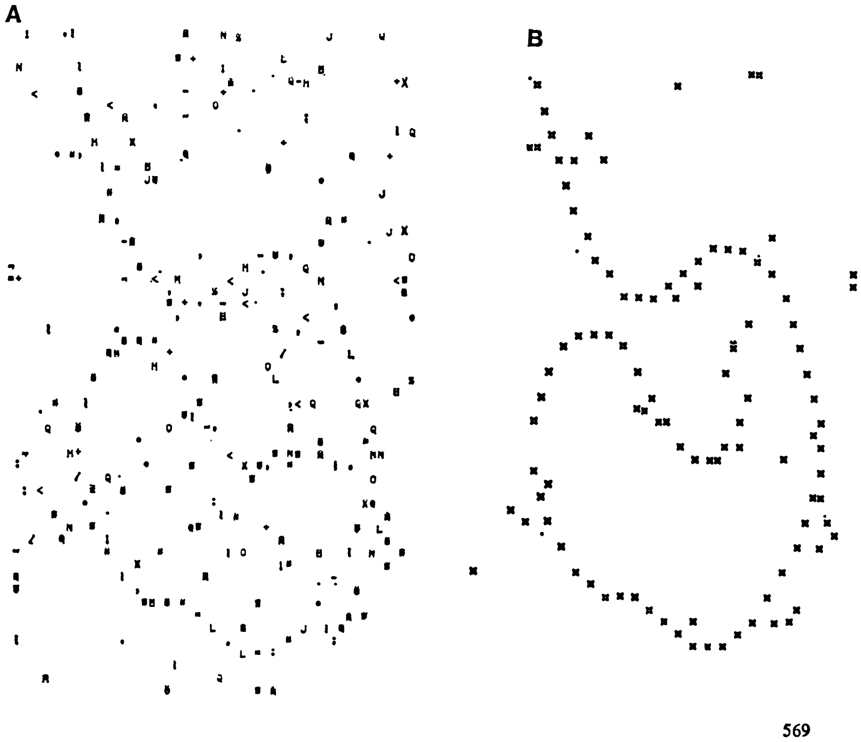
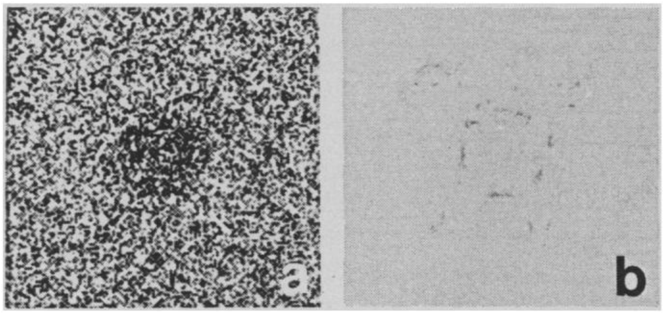


Fig. 3. Example of curve detection in noise by means of parallel local operations. (A) A dotted curve was concealed in a noisy background which was formed by selecting at random 6 percent of the points and assigning to each a randomly selected gray level (luminance). (B) Result of applying a line detection operation to (A) (*28*).



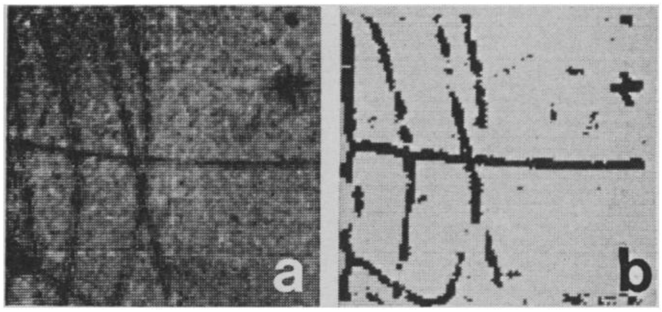


Fig. 4 (left). Horizontal and vertical edge detection for a probabilistically generated picture (probabilities of a black dot inside and outside the circle are .6 and .4 respectively). (a) Original. (b) Results of edge detection operation (*16*). Fig. 5 (right). Line detection does not work well on faint, broad streaks, such as those in (a), which is a portion of a photograph of a bubble chamber. (b) Results of streak detection operation in which the line detection operations were performed on patches rather than points (*30*).

signal at the lower stages of processing.

A variety of computer processing techniques has been developed for reducing the redundancy of images (*24*). Some of these will be considered here, and many physiological analogs in the animal visual systems will become evident. A technique that has long been used by photographers to obtain "outline drawing" effects from continuous-tone photographs is called *unsharp masking*. This consists in superimposing, in register, an in-focus positive and an out-of-focus negative of the same picture. An example of a picture processed in this way is shown in Fig. 2. This technique has also been used in a number of computer image processing systems (*25, 26*). In computational terms, subtracting a blurred image from a sharp one is equivalent to subtracting a weighted sum of the gray levels (reflectance values, or luminances) surrounding each point in the picture from the gray level at that point itself. This resembles the type of computation attributed earlier to the simplest detectors of local brightening and dimming in the visual system. The operation of such cells in the human has also been proposed as an explanation of "Mach bands," the false bright and dark streaks that appear where there are changes of luminance gradients (*27*).

Line-like features in a picture can be characterized as a succession of points, aligned in a given direction, which are brighter (or darker) than the points on either side of them in the orthogonal direction. This characterization can be used to define computer algorithms for local line detection in various orientations. Since any curve is locally straight-line-like, such a line detection operation can be used to extract arbitrary curves, for example, object boundaries, from a noisy background. An example of such a curve filtering is shown in Fig. 3 (*28*).

Apart from such "fine" operations as those described above for spot and thin-line detection, "coarse" operations of the same type are useful for detecting grosser features in a picture, for example, edges at which average gray levels change abruptly, and spots or streaks having higher or lower average gray level than their surrounds. A streak or coarse line detector operates on patches rather than points, and detects successions of patches aligned in a given direction that are brighter or darker than the patches on either side of them in the orthogonal direction. In the visual system there is both neurophysiological and psychophysical evidence (*29*) for channels with different spatial summating and resolving properties conveying information in parallel. This is presumably related to the finding in computer processing that it is advantageous to detect and operate with patches as well as points. Furthermore, by comparing the outputs of coarse and fine detection operations in the same regions, one can detect coarse features quite distinctly even in the presence of considerable noise. Figures 4 and 5 show some examples of the use of this technique (*30*).

The detection of local features can be further extended by applying the coarse operations to the outputs of the finer operations. In particular, one can detect a variety of "texture edges" (*30*) by allowing coarse edge detectors to operate on the outputs of fine edge, spot or line detectors. Related to this is the idea (*28*) that the noise cleaning effectiveness of a line detection operation is increased by iterating the operation (see Fig. 6). It may thus be useful to consider whether a neurophysiological feature detector can select and process alternative inputs, or if it can process the same input iteratively.

The operations described above are to a large extent concerned with rescuing the images of contours and edges from the ubiquitous noise that tends to submerge them. This proves to be a more difficult task than was anticipated, and it would not be unreasonable to suppose that a unit such as a "corner detector" (see Fig. 1) at the sixth stage of visual information processing is simply performing a reliable detection of an important element in the animal's visual space. But a corner is a relation between two simpler elements, and the problem of handling relations becomes important in computer programs capable of analyzing complex scenes.

Relational Structure Analysis by Computers

Up to a certain point, one can detect increasingly global features simply by using larger detectors. For example, to detect a square in a given range of sizes and orientations one could simply require the presence of four right-angle corners in approximately specified relative positions, and no corners or line terminations elsewhere. In a sense, this is analogous to a "corner detector" which requires the co-occurrence of two lines in the proper relative orientations. But there is a significant difference between the two types of detectors. In the case of the corner detector, the objects (that is, line elements) and their relationships are determined by strictly local computations (that is, computations limited to a local neighborhood of the two objects). In the case of a "square detector," the relationships may have to be computed over large regions. It would be very costly in terms of computer time to implement such operations for the wide variety of shapes, sizes, and orientations, that would be encountered in practice, and it seems unlikely that the brain would use such a system. It is more reasonable to assume that, with more complex features, spatial relations are encoded in a more abstract way, so that spatial arrangements can be recognized without operations having to be applied at all possible positions, orientations, and scales.

Figure 7 shows the strategy used in a computer program designed to recognize hand-printed English letters (*31*). There is clearly some analogy between these operations and those performed in the cat cortex (Fig. 1). In the cat, a specific element is detected, then the positional requirement is relaxed; in the computer program, the original representation in a 25 by 25 grid is reduced to a representation of elementary features in a 3 by 3 grid. There has also been redundancy reduction, for it is the presence of ordered structure in the original input that allows the more economical representation after analysis.

Global information about a visual scene can be inferred from local cues in a variety of ways. An interesting computer program provided by Guzman (*32*) analyzes line drawings of scenes containing polyhedra. This program infers, with quite high reliability, which groups of regions in a scene, such as that shown in Fig. 9, are the faces of a single body. The inferences are based on the properties and implications of vertices of the types shown in Fig. 8, each vertex being classified on the basis of the numbers and orientations of the lines that meet there. Of course, we can do at a glance what this computer program achieves; the interesting question is whether our neurophysiological mechanisms perform the same computational steps that Guzman's programs use, and if so, where and how they are performed.

Psychology of Perception

Scientists who investigate the neurophysiological activities of cortical neurons or design computer programs for scene analysis are obviously concerned with the same mechanisms that have been discussed extensively by Gestalt psychologists (*33*). Gestalt psychology attempted to define the organizational principles by which the global or general aspects of a visual scene were abstracted from the details. Attneave (*34*) pointed out the redundancy-reducing potentialities of these principles, and we indicated that feature detectors can be thought of as redundancy-reducing devices; they tend to disregard uniformities and accentuate discontinuities or singularities (for example, contours and corners). Analogous organizational principles may be thought of as detecting axes of symmetry and assymetry, periodicity, and the grouping or clustering of features in appropriate feature spaces, though the neurophysiological mechanism for such operations is unknown.

Various suggestions have been made as to how computer programs might be designed to simulate some of the Gestalt organizational principles (*35*), but very little systematic work has been done by computer scientists to determine an efficient set of relationships for the comprehensive analysis of pictures that work well in a variety of circumstances. It seems evident that such a set of relationships should include at least the following: right of, left of, above, below, inside, outside, at the center of, surrounded by, near, far, next to, attached to, overlapping (occluding), isolated (one), grouped (many), larger than, smaller than, longer than, shorter than, more than, less than; and perhaps more.

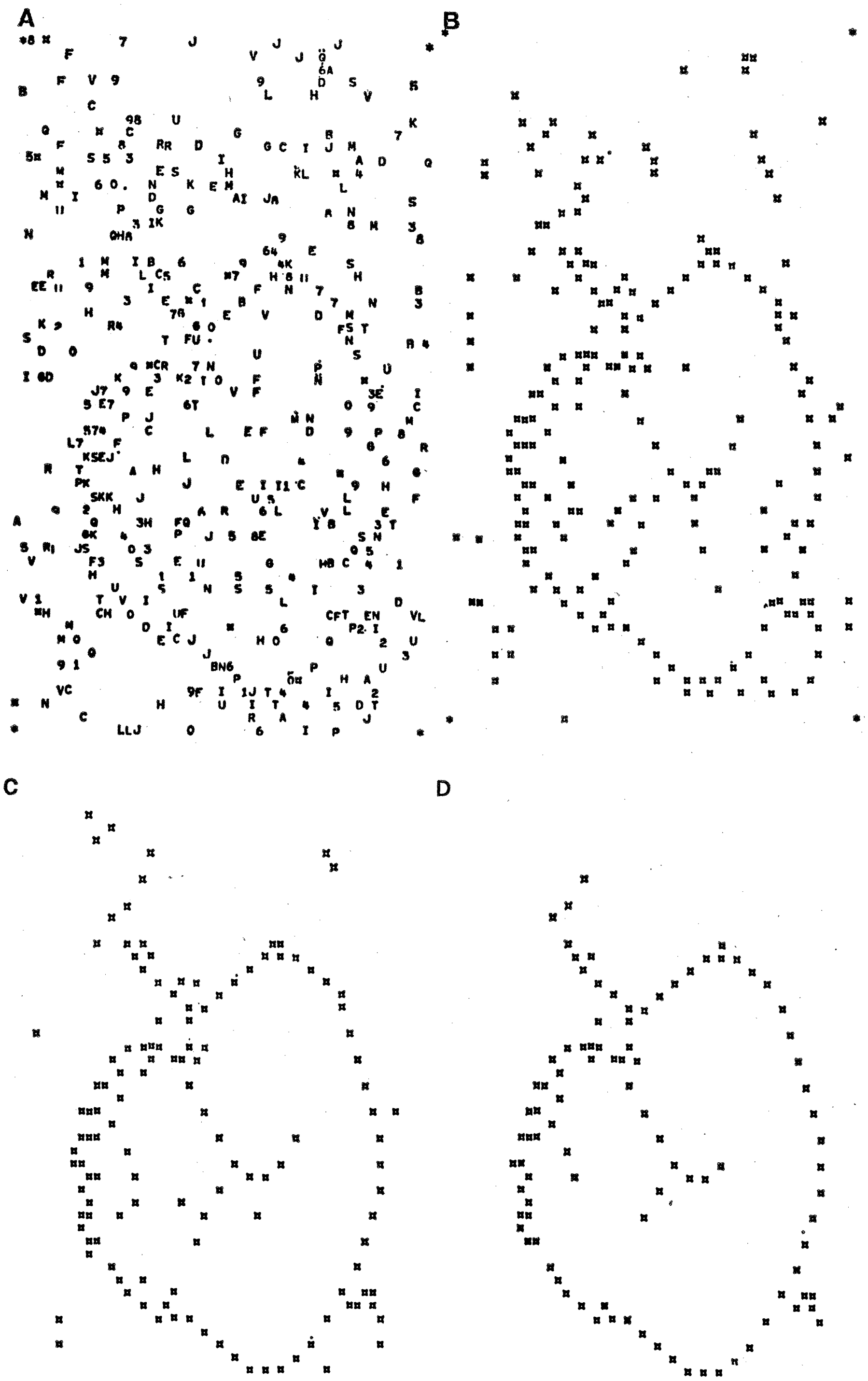


Fig. 6. Results of iterating a feature detecting operation. (A) Dotted curve in 12 percent noise (B–D). Results of applying a line detection operation to A, then B, then C (*14*). The characters are as in Fig. 3.

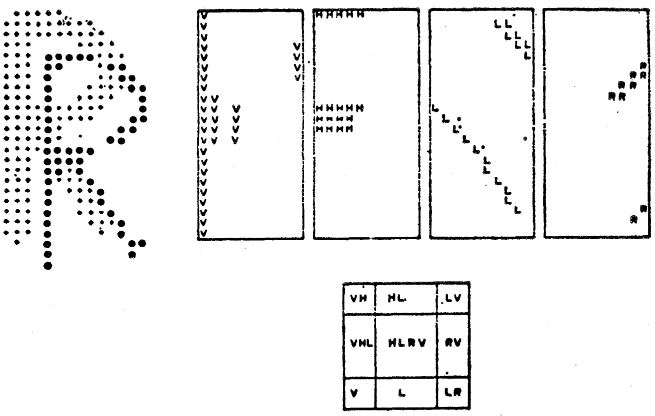


Fig. 7. Analysis of hand-printed English letters. The successive operations on the 25 by 25 binary input grid are thinning, oriented line detection, and representation in a coarser (3 by 3) matrix upon which logical operations are performed (*31*). At the left is the input image and its "skeleton"—that is, a symmetrically thinned version. At the right are the line element separated out along four principal directions: vertical, horizontal, right and left diagonal. At the bottom is an abstracted image of the visual field divided into nine gross regions. The occurrences of various global patterns can be inferred by looking for appropriate logical combinations of the labels in the 3 by 3 matrix.

Little is known about the neurophysiological mechanisms for detecting such comprehensive relationships. The organization of the visual system hints that each relational feature is first detected by many separate units on a local level, and is then generalized by combining the outputs of these separate units. This would require very extensive reduplication of the "relation detectors," but eye or head movements might lessen this requirement. The extraordinarily disruptive consequences of dissociating or disturbing the relation between self-initiated movements and visual experience (*36*) suggests that spatial relations in the visual field may be encoded in terms of "motor schemata," the movement associated with producing or removing a spatial relation (*37*).

Experiments on the visual discrimination of animals have provided much information on their ability to "generalize" from a specific training experience. Sutherland (*38*) has compiled a useful summary of these results and has proposed a set of conditions that any theory of visual perception should satisfy. However, it is hard to tell from such studies whether animals perform true scene analysis. Their laboratory-proven discrimination capabilities could be based solely on cluster analysis of first level features abstracted by their feature detectors, whereas observation of animal behavior in natural environments often suggests that they are able to perform competent and comprehensive scene analysis. These aspects of sensory performance are not well tested in current laboratory studies.

Possible Developments

In a rapidly developing subject even the most confident predictions are liable to be falsified, but we now try to bring together some of the suggestions that have emerged in the preceding discussions. Experience with computer programs for pattern recognition and scene analysis over the last 12 years has proved that these complex tasks are by no means easy to understand, even though our brains perform them so readily. The efforts that have led to successful solutions tell us what the natural hazards and difficulties are, and this new source of ideas suggests new questions. For instance, are there physiological mechanisms for detecting paired T-junctions and inferring the continuity of a partially occluded edge? Are vertices classified physiologically according to a scheme such as that used in Guzman's computer program? If not, by what means does the human brain perform the tasks that these programs perform?

Many of the feature detectors required for letter recognition by computer (Fig. 7) are already known to exist in nature, but perhaps some caution is needed before concluding that their functional roles are the same. For instance, a feature detector in an animal may be used for some very specific purpose, as in the "fly detector" of the frog's retina (*3*). Again, directionally selective units in the rabbit's retina may be the sensors of specific reflexes that adjust the contractions of the extraocular muscles, thus reducing the movement of the image over the retina (*39*). Even detectors for oriented lines or edges in the cat cortex may have the primary role of identifying the same part in the paired images of an external object in the two eyes, in order that its binocular parallax may be abstracted. The feature detector would then be acting as a means to link the two separate images of the same object; such linking devices play an important role in identifying the various parts of a single object in a complex scene [see Fig. 9 and (*32*)]. These examples show that one should not assume too hastily that feature detectors perform identical roles in animals and computer programs, because the environments for which the systems have been evolved may be very different.

Feature analysis by computer is often thought to depend on fixed, unmodifiable routines, but there is usually some degree of adaptability in a workable system. This may take a simple form such as an automatic gain control to adjust the intensity levels of the signals, and the mechanisms of light and dark adaptation in biological vision are sometimes regarded in similar terms. It has been shown, however, that these biological mechanisms are not quite so simple. For instance, the retinal units only develop their concentric, antagonistic, surrounds when the illumination level increases (*40*), and this gives the retina the property of emphasizing high spatial and temporal frequencies at higher light intensities (*41*). This may be

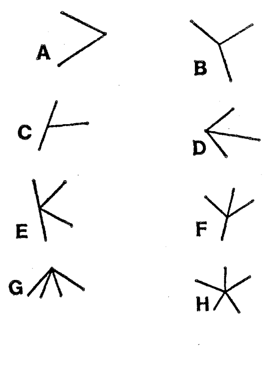


Fig. 8. Classification of vertices by means of a computer program that analyzes scenes containing polyhedra. The results are used to draw such inferences as the presence of hidden faces (*Arrow*) and occluded edges (*T*). (A) *L* form, vertex where two lines meet; (B) *fork*, three lines forming angles smaller than 180°; (C) *T* form, three concurrent lines, two of them collinear; (D) *arrow*, three lines meeting at a point, with one of the angles bigger than 180°; (E) *K* form, two of the lines are collinear, and the other two fall on the same side of such lines; (F) *X* form, two of the lines are collinear, and the other two fall on opposite sides of these lines; (G) *peak*, formed by four or more lines, when there is an angle bigger than 180°; (H) *multi*, vertices formed by four or more lines, and not falling into any of the preceding classes. [See (*32*)]

no more than a minor adjustment or a preprocessing trick, but the oriented line, slit, and edge detectors in the cat cortex also have a certain amount of plasticity. At first it was thought that the whole visual mechanism was established by developmental processes under strict genetic control, though it was shown that the mechanism could be disrupted by abnormal types of visual experience at a critical time (*42*). However, it now appears that the genetic determination is incomplete; the exact details of the specific response requirements of cortical neurons also depend upon what the kitten has seen (*43*). Binocular connections exist at birth, but the neurons have poor specificity for disparity and orientation, and the high degree of adult selectivity only develops with visual experience. Kittens which are prevented from using both eyes together lose the binocular connections to their cortical neurons (*42*). Preliminary reports show that, if kittens are exposed only to vertical stripes, they have no horizontal line detectors, and if they are exposed only to horizontal stripes they have no vertical line detectors (*43*). Furthermore, kittens raised with displaced vertical alignment of the images in their two eyes have cortical neurons connecting to correspondingly displaced retinal regions (*43*). Thus it seems that cortical neurons bias the selective sensitivity they develop in favor of the patterns of visual experience they have been exposed to, and against patterns outside their experience.

It is clear that our ideas about the nature of the genetic instructions to a cortical cell require a major revision. A cortical cell is not instructed to receive excitatory connections from a subset E of the axons growing around it and inhibitory connections from a subset I, or to perform some equally specific, inflexible procedure. Instead, the all-important conditional instructions that allow branching are included in the developmental program. The cell might thus be instructed to establish excitatory connections from a member of a subset E only if it is active in the absence of inhibition by any member of a subset I, or to implement some other rule for conditional connectivity. The genetic instructions may therefore resemble a computer program more than a wiring diagram. Genetic software of this sort may greatly reduce the amount of specific information about connections that is required to set up a functioning visual system, and make the system work in a greater variety of environments. It might be worth while to devise computer programs to explore the viability of such heuristics in achieving experience-dependent connectivities.

Coding to reduce redundancy is a theme that has recurred in describing physiological feature filters (*23*), feature filtering for pattern recognition programs (*24*), Gestalt organizational principles in psychology (*34*), and relational methods in scene analysis (*1*). The plasticity of the developing visual system also fits this paradigm; an individual's coding of his environment appears to be permanently modified in accordance with uniform aspects of his early visual experience. This occurs in the primary visual cortex, and one must suppose it also occurs in secondary visual cortex and beyond. Thus the result gives a little physiological support to the notion that redundancy reduction is important in the much higher level mental functions involved in solving problems that test for "intelligence," and in the formation of intellectual "concepts" (*23*). But there are difficulties in elevating redundancy reduction to a prime position in organizing sensory information. First, though it seems adequately descriptive both of feature filtering methods that have been found in natural systems, and of methods found useful in computer analysis, it is not sufficiently prescriptive to permit tests of its validity. Second, from a theoretical point of view, very nonredundant codes may be inconvenient to use because one usually has to look at the whole message before the meaning of any part is clear. Redundancy reduction appears not to be the only criterion; some other less easily defined criterion of convenience, or pragmatic adequacy, is required.

In spite of recent advances, it must be admitted that current achievements in pattern analysis by computer hardly begin to match those of human perception, and one wonders what gives the slow biological system its vast superiority. Computer scientists have so far devoted themselves to the empirical task of finding good sets of feature detectors, and algorithms for utilizing them to perform specific recognition tasks. Although success has so far come through this empirical approach, interesting theoretical questions now arise: What advantage is conferred by the plasticity of biological feature detectors? Can the "goodness" of a set of feature detectors be defined in terms of intermediate goals, such as reduction of redundancy, or effectiveness in recognition despite incomplete processing, or is success in

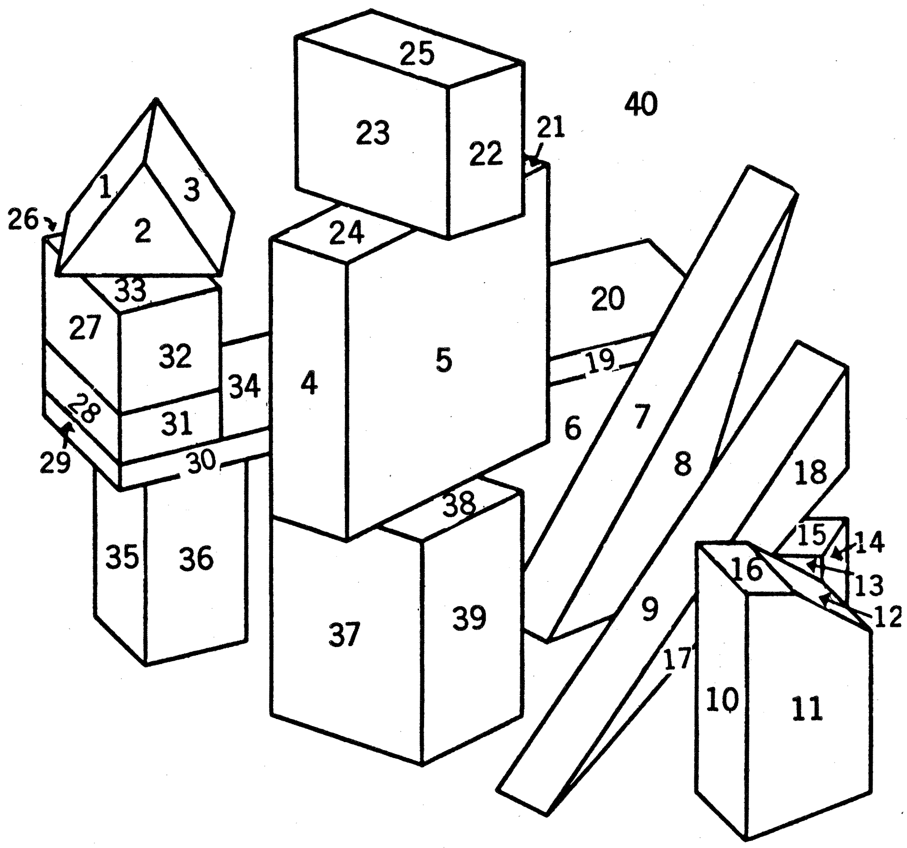


Fig. 9. Example of a scene analyzed by the program. Regions belonging to each body were correctly grouped together; the groups were (1,2,3), (4,5,21,24), (7,8), (9,17,18), (10,11,12,16), (13,14,15), (19,20,29,30,34), (22,23,25), (26,27,32,33), (28,31), (35,36), and (37,38,39). [See (*32*)]

the overall task the only valid criterion? How densely, or at what interval in the appropriate message space, must feature detectors sample the input—in other words, how many of them are required? A clearer understanding of such theoretical problems might narrow the gap between human perception and the computer's ability to handle pictures.

Conclusions

Neurophysiologists, psychologists, and computer scientists are all concerned with the problem of analyzing pictorial information. We have shown that there are similarities both in the details and in the general principles that underlie the operation of single neurons, computer programs, and perception. Among the obvious detailed similarities are the importance of detecting edges, lines, and the terminations or junctions of contours, in all types of pictorial analysis. A less obvious similarity is the empirical demonstration that, in order to overcome noise problems, a computer program must perform both coarse and fine feature detection; there is neurophysiological and psychophysical evidence for biological feature detectors having differing size specificities, and these presumably fulfill analogous roles. Because of such analogies, the detailed methods used by computer scientists to analyze a picture such as that shown in Fig. 9 are of obvious interest to those who wish to know how the human brain so readily interprets such scenes.

One general principle seems to be applicable at all stages of pictorial analysis, both in biological and computer systems: redundancy in the pictorial input is detected and utilized to create a simplified representation. There is also general agreement that in both systems the processing is hierarchical. A question of increasing interest is the appropriate role and relative importance of inflexible or genetically fixed mechanisms compared with others that are adaptive or environmentally modifiable. In natural systems the genetic instructions may not be so inflexible as was thought previously and may include conditional instructions that allow branching of the developmental programs.

We conclude that the difficulties encountered in pattern analysis by computer, and the means of overcoming these difficulties, may be of great interest to psychologists and neurophysiologists. Biological visual systems are the best general pattern analyzers known, and we therefore suggest that psychological and physiological investigations of natural systems may reveal points of interest to computer scientists.

References and Notes

1. Picture processing based on feature extraction has a long history. See G. P. Dinneen, *Proc. West. Joint Comput. Conf.* **7**, 94 (1955) for an early account. The use of syntax in the context of pictures is a more recent development. Early advocates of this approach were: R. L. Grimsdale, F. H. Sumner, C. J. Tunis, T. Kilburn, *Proc. Inst. Elec. Eng.* **106-B**, 210 (1959); M. Minsky, *IEEE* (*Inst. Elec. Electron. Eng.*) *Proc.* **49**, 8 (1961); M. Eden, *AMS Appl. Math. Symp.* **12**, 83 (1961); R. A. Kirsch, *IEEE* (*Inst. Elec. Electron. Eng.*) *Trans.* **EC-13**, 363 (1964); R. Narasimhan, *Inform. Control* **7**, 151 (1964). See also, R. Narasimhan, "Picture languages," in *Picture Language Machines*, S. Kaneff, Ed. (Academic Press, New York, 1970), p. 1.
2. See any modern physiology textbook for more details. Our description may not apply accurately to the chemical senses [see J. R. Gandrow and R. P. Erickson, *J. Neurophysiol.* **33**, 768 (1970)].
3. The frog retina was first investigated by H. K. Hartline, *Amer. J. Physiol.* **121**, 400 (1938); *ibid.* **130**, 690 (1940); *ibid.* p. 700. The relation between the specific requirements for excitation of its retinal ganglion cells and elicitation of the feeding response was suggested by H. B. Barlow, *J. Physiol. London* **119**, 69 (1953); R. Maturana, J. Y. Lettvin, W. S. McCulloch, W. H. Pitts, *J. Gen. Physiol.* **43**, suppl. 2, 129 (1960). See also D. Ingle [*Brain Behav. Evol.* **1**, 500 (1968)], O. J. Grüsser and U. Grüsser-Cornehls [*Z. Vergl. Physiol.* **59**, **1** (1968)], and J. P. Ewert and H. W. Borchers [*ibid.* **71**, 165 (1971)].
4. S. W. Kuffler, *J. Neurophysiol.* **16**, 37 (1953).
5. D. H. Hubel and T. N. Wiesel, *J. Physiol. London* **160**, 106 (1962).
6. ———, *J. Neurophysiol.* **28**, 229 (1965).
7. H. B. Barlow, C. Blakemore, J. D. Pettigrew, *J. Physiol. London* **193**, 327 (1967); J. D. Pettigrew, T. Nikara, P. O. Bishop, *Exp. Brain Res.* **6**, 391 (1968).
8. R. M. Gaze and M. Jacobson, *J. Physiol. London* **169**, 1P (1963). For a critical review of retinal receptive fields see W. R. Levick, in *Handbook of Sensory Physiology*, M. C. F. Fuortes, Ed. (Springer-Verlag, Berlin, 1972), vol. 7, part 2.
9. H. G. Wagner, E. F. MacNichol, M. L. Wolbarsht, *J. Gen. Physiol.* **43**, suppl. 2, 115 (1960); M. Jacobson and R. M. Gaze, *Quart. J. Exp. Physiol. Cog. Med. Sci.* **49**, 199 (1964); J. R. Cronly-Dillon, *Nature* **203**, 214 (1964); N. W. Daw, *J. Physiol. London* **197**, 567 (1968).
10. J. Y. Lettvin, H. R. Maturana, W. H. Pitts, W. S. McCulloch, in *Sensory Communication*, W. Rosenblith, Ed. (M.I.T. Press, Cambridge, Mass., 1961), p. 757; K. Fite, *Exp. Neurol.* **24**, 475 (1969).
11. H. R. Maturana and S. Frenk, *Science* **142**, 977 (19 → F. A. Miles, *ibid.* **170**, 992 (1970).
12. H. B. Barlow, R. M. Hill, W. R. Levick, *J. Physiol. London* **173**, 377 (1964); W. R. Levick, *ibid.* **188**, 205 (1967); W. R. Levick, C. W. Oyster, E. Takahashi, *Science* **165**, 712 (1969).

→ G. Horn and R. M. Hill, *Exp. Neurol.* **14**, 199 (1966).

14. C. R. Michael, *J. Neurophysiol.* **31**, 249 (1968); *ibid.*, p. 257; *ibid.*, p. 268; *Brain Behav. Evol.* **3**, 205 (1970).
15. W. Kozak, R. W. Rodieck, P. O. Bishop, *J. Neurophysiol.* **28**, 19 (1965); J. Stone and M. Fabian, *Science* **152**, 1277 (19 → R. W. Rodieck, *ibid.* **157**, 90 (1967); N. W. Daw and A. L. Pearlman, *ibid.* **167**, 84 (1970).
16. P. L. Marchiafava and G. Pepeu, *Experientia* **22**, 51 (1966); M. Straschill and A. Taghavy, *Exp. Brain. Res.* **3**, 353 (19 → J. T. McIlwain and P. Buser, *ibid.* **5**, 314 (1968); J. M. Sprague, P. L. Marchiafava, G. Rizzolatti, *Arch. Ital. Biol.* **106**, 169 (1968); P. Sterling and B. G. Wickelgren, *J. Neurophysiol.* **32**, 1 (1969).
17. D. H. Hubel and T. N. Wiesel, *J. Physiol. London* **154**, 572 (1960); P. Gouras, *ibid.* **192**, 747 (1967); *ibid.* **204**, 407 (1969).
18. R. L. DeValois, *J. Gen. Physiol.* **43**, suppl. 2, 115 (1960); I. Abramov and G. H. Jacobs, *J. Opt. Soc. Amer.* **56**, 966 (1966); T. N. Wiesel and D. H. Hubel, *J. Neurophysiol.* **29**, 1115 (1966).
19. D. H. Hubel and T. N. Wiesel, *J. Physiol. London* **195**, 215 (1968).
20. C. G. Cross, D. B. Bender, C. E. Rocha-Miranda, *Science* **166**, 1303 (1969); C. G. Cross, C. E. Rocha-Miranda, D. B. Bender, *J. Neurophysiol.* **35**, 96 (1972).
21. Linear combination of excitatory and inhibitory inputs was shown in some retinal ganglion cells by R. W. Rodieck, *Vision Res.* **5**, 583 (1965); C. Enroth-Cugell and J. G. Robson, *J. Physiol. London* **187**, 517 (1966). But this is not always found; see L. Cervetto, A. Fiorentini, L. Maffei, *ibid.* **204**, 121 (1969); O. J. Grüsser, D. Schaible, J. Vierkant-Glathe, *Pfluegers Arch.* **319**, 101 (1970).
22. The "veto" function of inhibition was found by H. B. Barlow and W. R. Levick, *J. Physiol. London* **178**, 477 (1965).
23. H. B. Barlow, in *Mechanization of Thought Processes* (Nat. Phys. Lab. Symp. No. 10, Her Majesty's Stationery Office, London, 1959), p. 535; H. B. Barlow, in *Information Processing in the Nervous System*, K. N. Leibovic, Ed. (Springer, New York, 1969), p. 209.
24. See the special issue on redundancy reduction, *IEEE* (*Inst. Elec. Electron. Eng.*) *Proc.* **55**, 251 (1967).
25. L. S. G. Kovasznay and H. M. Joseph, *IEEE* (*Inst. Elec. Electron. Eng.*) *Proc.* **43**, 560 (1955). For recent general reviews of image processing by computer see A. Rosenfeld, *Picture Processing by Computer* (Academic Press, New York, 1969); T. S. Huang, W. F. Schreiber, O. J. Tretiak, *IEEE* (*Inst. Elec. Electron. Eng.*) *Proc.* **59**, 1586 (1971).
26. H. B. Barlow, in *Sensory Communication*, W. A. Rosenblith, Ed. (M.I.T. Press, Cambridge, Mass., 1961), p. 782.
27. F. Ratliff, *Mach Bands: Quantitative Studies on Neural Networks in the Retina* (Holden-Day, San Francisco, 1965).
28. A. Rosenfeld, Y. H. Lee, R. B. Thomas, in *Picture Processing and Psychopictorics*, B. S. Lipkin and A. Rosenfeld, Eds. (Academic Press, New York, 1970), p. 381.
29. T. N. Wiesel, *J. Physiol. London* **153**, 583 (1960); F. W. Campbell and J. G. Robson, *ibid.* **197**, 551 (1968); F. W. Campbell, G. F. Cooper, C. Enroth-Cugell, *ibid.* **203**, 223 (1969); C. Blakemore and F. W. Campbell, *ibid.*, p. 237.
30. A. Rosenfeld and M. Thurston, *IEEE* (*Inst. Elec. Electron. Eng.*) *Trans.* **C-20**, 562 (1971).
31. R. Narasimhan and V. S. N. Reddy, *Pattern Recog.* **3**, 345 (1971).
32. A. Guzman, *Proc. Fall Joint Comput. Conf.* (1968), p. 291. Helmholtz appreciated the implication from a T-junction that the object with the continuous border lies in front of the object with the discontinuous, interrupted border. See discussion → P. Ratoosh, *Proc. Nat. Acad. Sci. U.S.A.* **35**, 257 (1949).

→ M. Wertheimer, *Psychol. Forschungsber.* **4**, 301 (1923). For a translation of this paper and for later work see D. L. Beardslee and M. Wertheimer, Eds., *Readings in Perception* (Van Nostrand, New York, 1958); W. Kohler, *Gestalt Psychology* (Mentor, New York, 1959); L. Zusne, *Visual Perception of Form* (Academic Press, New York, 1970).

→ F. Attneave, *Psychol. Rev.* **61**, 183 (1954).

35. R. Narasimhan, *Digital Comput. Lab. Rep. No. 142* (University of Illinois, July 1963); see also (*30*).
36. R. Held, *Sci. Amer.* **213**, 84 (1965).
37. The importance of eye movements in visual perception is discussed in D. O: Hebb, *Organization of Behaviour* (Wiley, New York, 1949). That motor schemata play a central role in speech recognition has also been hypothesized; for papers, see W. Wathen-Dunn, Ed., *Models for the Perception of Speech and Visual Form* (M.I.T. Press, Cambridge, Mass., 1967).

→ N. S. Sutherland, *Proc. Roy. Soc. Ser. B* **171**, 297 (1968). See also, R. M. Gilbert and N. S. Sutherland, Eds., *Animal Discrimination and Learning* (Academic Press, New York, 1969).

39. C. W. Oyster and H. B. Barlow, *Science* **155**, 841 (1967).
40. H. B. Barlow, R. FitzHugh, S. W. Kuffler, *J. Physiol. London* **137**, 388 (1957); V. D. Glezer, *Vision Res.* **5**, 497 (1965).

41. H. B. Barlow, in *Handbook of Sensory Physiology: Visual Psychophysics,* D. J. Hurvich and L. M. Hurvich, Eds. (Springer-Verlag, Berlin, 1972), vol. 7, part 4, page 1.
42. D. H. Hubel and T. N. Wiesel, *J. Neurophysiol.* **26**, 994 (1963); T. N. Wiesel and D. H. Hubel, *ibid.* **28**, 1029 (1965); D. H. Hubel and T. N. Wiesel, *ibid.*, p. 1041; *J. Physiol. London* **206**, 419 (1970).
43. H. V. B. Hirsch and D. N. Spinelli, *Science* **168**, 869 (1970); C. Blakemore and G. F. Cooper, *Nature* **228**, 477 (1970); R. Shlaer, *Science* **173**, 638 (1971); H. B. Barlow and J. D. Pettigrew, *J. Physiol. London* **218**, 98P (1971).
44. This article is the outcome of the Seminar on Visual Mechanisms and Form Perception organized by the Tata Institute of Fundamental Research, Colaba, Bombay, India, 25 January to 6 February 1971. H.B.B. was supported by PHS grant EY00276; A.R. was supported by Office of Naval Research contract NONR-5144(00).

The Non symbiotic Origin of Mitochondria

The question of the origin of the eucaryotic cell and its organelles is reexamined.

Rudolf A. Raff and Henry R. Mahler

The greatest evolutionary discontinuity between living organisms is that separating procaryotic from eucaryotic cells. While there is ample biochemical evidence demonstrating that these two classes did not arise independently, the fundamental differences in their basic organization has made it extremely difficult to reconstruct their evolutionary history.

One of the most puzzling features of eucaryotic cellular organization is the existence of semiautonomous cytoplasmic genomes in such organelles as mitochondria and chloroplasts. The presence of these self-replicating organellar genomes, and the resemblance of the associated organellar systems of protein synthesis to bacterial systems has led to the wide acceptance of a theory originally propounded in the late 19th century that these organelles had their origin in a symbiotic association of bacteria and blue-green algae with the ancestral eucaryotic cells (*1–5*). This theory requires that various organelles were actually generated in several symbiotic events (*3, 4, 6*).

Since this view has gained wide popularity, we chose to reexamine the data used in its support with respect to the origin of mitochondria.

In our opinion there is no a priori reason why the eucaryotic cell, which has proved capable of remarkable evolutionary innovations, should have originated as a collage of procaryotic cells and parts of cells rather than have evolved in a more direct manner from a particularly advanced type of procaryotic cell. While symbiosis may have been of some evolutionary significance, overdependence on it as an explanation for the origin of the eucaryotic cell and its organelles may leave interesting questions unasked. Furthermore, dogmatic adherence to this theory leads to such improbabilities as the postulation of the origin of cilia from symbiotic spirochaetes (*1, 3, 4*), implying a nonexistent homology between flagellin and microtubule protein (*7*), and the multiple origin of chloroplasts from three entirely separate groups of photosynthetic procaryotes (two of them hypothetical) (*6*).

We shall show that while the symbiotic theory may be esthetically pleasing, it is not compelling, and we will propose an alternate hypothesis for the origin of the eucaryotic cell (*8, 9*).

Outline of the Symbiotic Theory

The current symbiotic theory for the origin of the eucaryotic cell and its mitochondria is succinctly discussed by Stanier (*5*).

By this theory, as the primitive earth atmosphere began to change from anaerobic to aerobic as the result of photosynthetic oxygen production, procaryotes which had utilized a wide variety of anaerobic metabolic pathways were forced either to adapt to aerobic conditions or to become restricted to the few anaerobic environments remaining. Since eucaryotes are restricted to glycolysis for their anaerobic energy supplies, the ancestral protoeucaryote likewise utilized glycolysis. This protoeucaryote, by various adaptations, escaped from the selective pressure of free oxygen, the determinant driving the evolution of advanced oxidative metabolic pathways in other contemporaneous procaryotes. By evolution of larger cell size, intracellular translocation, advanced mechanisms for motility, and the ability to phagocytize, the protoeucaryote became able to ingest procaryotes as prey to provide substrates for glycolysis. Related and subsequent to these advances was the establishment of stable intracellular symbiotic relationships between the protoeucaryote and certain ingested aerobic procaryotes. Such relationships exist in present-day organisms. The terminal stage in eucaryote evolution was thus the acquisition of oxygen mediation (photosynthesis and respiration) by several quantum steps.

The theory further requires that in the course of time the symbiotic association has become extremely intimate. Most of the genetic information required for assembly of the organelle-symbiont has been transferred to the nuclear genome. The informational content of the organellar genome has been concomitantly much reduced and this genome as well as the organellar protein synthesis systems are evolutionary relicts.

This hypothesis has two particularly awkward aspects. The first is that the postulated protoeucaryote possessing many advanced cellular adaptations should have been so primitive and inefficient metabolically. In the face of competition from conventional procaryotes possessing more efficient aerobic, energy-yielding pathways already foreshadowing the patterns observed today, this should have left it at a considerable disadvantage. Second, the integration of the endosymbiont-proto-

---

Dr. Raff is assistant professor of zoology and Dr. Mahler is research professor of chemistry at Indiana University, Bloomington 47401.