

**Toward a Computational Information-
Processing Model of Object Perception**

by

Leonard Uhr

Computer Sciences Technical Report #651

July 1986

Toward an Information-Processing Computational Model of Object Perception

Leonard Uhr
Computer Sciences Dept.
University of Wisconsin, Madison 53706

Abstract: Integrating What Is Known Into a Running Model

This paper describes the first version of an information-processing model (to be programmed for an appropriately structured parallel-serial multi-computer network) for the visual recognition of complex objects. The model will be run to exhibit its behavior, and tested, evaluated, and compared with living visual systems, and with other models. It should also serve, where the brain's mechanisms are not yet known, as a test-bed to evaluate alternate possible structures.

Such an enterprise now appears to be feasible. Consider these highlights of pertinent facts:

- . The retina and primary visual cortex, with their massively parallel and to some extent serial structure of processes, detect spots, colors, families of oriented edges, and other features.
- . Much is now known about the cortex's columnar structure, and the topology of links within and between columns, hypercolumns, modules, and areas.
- . Roughly 20 cortical visual areas have been discovered. A great deal is known about the many links between them, the way they map information, and the processes each effects.
- . The recent firm evidence for neurons in the temporal lobe that respond selectively, in 70 to 200 msec, to face, different particular faces, complex parts of faces, and other complex objects, strongly suggests that these neurons are involved at a culminating stage in the complex structure of processes that perceives patterned objects. Temporal lobe lesions make complex objects like faces unrecognizable, while leaving other visual processes largely undisturbed.
- . The brain's massive parallelism makes this enormous speed possible. The relatively slow msec response times of its basic computing elements, the neurons, means that the maximum possible serial depth of processes is a few dozen to a few hundred at most.

This paper first organizes key known facts about the visual system and describes the model. (Later sections give more detail, and pinpoint important facts still unknown.) It also briefly examines this model's relation to parallel computer architectures, to other models for visual perception, and to computer vision programs, emphasizing those from which this model grew.

The goal is to develop a model/theory that exhibits the brain/mind's behavior on difficult recognition tasks, and suggests plausible mechanisms where facts are not yet uncovered, or firm. The running program should show that the model is precise and consistent, and how well the separate parts work together. It should suggest new questions and experiments. This first version incorporates only what appears essential to achieve a simple yet realistic working model. The final discussion examines possible improvements.

NOTE: This is a rough draft of a working paper. It is incomplete in several ways: Additional important facts are known (this is an extremely complex and active field, and a more extensive examination is needed than I have had time for to date). One must state and take into account the certainty of each fact. Often facts are only partial and need to be tracked down - e.g., intracolumnar links between neurons, the spread and number of links from each column to related columns in other hypercolumns. [2 questionmarks (??) flag missing facts, or raise questions.]

The model attempts to develop a best fit to the facts, guided by an underlying logical analysis of what might best be done (from the point of view of efficiency, simplicity, speed, generality). I'd greatly appreciate any comments and suggestions, especially about facts or interpretations that are missing, or can be extended or made more precise, or are wrong or misleadingly stated - and about how to improve the model, and how to make it more useful.

The Gross Behavior of the Model, and the Key Facts about Vision it Incorporates

This section presents the salient facts that will be programmed into the model. (See Table 1 for an overview.) *{Squiggly brackets indicate facts that will be simplified, or not addressed, in the first version.}* The brain is so immense that it will usually be necessary to use smaller numbers. For example, instead of 6 to 10 million cones, retinal arrays will have 64x64 to 512x512 or so (the reduction to the number of ganglia, columns, and hypercolumns will be commensurate).

Overview: The Total System and the Individual Neuron Building-Blocks

The model's overall structure will be (a necessarily simplified version of) the following:

The visual system's many millions of rod and cone sensors input information about images focused on the retina by the eye's lens. Then several layers of neurons in the retina and in the cortex's 20 or more visual areas process this information in a massively parallel, to a shallow extent serial, hierarchical, way, with many feedback loops and complex interactions. Convergence, divergence, and overlap of information are pervasive. Processes have the flavor of probabilistic threshold operations, and sub-networks or ensembles of such operations. Coordinated structures of processes (often called feature-detectors or filters) lead to individual neurons that are sensitive to simple features like different edges at different slopes. A parallel-serial structure of processes continues, achieving several levels of different types of characteristics and more complex recognitions, until complex objects like hand, face, and particular faces, are recognized.

Neurons {and possibly glial cells} do the processing. Each neuron fires into synapses to other neurons, and is fired into by other neurons. {A neuron typically has thousands of synapses from other neurons. [It will be too costly to run the model with more than a few, or a few dozen or hundred at most.]} When sufficiently excited {either chemically or electrically} across synapses the neuron will fire. {Chemical synapses are typically unidirectional. They take .3-5 msec to fire. Electrical synapses are typically bidirectional, the neurons affecting one another at essentially electronic speed. Impulses move slowly (.5-2 m/sec) through unmyelinated processes, and much faster (3-120 m/sec) along myelinated axons.} Neurons often fire most strongly to particular patterns of inputs; but almost always they will also fire (successively less) to successively less similar inputs.

The two eyes' retinas converge and diverge information, building a map of spots: Cones and rods synapse on horizontal and bipolar cells; bipolars synapse on amacrine and ganglion cells. The "on-center, off-surround" ganglia respond to bright spots surrounded by less illumination; an equal number of "off-center-on surround" are the opposite. A million ganglia carry this map to the striate cortex (also known as V1, area 17, and primary visual area) in the two hemispheres of the cerebral cortex, via one synapse at the lateral geniculate nucleus (LGN). {Most are X-cells with small receptive fields that are responsive to spots and (in monkey) color; these are probably involved with form perception. Some are larger, fast-conducting Y-cells with larger fields sensitive to motion and to changes in intensity. The rest are slow-conducting W-cells; their functions are largely unknown, although some may be responsive to both light and dark spots.} [Information is also sent over other pathways, including the superior colliculus and pulvinar, to both sub-cortical and cortical areas. These appear to be only peripherally involved with form perception, and will not be incorporated into the first model.]

The cortex is a {1 to 4.5 mm thick convoluted} sheet of 6 layers of grey matter (unmyelinated processes) arranged in structures of columns below which are the white matter (myelinated axons). Each roughly 80,000 mm² hemisphere in man contains 10-100 billion neurons. The primary path from the LGN is to Layer 4 of V1. Roughly 2/3ds of the neurons are excitatory py-

ramidals. Almost all make cortico-cortical links, as do some of the interneurons. Fewer than half of the interneurons are inhibitory. Many neurons link several mm laterally through the grey matter. *{The W cells project to Layer 6. In cat (but not primates) the LGN has major links to other cortical and sub-cortical areas.}*

*Layer 4 in each V1 column (in primates but not cat) contains cells that respond to spots (much like LGN and ganglion cells), and also simple cells sensitive to long or tiny (often colored) moving edges with a particular slope at a particular location. They link to cells in layers 2, 3, and 5, which contain complex cells that detect edges (with 0, 1, or 2 stops) with the same slope, but anywhere in a much larger region. Columns for slopes about $10-15^\circ$ apart are arranged in slabs. Information from the two eyes and the two hemispheres projects to adjacent columns, competing and merging together. There appear to be about 50 hypercolumns stretching along the horizontal axis. *{About 20 of them appear to be within the central 3° .}* Blobs that appear to be processing color are inserted periodically in the hypercolumns. About 1000 modules, each containing rows of blobs, appear to be major processing units. *{These hypercolumns and modules are sensitive to fields that, moving toward the center, are up to 40-fold smaller and overlap more.}* Columns are linked laterally to related columns, and blobs to blobs.*

More than a dozen visual areas have been identified, in several different species. Each of the macaque's 20 or so visual areas has 2 to 8 major links (usually 2-way, one ascending, one descending) to others (over 40 in toto), giving a parallel-serial structure with hierarchical depths from 3 to 6. These are organized into two major systems, for shape and color (V1-V2-V3-V4-IT [inferior temporal]; see below), and for motion and spatial relations (V1-V2-MT-MST-Parietal Lobe). Neurons' receptive fields grow successively larger. Links within an area, and from one area to another, spread regularly, in small clusters. The different areas contain different, possibly successively less detailed and more abstract, transformed images, some 2-, some 1-, some 0-dimensional. Some contain hypercolumns, blobs and modules.

Individual neurons sensitive to very complex patterned objects have been found in IT and at still higher levels of the hierarchy, where shape- and color-related processes appear to merge back together, in STS (superior temporal sulcus). Here a small percent of neurons have been found that fire selectively to very complex objects, like hand or face, or to the face of a particular person, or to particular complex sub-parts of faces. Latency is 70-200 msec, whereas cells in MT and V4 fire in 40 or 49 msec. Humans with bilateral lesions of the temporal lobe have relatively normal perception - except that they cannot recognize particular faces, or other complex objects (e.g., a particular car, piece of clothing, or bird).

*Different patterns of neuronal firing to different stimuli, each neuron participating in many, have been found using multi-electrode techniques. Many neurons in parts of the temporal lobe respond most of the time (one exception being the highly selective face neurons). This suggests that distributed processing by networks, rather than individual neurons, may be the basic recognition units. *{The model will use simple networks of this sort, but it will often use structures of transforming functions where each function is a simplification of a network.??unclear?}**

There are a number of other areas involved with visual perception; most will not be included in the model. The problem is already so big that major simplifications must be made. The system that processes information (images) sensed by the eye and sent through the LGN to V1 and through the other visual areas of the cortex, culminating in the temporal and parietal lobe areas, will be the subject of study, since it does indeed appear to be the major system for object percep-

tion. *{But it is important to keep in mind the many other parts of the larger structure, including:*

- . Binocular vision and motion are slighted, as secondary to shape and object recognition.*
- . The path to the cortex through the superior colliculus and pulvinar is primarily involved with ocular-motor control. But it may also play a role in perception, if only to serve as an alternate path of information.*
- . Still higher-level areas, apparently chiefly in the parietal and prefrontal lobes, and in the amygdala, are involved in combining information about shape and color with information about motion and spatial relations, in combining vision with other sense modalities, in linking perception with language, in merging information into the motor system, and in the motivational and emotional aspects of perception.}*

There are many variations between different species. The microcircuitry also varies between, and within, individuals. These variations might mean there are many viable solutions and that details are unimportant, or/and that different systems are optimized for different purposes.

Issues Involved in Programming the Model for the Computer

The model is posed a very difficult problem - recognizing and discriminating among different faces and other complex objects. Further, it must do this in from 20 to 200 time-steps (where a time-step is the time needed to execute a basic neuron-like threshold element process). Finally, it must do this with a structure of processors that models (albeit in simplified form) the global retinal and cortical structures, and what is known about their microcircuitry.

The first version of the program that embodies the model will contain many simplifications - otherwise it would not be possible to fit everything together and examine the behavior of the whole system. It will be coded in modules, so that the effect on the total system of varying individual procedures can be examined.

Overview of the Program and the Model that It Embodies

Typically, procedures modeling arrays of neurons (e.g., the bipolars, or V1's simple edge detectors) will execute a process everywhere in parallel and pass the result on via output links. These results may also be stored, with an appropriate decay, at that point (which can be thought of as a local memory buffer??). The program will be coded to execute in two modes, static or dynamic. The static (which will often be used, chiefly for economy and simplicity) will handle input images as though viewed in a brief (e.g., 1 msec, essentially instantaneous) flash. [This is the mode of computer vision programs. It can be interpreted as one where there is an immediate adaptation - that is, as soon as a firing occurs it is damped out. Its main purpose is to reduce the enormous computational burden.] The dynamic mode will have images continually firing. [This is computationally expensive, since it necessitates continual up-dates of the states of all parts of the system, using an appropriately fine temporal resolution.]

Where a particular subsystem's behavior is known, but facts about its structure or processes are incomplete, more or less "black-box" procedures will be used, and compared with procedures that attempt to incorporate the particular underlying structures. [For example, columns in V1 in primates will be coded as a several-level structure that transforms maps of spots to simple cells' slopes at particular locations, and complex cells' stopped edges and slopes located anywhere - using whatever structure of processes appears to be most plausible, simple and general.] This kind of black-box process will be compared with one or more processes that attempt to achieve the same results using more realistic models of the actual neuron network. Since complex net-

works like the hundreds of neurons in each column, plus large numbers of links to other columns and other areas, are far from being completely mapped out or understood, a procedure that attempts to fill in and interpret unknown facts will almost certainly be wrong in many details - but it can be of value in suggesting alternatives and showing their repercussions.

The Basic Processors: Synapsing Neuron-Like Threshold Elements

The neuron will be simplified greatly for the first program, to a threshold element of the following sort: One or more neurons input weights (via synapses) to an output neuron. These weights are combined according to whatever function is desired for a particular run (to start, addition will probably be used). If the combined input exceeds its threshold, the output neuron fires into one or more neurons to which it inputs. [Note that the use of thresholds, and of simple circuits that choose the most intensely firing neuron, build into highly non-linear processes.] All neurons will be clocked at the same speed (assume this is 1 msec). *{NOT modeled: different types of electrical and chemical synapses; the thousands of synapses to each neuron (much smaller numbers will be used); different times for synaptic transmission and conduction; different times for graded and action potentials; effects of neurotransmitters.}*

The sharpness of tuning (e.g., of a simple cell to a particular slope), amount of overlap (e.g., between ganglion or simple cell receptive fields), amount of convergence (e.g., from cones to bipolar cells), and amount of divergence (e.g., from Layer 4 to cells in other layers) will be a function of neurons' weights and thresholds, and the links between them. In most cases these will be greatly simplified, and handled in a completely regular manner, in contrast to the pervasive minor variations found in living systems. For example, overlap will be the same everywhere.

Input of the Image and Retinal Processing

The input array can be any 2-dimensional array. For convenience, the smallest size that adequately resolves images of interest will be used - probably 64x64 or larger. (Binocular images can be input, but this will be done only for runs that examine binocular issues, like merging, competition, and dominance.) The program will input an image into a "retinal array." This can either be a grey-scale or a 3-color array, with whatever range of intensities is desired for a particular test run. *{NOT modeled: breakdown of rhodopsin and details of transduction; the rod system; the cone system's finer resolution at the center, hexagonal packing, unequal placement of cones sensitive to different colors; pervasive small irregularities; many other known details of the microcircuitry.}*

Several procedures will be coded for retinal detection of on-center off-surround and off-center on-surround spots. These will include black-box computations (e.g., Difference of Gaussians; appropriate gradient detectors), and procedures (parametrized so that tests can be made as to, e.g., the effect of varying the size of the surround), that gather the surround information using horizontal cells. One variant to be investigated will have horizontal cells synapse not only on bipolar cells, giving information about the surround; but also with other horizontals, forming a mesh that introduces a global component to this information. It will be assumed to start that X-cells handle shape and color, so that the Y- and W-cells can be ignored. *{NOT modeled: most of the complexities of this structure, which is built from at least 60 types of neurons and a number of different microcircuits.}*

The retinal processing will result in an appropriately reduced (as well as transformed) image. This will be accomplished by merging (with appropriate overlap) information output by neighbors. The reduction will be on the order of 2-, 4-, 8-, 16-fold.

The lateral geniculate will be modeled as a single level of parallel transformations where one local threshold operation can be made, with one, or when desired several, outputs. (This will usually not be used, since LGN projections in V1 in monkey appear to be the same as in ganglia and LGN. But the model can be used to probe what roles the LGN might serve.)

Primary Visual cortex

The striate cortex (V1) will (as found in primates) receive LGN projections in Layer 4, unchanged. These will be transformed by detectors of (simple) local sloped edges, and (complex) stopped edges and sloped edges anywhere, arranged in orderly hypercolumns with inputs from the two eyes side-by-side. Simple networks of elements will combine LGN spots (even, as experiments indicate, 2 or 3 contiguous spots) into simple sloped edges, and simple edges and spots into more complex features. More realistic networks, that attempt to pull together the many facts about columns, will be formulated, coded, and compared. Columns will be arranged in modules with inserted blobs; these will be taken as the major processing units. This gives several complex arrays of information, each containing sub-arrays in an orderly arrangement. *{NOT modeled: links between LGN and Layers 6 or 1; finer resolution moving toward the center; and (except as a test-bed for plausible conjectures) those many actual links between neurons still largely unknown.}*

Columns will be linked to one another, as indicated by anatomy and physiology. The known facts about these links (e.g., that they are between similar columns, over several mm distances) will be incorporated, along with the most plausible conjectures (e.g., that similar columns excite each other, possibly serving to build longer contours; that dissimilar columns inhibit each other, possibly serving to choose among competing alternatives). *{NOT modeled: The local and more global variations in links; the great amount that is not yet known.}*

Additional Visual Areas, and Links Between Them

The program will be coded so the different visual areas can be linked to one another using whatever topology is desired for a particular run. This means that each area will output in parallel to each area to which it is linked. The links used between areas will be those actually found (Van Essen, 1979, 1985; Van Essen and Maunsell, 1983; Mishkin et al., 1983; see below). Output will go in both directions (one will usually be for feedback, sending different types of information). Feedback will chiefly serve to continue to cycle and evaluate information. *{NOT modeled: Details of different sizes and mappings of areas; any but the simplest feedback (since little is known).}*

The actual processes executed in these areas are beginning to be identified. But where they are still unknown they will be chosen on logical grounds, to fill in the overall structure of a loosely hierarchical parallel-serial distributed system. Different types of processes (e.g., linked edges, shape features, Fourier descriptors, region-growers), used separately and in combination, will be compared. Essentially, they must fill the gap between the features to which cells in V1 are sensitive (sloped lines, stopped on 0, 1, or 2 ends) and the complex objects recognized in the temporal lobe.

Among the facts and possibilities to be investigated:??expand? V2, V3, and V4 have many complex cells that appear to fire to the same stimuli as those in V1. They have not been systematically tested for responses to more complex shapes (Van Essen, 1985). Columns appear to link in regular fashion both within and between areas. V1 and V2 extract and send different types of information to different areas, for further processing. MT is sensitive to motion and

spatial relations. V4 appears to fire selectively to color (and also slope), uses global information, and may be building up more complex shapes.

Temporal Lobe Recognition of Faces and Other Complex Objects

The selective firing of individual neurons in the inferior temporal cortex (IT) and the superior temporal sulcus (STS) to face and other complex objects, and to particular faces and complex parts of face, will be modeled as follows: Processes, that can be interpreted as individual cells or as ensembles, will be coded to be sensitive to different features, and compounds of features. These will be linked in a loose hierarchy of complexity, some of whose stages will be embedded in areas anterior to IT and STS.

Information about motion and spatial relations will be used to the extent that it appears to be necessary, or helpful, in the recognition of objects. For example, the spatial relations between the parts (e.g., between the two eyes, nose, mouth and chin) contains a great deal of useful information about the whole. The visual areas culminating with the parietal lobe will therefore be included, but not as ends in themselves. Rather, their results will be merged together with the information in IT and STS.

The Total Structure, and Future Possibilities

The purpose is to achieve a structure of cells that process complex images to levels of abstraction where cells are sensitive to, and in some sense recognize, complex objects like faces and hands. They must do all this within the extremely stringent time constraints, using an anatomy that is dictated by the living system and processes that are similarly dictated (or when not known are plausible, simple, and general).

If this much of the model can be programmed and run successfully, the next phase can begin to relax the many simplifications and omissions. Subsequent steps of processing can be added, in the parietal and prefrontal lobes, the amygdala, and the reticular formation. This also means adding mechanisms that combine information from several sense modalities, and address problems of motivation, perceptual-motor coordination, and the focusing of attention.

Previous Related Work: Parallel-Serial "Recognition Cone" Programs

This model grows out of a type of program being developed for highly parallel shallowly serial perception of complex real-world objects (Uhr, 1972, 1976, 1978; Uhr and Douglass, 1979; Schmitt, 1981; Li and Uhr, 1985a, 1986). These programs have successfully recognized a variety of complex real-world objects (e.g., neurons, cars, trees, windows, houses) in real-world photos and TV images.

The long-term goal of this research is to develop programs that will recognize and describe scenes of objects as they move about in real time. These include two major types of programs: those designed for speed, efficiency, and power using appropriately-structured multi-computer networks made from existing microelectronic technologies, and those that model the structures, processes, behavior, and performance of the brain.

Even the fastest of today's super-computers (Riganati and Schneck, 1984; see Uhr, 1984,

1986 in press) are far too slow to be programmed to recognize moving real-world objects. But since the programs being developed are designed to run on a suitably structured pyramid multi-computer (see below) that has large numbers of computers all working in parallel they offer real hope of achieving this goal - if a good structure of good processes can also be developed.

These programs have been developed with the brain's structure serving, loosely, as a model (see Uhr, 1980, 1974, 1972). Chiefly, the visual system's generally converging structure of threshold element processors arranged in layers was taken to suggest a pyramid- or cone-shaped structure of arrays of locally linked transforming processes reminiscent of neurons. The kinds of transformations found in the visual system, where gradients, spots, simple oriented edges, complex edges, etc., are discriminated, suggests - as does a simple logical analysis - a hierarchy of transformations that build a successively more global and more abstract assessment of the scene.

The Kinds of Processes Executed by These Programs

These programs start by inputting the image (e.g., a 512 by 512 TV picture). They apply one or several relatively local operations (e.g., to difference, detect gradients, smooth) in parallel everywhere, and continue to apply operations that appear to be appropriate (e.g., to detect short edges, to combine colors, to assess simple local textures). At higher levels more global characteristics, including contextual information, are examined, combined, and used to imply more complex objects (e.g., curve-fragments and textures imply banana-skin or cheek; properly inter-related cheeks lying below a region raises confidence the region is the eye-ensemble). These processes are chosen primarily because they extract, abstract, and combine useful information fast and efficiently, applying simple, local operations in parallel.

Whenever the programmer thinks it appropriate, this information is converged, using whatever combining operation (e.g., sum, median, max, and, or) is defined (followed, if desired, by thresholding and/or choosing operations), and input to the next-higher level of the pyramid.

Operations of this sort are going on at every level, in parallel. At the same time, higher-level operations feed their results back down to lower levels, to influence processing.

When one codes the program to send information between levels that are not adjacent, the intervening levels serve as way-stations that simply pass that information through. This is often conveniently effected by a simple logical-or, that produces an output whenever the transform receives an input. (These in effect give pass-through neuronal links, where one neuron fires another and no actual transformations of information are effected.)

Developing More Realistic Models

These systems loosely model some of the brain's structures, but they deviate in many ways. This is necessary to make it possible to run them at all; the brain is so complex, and so many of its processes are largely unknown. Although probabilistic, thresholding, neuron-like processes are used, they are kept very simple, and given far fewer than the living neuron's thousands of links. The network linking processors is much more regular than the brain's. It makes use of topologies that can economically be built into computer hardware and software, and are known to be efficient in terms of the processes being coded (mesh links for local operations, tree links for global operations and for convergence and divergence of information). But the overall structure is amenable, and even friendly, to modifications that move it closer to the brain.

A recognition cone program run on a hardware pyramid (or on a software simulation of a pyramid using a conventional single-CPU computer) must pass information via links to siblings, parents and children, and therefore laterally through one array and up and down to adjacent arrays. But that is a hardware detail; such a system can be used (although possibly inefficiently) to pass information from any region to any other region. Therefore one could code a program that assigned successive levels of the pyramid to different areas of the brain (e.g., V1, V2, V3, V4); now, for example, information can be sent from V1 to V3 by simply passing it through V2 unchanged. Alternately, several areas of the brain can be assigned to the same level of the pyramid, all the areas processing information in parallel. If the hardware has N processors (each an array of parallel processors) that can be assigned to N different areas this can be a hardware-parallel process, otherwise it will be a simulated-parallel process. But the results are exactly the same - except that the simulation will take longer, since what are really parallel processes must be executed one at a time.

The Kind of Hardware Structure (Anatomy) Most Appropriate for These Programs

These programs, when executed on conventional single-CPU serial computers, simulate parallel processes by iterating through them one at a time. This gives exactly the same result as would the appropriate parallel computer, except that it takes orders of magnitude more time (Li and Uhr, 1985b). The programs are designed to be executed on appropriately structured parallel networks built from many thousands, or millions, of computers. The present state of technology makes this kind of parallel computer entirely feasible. This is so because (to boil a complex issue down to a simple argument) the cost and difficulty of building a computer is directly related to the number of separate components that have to be wired or plugged together.

A Quick Look at Feasible Sizes of Computers

A computer that has 10^2 components is simple; 10^3 is comfortable; 10^4 is getting pretty large; 10^5 is pushing things. Today each basic component is a silicon chip that can have 10^6 transistors, and this number doubles every 12 to 18 months. A computer's general-purpose processor needs only 10^2 to 10^5 transistors. So millions of processors can be achieved.

Actually, memory takes up most of the silicon area in today's computers. But a computer that worked like the brain would flow information through an ensemble of transforming processors, rather than at each step store it back into memory (the technique used in the "stored-program computer"). It would therefore no longer need large amounts of memory.

Large networks are today just beginning to be built, and a great deal of research is still needed to arrive at a good understanding of what is the most appropriate design (see Uhr, 1984, 1986 in press). However, it is not crucial that the appropriate design be achieved. Any design that offers significantly more power and speed than a conventional serial computer is worth using. The more appropriate the design the faster and more efficiently will it execute the program.

Large Specialized Multi-Computer Arrays

For example, 2-dimensional arrays of thousands of computers have already been built and used for local feature detection and other image processing tasks, giving orders of magnitude speed-ups (Duff's 1976 CLIP4; Reddaway's 1978 DAP; Batcher's 1980 MPP). But some of the

processes they must execute can only be done slowly and inefficiently. For example, to send information from the processors responsible for particular columns in one hypercolumn to processors responsible for similar columns in other hypercolumns these arrays would have to shift information step-by-step through all intervening computers. Direct hardware links could be added to specialized versions of these arrays (presently they link each cell directly to its 4, 6, or 8 nearest neighbors). But links and interfaces are expensive using today's microelectronic technologies. It is not clear what particular links at a distance would be most useful, and it seems unlikely that the additional costs would be worth paying, when the program could be run without them, the only price being occasional slow-downs.

Pyramid-Structured Stacks of Linked Arrays

Pyramids of arrays are now being built (Cantoni et al., 1985; Schaefer, 1985; Tanimoto, 1983, 1984); in fact Schaefer has a 16x16-based prototype running. Essentially, a pyramid links successively smaller arrays via a tree. This makes possible multi-level pyramid operations as well as single-level array operations. It also allows information to be converged up the pyramid, diverged down to larger layers, and passed between different nodes via the logarithmic distances up and down the tree.

A pyramid needs at most only $1/3d$ more processors than an array, and therefore is economically feasible. More powerful systems could be built by using multiple pyramids, or by placing several processors at each node. Now each processor might in parallel assess a different simple feature, for example the different sloped edges of simple cells. The only problem is the extra expense: Is it worth it, since the simpler structure can do the more complex operations when they are needed, but more slowly? Once we know exactly what we want to do we can build the appropriately configured system. Until then it is often better to make do with an anatomy that is usable even if less than ideal.

Multi-Computers with a Wider Variety of Topologies

Arrays already give many thousand-fold speed-ups, and they can easily be increased in size, potentially without limit. Arrays and pyramids use very simple processors; this allows them to have thousands (or, already becoming feasible, millions) of processors all working in parallel. They thus achieve orders of magnitude increases in speed. Their very regular, micromodular design, along with their simple, synchronized control make it relatively easy to build them, and to program and use them.

An enormous variety of other architectures are possible - potentially, any desired graph topology (see Uhr, 1984, 1986 in press). A number of potentially interesting designs are being explored, including clusters of clusters (Kuck et al., 1983), trees (Shaw, 1986), augmented trees (Despain and Patterson, 1978), hypertrees (Goodman and Sequin, 1981), and pyramids embedded in larger systems (Tanimoto, 1985; Uhr, 1985).

However, other than arrays, the largest multi-computers built to date have only a few hundred processors (e.g., shuffle networks - Crowther et al., 1985; Hypercubes - Fox and Otto, 1984; Peterson et al., 1985). They are far more costly, and far harder to program and to use efficiently. Several deep and difficult problems must be solved before large networks of unsynchronized computers can be used effectively. For example, almost all have major problems handling message-passing, which typically takes thousands, or at best hundreds, of times longer than pro-

cessing. This degrades performance substantially, unless programs rarely send messages. But brains are always sending messages, to near and distant places, and that is a very difficult constraint within which to program a model of the brain.

Networks, Including Ones that are More Specialized to the Brain's Structure

One might consider designing specialized hardware architecture that assigned an array or other appropriate structure of the appropriate size with the appropriate linkage between siblings to each part of the brain being modeled, with links between these arrays that similarly modeled the brain. Since these are not yet firmly mapped out or well understood, it is much easier to simulate them, changing the structure as we learn more, using the computer whose structure makes this simulation as easy, fast, and efficient as possible.

Possibly the best design for a multi-computer that would be reasonably general yet specialized enough to run programs modeling the visual system (and other brain structures) with efficiency and speed would link a small number of suitably specialized arrays, pyramids, and truncated pyramids (that is, with missing layers) over a topology that expedites efficient passing of information from structure to structure. Bermond et al, 1983, summarize a number of elegant and efficient recently discovered topologies. Among the most interesting from the point of view of multi-computer architectures (see Uhr, 1984, 1986 in press) are Moore graphs (Hoffman and Singleton, 1960), De Bruijn networks (Imase and Itoh, 1981), and also shuffle networks (Siegel, 1984). However, almost all have structures that are probably too intricate for evolution to achieve, and are not fault-tolerant. Fortunately, simpler structures, although less efficient, will work. For example, a number of suitably specialized massively parallel modules can be linked via a simple array, tree, or augmented tree. This is not as efficient as, for example, the proved-optimal Moore graphs, but that merely means that more time will be needed to send information.

Fitting the Model's Structure to the Program's Structure to the Computer's Structure

The program describes and in a real sense embodies the model. That is, the program is (a specification for) that model. The computer interprets and runs program, executing the model. The situation is much as though we could draw a blueprint of a factory assemblyline, and feed this into a gadget that then actually built that particular factory and started running the assemblyline. Any structure we build into the gadget and the modules that it uses to construct the running assemblyline will make the gadget's job easier or harder, depending on how similar or dissimilar it is to the structure of this assemblyline. Similarly, the programmed model will run faster or slower depending upon how well it fits into and can be executed by the computer being used.

Consider how best to handle the lateral spread of information by horizontal or amacrine cells, or the clustered links between hypercolumns. A special-purpose array could be built whose processors linked (as in the brain) to a regular lattice of more distant processors. Or an array like CLIP could be programmed to shift information step-by-step to the distant processors, taking D steps for a distance of D . Or a pyramid could pass information up, then down, in $2\log D$ steps. All will give exactly the same results.

The following alternative ways of handling N different areas will all give identical results. [Assume the hardware system is built from one or more networks (arrays, pyramids, or whatever is most appropriate). Assume each area has M micro-areas (e.g., modules, or columns).]

A. One network of M processors processes all N areas. Information about each different area

is stored in different memory locations. (This system processes within an area in parallel but processes the areas one at a time, simulating parallel processing of areas. The results are exactly the same as if all areas were processed at the same time, but it will take N times as long.)

B. A different network of M processors processes each area. (Now hardware links are needed, where appropriate, between these different networks. This system will be N times as fast as the single network; but it has N times as many processors, plus the extra links.)

C. One network of M processor-clusters processes all N areas, where each processor-cluster is actually a set of N processors, and processes all N areas in parallel. That is, the different areas have been merged together. (This system will also be N times as fast as the M -processor network; but it has N times as many processors. It does not need extra links between areas, but it does need extra links between the processors within a cluster and their shared memory.)

Both B and C are N times faster than A, which is important, and possibly crucial. (Note that a conventional single-CPU serial computer will also give exactly the same results. But it will be at least M times slower still - which is surely crucial since M is a large number, at least in the thousands or millions.)

Both B and C need at least N times as much hardware. The different visual areas suggest B, but the several different processes that the striate cortex appears to handle (edges, intensity, binocular information, motion, color, possibly texture, etc.) suggest C. I tend to prefer C, since it can at every step compute a much larger set of functions over any combination of information. Evolution seems to prefer a combination, although this may simply mean it's easier to grow a new area than to squeeze new processors into an already-crowded space. Or are these unimportant implementation details, depending on local costs for raw materials, since the results are the same?

There are similar ranges of possibilities, and tradeoffs, for the structure and function of columns and hypercolumns, and the ways that the different areas send and process information.

Brain Models, and Programs for Pattern Recognition and Computer Vision

A great variety of verbal, mathematical, neuronal, and computer-programmed (computational, information-processing) models are potentially relevant to this enterprise (e.g., Rashevsky, 1938; Craik, 1943; McCulloch and Pitts, 1943; Hebb, 1949; Selfridge, 1955; Beurle, 1956; Uttley, 1959; Rosenblatt, 1962; Young, 1964; Deutsch, 1967; Konorski, 1967; Dodwell, 1970; Arbib, 1972; von der Malsburg, 1973; Szentogathai and Arbib, 1974; Uhr, 1974; Braitenberg, 1977; Edelman, 1979; Fukushima, 1980; Cooper, 1981; Julesz, 1981; Hopfield, 1982; Kohonen, 1982; Marr, 1982; Grossberg, 1984; Orban, 1984; Braitenberg, 1977, 1985; Eccles, 1952, 1985; Feldman, 1982, 1985; MacKay, 1956, 1985; Szentogathai, 1985; Ballard, 1986; Nevatia, 1986; Rumelhart et al., 1986). [I hope to examine these in a separate paper. There are many complex issues involved in using, judging and building models, but space here for only a few brief comments.]

Verbal, Mathematical, and Partly Mathematical Models

A model that runs on a computer is precise in all details (of anatomy, physiology, and behavior) addressed. It and its consequences can be examined, tested, and evaluated, and compared with other models. It is important to stress that the computer is not the model. It is merely

the medium for expressing the model, just as pencil and paper are the medium for expressing mathematical or verbal models.

Verbal models that predict empirical observations, or are built around a particular set of processes, can address only a few of the many details. It is hard to be entirely certain how they would behave; they rarely lead to testable predictions. Mathematical and network models typically address only parts (e.g., V1) of a complex process like object perception. Complex problems must often be simplified and divided into more tractable pieces; but the original problem remains the goal, and organizes and gives meaning to the parts (see Maxwell, 1985). Computers make feasible the largest possible range of more complete models (any mathematical or verbal model can be expressed as a computer program). ??eliminate this whole section - dogmatic unless examines and explains?

Computer Programs for Pattern Recognition, Image Understanding, Etc.

Virtually all of the computer programs for what is variously called pattern recognition, computer vision, image processing, or image understanding have largely ignored much of what is known about living perceivers. [However, they often rely heavily on the intuitive insights into perception that we all have. After all, the living perceiver is the only successful system around. Even people with no interest in modeling the brain, and with a particular practical problem that must be handled efficiently using available technology (2-state transistors on chips packaged into precisely controlled serial digital computers, rather than many-state analog/digital neurons intertwined in scarcely controlled, massively parallel, shallowly serial structures) often draw ideas from nature.] Most are extremely serial, and slow, the issues of speed and parallelism ignored.

From the point of view of modeling the brain, people writing computer programs for pattern recognition, even when they are interested in brain processes, will at least occasionally do things that are not at all brain-like. Often (but not always) this is an implementation detail - the process could have been done in a brain-like manner, but given the economics (costs of programming and hardware) of the computer being used there was a cheaper, simpler way.

Most programs apply whatever set of processes seems most appropriate, using whatever computational procedures seem most efficient. For example, the image might be assessed in a variety of ways, using moments, or/and measures of discontinuities, smooth regions, jaggedness. A variety of different procedures might be used to decide which of these features to examine next, and where, and how to combine results and make decisions. Often, each object to be recognized is stored as a graph-like model, and the attempt is made to graph-match each with the input image. This, at least as handled today, is an extremely serial (and difficult) process.

Probably the most brain-like systems are massively parallel, shallowly serial pyramid systems (see Tanimoto and Klinger, 1980; Rosenfeld, 1984) of the sort described above (see also e.g., Hanson and Riseman, 1978; Levine and Leemet, 1976; Sloan, 1977; Tanimoto, 1978; Bajcsy and Rosenthal, 1980; Rosenfeld, 1983; Neveu et al., 1985). Less obviously, but to some extent, brain-like are the large number of image processing systems for the earlier stages of perception that make use of the type of massively parallel local array operations that can be executed with great speed and efficiency on a cellular automaton (Von Neumann's brain-motivated model).

There is no space to describe these different systems here, except to make the general comment that they input the large image and then apply parallel-serial distributed structures of mas-

sively parallel processes, attempting to recognize.

A More Detailed Description of the Living Visual System

The following attempts to give a fuller picture of the living visual system. So much is known, and so much unknown, that it is still selective and incomplete. But it tries to cover all aspects of the problem that are pertinent to the way the brain recognizes complex objects.

The Basic Processor, the Neuron

Neurons send information to other neurons over synaptic gaps between them. In the classical chemical synapses, one neuron's axon synapses on another neuron's dendrite, over a roughly 30 nm gap. The efferent (sending) neuron must pump and diffuse transmitter substances, in a complex process that finally builds to a charge that exceeds the afferent (receiving) neuron's threshold, hence fires it. (There are several hundred transmitters, including a few that are used widely. They are secreted by many neurons, and also by more global sources like the endocrine system. They appear to play an important, but still largely unknown, role.) It takes roughly 1-2 msec to fire (the range can be from .3 to 5 msec or more) over a chemical synapse.

Electrical synapses occur where fibers come very close together, typically about 2 nm apart. They are much faster, since the neurons interact at electronic speeds; but they are more analog and less controlled. Electrical synapses can link dendrite to axon, cell body, or dendrite, or axon to axon. The interneurons in the retina fire electrically, as do at least some ??which? of the cortical interneurons.

The typical output of a chemical synapse is a digital volley of "all-or-none" spikes that travels quickly (3-120 m/sec), often over long distances, through myelinated axons. This is the classic mode of conduction, from one region of the cortex to another through the white matter, and to and from sub-cortical regions. The impulses integrated by unmyelinated processes move slowly (.5-2 m/sec); as they move through branching fibers they appear to be capable of computing complex, analog functions.

The Retina and the Lateral Geniculate

Each of the two eyes has roughly 100,000,000 rods and 6,000,000 cones. The cones are sensitive to 3 different color frequencies (red, green, and a much smaller number to blue), and are the sensory input transducers for the recognition of complex shapes and objects. Rods are exquisitely sensitive - to single photons, and to slight changes in intensity and location (motion). Cones and rods fire into short bipolar cells that in turn input to ganglion cells whose axons leave the retina in the 1,000,000 neuron optic nerve and synapse in the lateral geniculate. Horizontal cells ??give numbers of all these? spread laterally, often with large gap junction synapses to one another, so that they appear to form a single mesh throughout the retina. They receive inputs from the cones and rods and output to the bipolars. Amacrine cells spread laterally over a region of roughly ?? bipolars. They receive inputs from the bipolars and output to the ganglion cells and also to the bipolars.

The ganglia contain a detailed map of the image, emphasizing spots of light (Kuffler, 1953) - and therefore gradients and edges. Roughly 60% are X-cells most sensitive to spots of bright light surrounded by less light (or the opposite).??the others? In cat (Sterling, 1983, Sterling et al., 1986) about 60 different types of cells and several major circuits have been identified, for daylight, twilight, and starlight vision. E.g., Y/alpha-ganglia have 150 um diameter dendritic fields; each collects signals from almost 200 bipolars; X/beta ganglia have much narrower fields??what? and fire to smaller spots. The cone-bipolar circuit of On-X/beta ganglia has 26,000 cones, 6,500 bipolars and 2000 ganglia per mm² at the center. Each cone inputs to 4 bipolars, each of which inputs to 4 ganglia. It appears that networks of horizontal cells collect additional cone signals.??incomplete; but too long!

Cones are concentrated and tightly packed in the fovea, forming a regular triangular/hexagonal mosaic. They are even miniaturized in the center, to about half their typical diameter (Williams, 1986). In the central fovea a bipolar will link to a single cone and a single ganglion, and respond to only a few minutes of arc. Moving away from the center, rods are interspersed between cones, giving less regular arrays; then rods take over. Bipolars link to increasing numbers of cones and rods; at the periphery each responds to 10-20⁰.

The Lateral Geniculate Nucleus (LGN) and its Links to the Primary Visual cortex

The LGN serves as a way-station where retinal ganglia synapse with cells that input to Layer 4 of V1 and Layer 6 of V1; the latter link back to the LGN. Most are in 4 groups of X-cells (which continue to fire to a sustained input, have small receptive fields, are sensitive to shape and color, and have moderate velocity) that go to Layer 4A, 4Ca, 4Cb (for orientation) and 3 (for color) of V1 (Hendrickson, 1985; Rodieck, 1979; Weller and Kaas, 1981). A number of Y-cells respond briefly to changes, have larger receptive fields, and are fast conductors; they project to the superior colliculus as well as to Layer 4. The rest are possibly numerous slowly conducting W cells, with unknown, complex functions; some may fire to both on-center and off-center fields; they project to Layer 6 of V1.

The Cerebral Cortex

The chief areas of the brain involved with form perception appear to lie in the neocortex. There are a large number of sub-cortical visual areas that appear to be primarily involved with motor control and motivation that for the present will be ignored.

Each hemisphere of the cerebral cortex would, if spread flat, be about $80,000 \text{ mm}^2$ in area in man (??but Carpenter says 1.3 ft^2) $8,000 \text{ mm}^2$ in macaque monkey, and $5,000 \text{ mm}^2$ in cat. The 1 to 4.5 mm thick grey matter is arranged in 6 layers. (It has the same number of neurons no matter what the thickness.) In volume it is about 50% neurons, 50% glia. The much smaller glia are roughly 10 times as numerous as neurons. (Their function is largely unknown, although they guide developing neurons, insulate axons, and may absorb and diffuse neurotransmitters.) Grey matter is underlaid by white matter - the mass of myelinated axons that make cortico-cortical links ??how many, pattern, distances, etc?.

The columns in the striate cortex (the binocular part of Area 17 in primates) each contain about 150% more neurons than columns in other areas. Using electron microscopes, Rockel et al. (1980; see Powell, 1981), averaging 40 counts of cells in column-sized 30 um wide by 25 um thick sections in motor, somato-sensory, frontal, temporal and parietal areas in macaque, found 109.4 ± 9.4 to 114.6 ± 9.9 . In contrast, striate, had 267.9 ± 13.7 . Similar counts in other species showed 258.9 ± 15.8 cells in Area 17 in man, but not cat, rat, or mouse. Five other monkeys gave similar results for striate cortex (the largest was 279.0 ± 5.66 in chimpanzees); one, Tupaia (tree shrew) had 109 in monocular and 192.8 in binocular regions - an interesting hint as to the genesis and purpose of the extra neurons. (Note that binocular regions in cat have only 110 neurons.) The range across all areas for all species (except Area 17 in primates) was from 102.3 ± 9.5 to 114.6 ± 9.9 . Thus there is a striking uniformity in a great variety of columns. All contain about 110 neurons, with the major exception of striate cortex in primates, which contain 260.??what in-between evolutionary steps are there?? Beaulieu and Colonnier, 1984, got somewhat different results, although the variations are still narrow. They found, in 1 mm^2 areas in cat, 78,000 neurons in 17B (binocular areas of 17), 59,000 in other sensory areas, and 44,000 in motor areas. Most, but not all, the extra 17B cells (and of primate striate cortex) were small stellates in Layer 4.

Winfield et al. (1980) found in cat, rat, and monkey, that from 62% to 69% of the cells in the visual and motor cortices were pyramidal, from 3% to 7% were large stellates, and from 21% to 33% were small stellates.??expand, with functions? Other types of interneurons (often with characteristic dendritic arborizations) exist, including spindle and basket cells. Some anatomists recognize only 4 or 5, others several dozens.??elaborate? Almost all (??) pyramidal cells have axons making cortico-cortical links via the white matter, as do many (??) of the interneurons. All the pyramidal cells and most of the interneurons are excitatory. (??, more info of importance about types??)

The Primary Visual Area (V1), the Columns, and Hypercolumns

V1, the primary (and largest) visual area has the most extensive and the most detailed representation of the (by now transformed and abstracted) image. It may also be the area that does the most processing.??why other than that more cells in a column?each area's size,neurons in each?? In man it is roughly $24\text{--}3200 \text{ mm}^2$,??jibe with 1000 mm^2 columns? about 2-4% of the total neocortex, 25% of the occipital lobe, and twice the size of macaque (Van Essen, 1985)??cat?.

The following are some of the relatively firm and more important details. The cells (primarily in Layer 4, with some in Layer 6) that receive inputs from the LGN are (in monkey) like LGN and ganglion cells in that they respond most strongly to spots. Layer 4 contains many tightly-packed tiny (4-8 um) stellate cells.

Hubel and Wiesel (1959, 1962, 1968, 1972, 1977; see also Kuffler et al., 1984) have discovered an impressive amount about the types of features individual cells in the striate cortex are sensitive to, and the functions of the larger columns and hypercolumns into which they are linked. There are many cells in Layers 4 and 6 that are Hubel-Wiesel "simple" cells, sensitive to a localized edge with a particular slope moving at a particular speed. (In cat and other mammals LGN cells link directly to simple cells.) Each simple cell is tuned to a particular slope, but also fires with decreasing sensitivity over 10-30 degrees in either direction. Swindale and Cynader, 1984, found great sensitivity to tiny, local sloped edges: bars only 2-3 ganglia long fired cells, and flashed spots evoked the largest response when they fired contiguous ganglia.??can this be extended to longer curves? Sillito (1975, 1984; Sillito et al., 1985) found that blocking GABA (the neurotransmitter used by inhibitory synapses) eliminates simple cells' sensitivity to slope, although the effects on complex cells varied greatly.

Simple cells in Layer 4 project to Layers 2 and 3, which project to Layer 5. All 3 layers contain many Hubel-Wiesel "complex" cells, including ones that fire to the same moving edge anywhere, and "hypercomplex" cells (roughly 15%) that fire to edges that end in one, or both, directions. Layer 6 also contains complex cells with greatly elongated receptive fields. Michael, 1981, 1985, found columns of simple, complex, and hypercomplex cells were responsive to sloped edges of a particular color. Most cells in Layer 4 are driven monocularly, whereas cells in other layers are driven by both eyes. Large pyramidal cell bodies (roughly 30 μm in diameter) are found in Layers 2, 3 and 5. Pyramidal dendrites ascend and descend vertically through most layers. Pyramidal and many other cells' axons descend to link, via the white matter, to other cortical areas.

Complex cells in Layers 2 and 3 have small receptive fields and low spontaneous activity. They synapse on cells in Layer 3, and also provide most of the output to other cortical areas. Many cells in Layer 6 respond to long lines; some cells (??%) in Layer 5 respond to short end-stopped lines. Some may be simple X cells, some complex Y.??fill this out) Most simple cells are stellates in Layer 4; most complex and hypercomplex cells are pyramidal in Layers 2, 3, and 5.

Many more details are known; and there are many complications. The first experiments suggested there was evidence for a relatively simple hierarchy of detectors of 1. spots; 2. sloped edges; 3. sloped edges everywhere, single-ended, 4. double-ended; 5. angles and more complex ?? features, 6 - with a lot farther to go.?. But subsequent findings of parallel paths make such a simple hierarchy unlikely (although they do not deny that important hierarchical, as well as parallel, processes exist).

For example, LGN cells have been found that synapse directly on complex cells (Toyama et al., 1977; Bullier and Henry, 1979). White and Herch, 1982, found LGN inputs to particular types of pyramidal cells in Layer 5 (see Gilbert, 1983, for other examples). Malpelli, 1983, found that inactivating Layer A of the LGN virtually abolished all activity in Layers 4 and 6, yet complex cells in Layers 2, 3, and 5 were not affected. Hammond and MacKay, 1977, found that some complex cells (probably in Layer 5) but not simple cells respond to textures. To explain the puzzling fact that monkeys have very few simple cells, Hubel, 1982, suggested that different dendritic branches of a single neuron might serve their function. Bishop et al. (1971a, 1973) found complex responses of simple cells to slits of different widths and speed which greatly obscure the distinction between simple and complex cells. Cells respond separately to each edge of wide slits, most responding to both, some to both in both directions.??more details; others? They also (1971b) propose a simple model for achieving these results. Stone and Dreher, 1973, found (in cat) that many cells receive direct input from the LGN and that these, rather than hierarchical processing, can account for the differences between complex and simple cells. Stone et al., 1979, comparing hierarchical and parallel models, showed that the different properties of complex and hypercomplex cells probably result not from a hierarchy but from the different properties of segregated X-, Y-, and W-cell pathways. They suggested that simple-to-complex pathways (e.g., Henry et al., 1974) occur within each of these. They pointed out (see also Zeki, 1978b) that in primates information diverges from V1 and V2 to areas that process information in parallel. Thus rather than a strict serial hierarchy, both parallel and serial processing are combined.??More details, problems? [This seems reasonable, since time-space-economy issues determine parallel-serial tradeoffs. No system is purely one or the other.] Leventhal, 1985, suggested that retinal ganglion cells' sensitivity to slope (Levick and Thibos, 1982) underlies V1 simple cells'.

Columns are arranged in rows, each adjacent column (about 20-50 μm thick) firing to an edge rotated about 10-15°. ??Does angle move in jumps, or more continuously (interpolating??)? These cells are rather broadly tuned, each firing to a +-10-30° range of slopes. [This would seem to suggest that each receives input from and integrates over a large set of spot detectors (firing to many different sub-sets) and/or other simple cells (averaging these to interpolate and smooth).] Slope usually (but not always) progresses in an orderly manner, but fractures typically ap-

pear, somewhat irregularly??, every 1 mm or so. Two rows (each about 250-500 μm wide) stand side-by-side, one for each eye (in old world monkeys). [??Since macaque V1 hypercolumns contain 300 μm by 30 μm slabs for each slope for each eye; each has $260 \times 10 = 2600$ cells. Does this suggest that several features (e.g., color, intensity, motion, texture, curves) are being processed along with slope.??] The total size of a hypercolumn is about .5 to 1 mm^2 . LeVay et al., 1985, found that each hypercolumn at the periphery is mapped to by a retinal region sensitive to several degrees of angle; at the center there are about 8 hypercolumns per degree. Thus the resolution is roughly 40 times as great at the center as at the periphery. [My rough count in Figure 6 of their paper gives 52 hypercolumns over the entire 120° field, and about 20 in the central 3° field.]

Rows of color-sensitive 150-200 μm blobs spaced 500-550 μm apart lie along the center of ocular dominance columns (Hendrickson, 1985). Livingstone and Hubel, 1984a, found that blobs are inserted without disturbing the columns, and blobs are bigger moving toward the center, "but not dramatically." Blobs make up 35% of the area of Layers 2 and 3 in the central fovea. Blasdel and Salama, 1986, using TV camera input to an image-processing computer to monitor voltage-sensitive dyes, found large numbers of regular fractures that suggest the existence of a modular structure that may organize ocular dominance, slope sensitivity, and blobs.

Multi-Electrode Recordings of Patterns of Activity

Gerstein, 1985, 1986, recording from auditory areas of the cat, and Kruger, 1986, recording from visual areas V1?? of macaque??, using arrays of 20 to 30 electrodes, have found that each individual neuron fires most of the time, but that different patterns of firings result from different patterned inputs (e.g., 3 notes, or simple edges??).

These results suggest that the unit that detects a feature or recognizes an object is a whole ensemble of neurons, rather than an individual. (The fact that an appreciable percent of neurons are found in the temporal lobe sensitive to an object-class like face or hand, and that neurons differentially sensitive to different faces were found at all, suggests that these found neurons were parts of larger networks, for redundancy and fault tolerance and/or for distributed processing.)

Links at Greater Distances, between Hypercolumns and between Areas

Livingstone and Hubel, 1984b, found that blobs link selectively to other blobs up to 1 mm away, and to stripes in V2. Non-blob regions link to non-blobs. [It appears that there are on the order of 1000 modules, each structured around several blobs. This means that a major convergence in V1 drops resolution from the roughly 1000 by 1000 array of spots that a million ganglia project to Layer 4 of V1 down to 50 by 50 or so. Yet the system still appears to be at the rather low level of straight edge detection by complex cells. ??or is it getting features, eg curves, that are too subtle to find? Even the resolution at the central fovea seems quite coarse with respect to the task of combining edges into more complex features. This suggests, along with its large size, many projections, and complex columns and hypercolumns, that striate cortex may be doing a good bit of yet unidentified higher-level processing.]??give sizes (area, neurons, cols) of each visual area?

Powell, 1981, summarizes several different researchers' findings (e.g., Gatter and Powell, 1978) for different areas and species with respect to lateral links through the grey matter: After minimal lesions there is dense degeneration in all directions for 2-300 μm , then moderate degeneration for up to 2-3 mm.??give more?number and synapses? Both anatomical and physiological techniques that are more sensitive than degeneration after lesions find that lateral links within the grey matter can go long distances, up to 6 mm (Fisken et al., 1975; Creutzfeldt et al., 1977 - reported in Gilbert and Wiesel, 1983). Rockland and Lund, 1983, found periodic lattice-like patterns extending 2-3 mm in different layers. Rockland, 1983, found similar patterns extending 2-3 mm in V2 and 3.5-4 mm in V4. Gilbert and Wiesel, 1983, and Ts'o et al., 1986, have found lateral links between similarly oriented columns in nearby hypercolumns up to 2-4 mm distance (see also Gilbert, 1985).??what percent?

Allman et al., 1985, and Desimone et al., 1985, described global interactions for which mechanisms such as these are needed. In addition, the white matter contains large numbers of cortico-cortical links between different areas. There are also major radiations between different parts of the cortex, and major pathways to subcortical areas.??details;continue?and other links?

Secondary Visual Areas

Multiple secondary cortical visual areas (up to 20) have been identified, in many species. There are often major

differences between even closely related species, e.g., new world and old world monkeys (Weller and Kaas, 1981). Given the amount of still-uncharted grey matter, it is expected that this number will increase substantially.

Several researchers have mapped out the links between areas, using different techniques, in most cases with similar results. In addition (see below), there are major links to areas in the superior temporal sulcus (STS), parietal and frontal lobes, and the amygdala that appear to be involved in the highest levels of object recognition. Van Essen, 1979, 1985, described roughly 40 firm links (over 80 possible) between roughly 20 cortical visual areas in macaque (2 to 8 links to each). These organize into a hierarchical depth of roughly 4 to 6. (See also Allman and Kaas, 1974, 1975; Zeki, 1978b; Weller and Kaas, 1981). Mishkin et al., 1983, proposed two major pathways, for shape and color and for motion and spatial orientation. Thus there is relatively rich interconnectivity, but far from a complete graph where every area links directly to every other.

The following are among the major linkages found anatomically:

- . V1 (the striate cortex) links chiefly to V2 and also to V3 (in cat but not monkey these also receive inputs from the LGN) and MT (the middle temporal lobe).
- . V1 and V2, by far the largest visual areas, share a long common border.
- . V2 links chiefly to V3 and MT, and also to VP (ventral posterior), V4, and Area 8 (frontal eye fields in the prefrontal lobe).
- . V3 links to V4 and MT.
- . V4 links chiefly to IT (inferior temporal lobe), and also to MT and POa (parietal-occipital).
- . MT projects to V1, V2, V3, V4, VP, MST (medial superior temporal), and VIP (ventral intraparietal), and possibly other areas.
- . V4 and VP (ventral posterior) project to IT.
- . MST projects to 7a (in the parietal lobe) and 8.
- . There are additional tentatively identified areas and minor projections, and it is expected that more will be found.

Links are almost always 2-way, one ascending, 90% from superficial layers primarily to Layer 4, the other descending, from both superficial and deep layers, mostly to Layer 1 and/or 6. Six hierarchical levels can be counted by moving through the deepest ascending path: V1-V2-V3-MT-MST-(7a,8). (The shallowest paths are: V1-MT-MST-(7a,8) and V1-(V2,V3)-V4-(IT,8).) Desimone et al., 1985, suggested that the major path for object recognition is largely serial: V1-V2-V3-V4-IT.

Removal of V1 eliminates recognition, although sometimes primitive form perception continues. Deficits in other areas sometimes have minor effects, but can eliminate one function or another (stereopsis, form perception, color) - suggesting that each is handled in a different area.

Zeki, 1978b, summarizing a number of findings, suggested that V1 sends out different types of information to V2, V3, V4 and MT, for different types of processing. He also found successively larger receptive fields and higher-level processes, moving from V1 to V2 to V3 (for example, V2 cells are truly binocular, while V1, where cells simply prefer one eye, are not). Shipp and Zeki, 1985, found that V2 appears to send different types of information in a regular pattern to different secondary visual areas: color information to V4, information about motion to MT. [But typically each area has cells responsive to a variety of features. What percents? They might serve for contextual interactions, landmarks for merging, fault tolerance, ensembles.] Kennedy and Bullier, 1985, found that all visual areas have more projections to V2 than to V1 except for STS, which has an equal number, and also has the highest percent. They suggest that V1 projects to, and is projected to from, all areas. Rockland and Pandya, 1979, found that different afferents project to different layers in a regular pattern, and that STS links directly with V1, V2 and V3. Visual information must all come together at still-higher levels, apparently in the parietal lobe, prefrontal cortex, and amygdala - all of which are richly linked to visual areas (Jones and Powell, 1970). Seltzer and Pandya, 1978, found many such links from STS. Seltzer and Pandya, 1978, found many such links from STS. Van Essen?

Hubel and Livingstone, 1985, found that blobs in V1 link to regularly spaced patches in V2. (Gilbert, 1983, points out that cortico-cortical links between areas are usually clustered, in visual, somatosensory, auditory, and frontal areas.) Cynader et al., 1984, found with electrodes placed in a grid 300 um apart that cells are clustered for slope, eye, and motion (slope was arranged most regularly). Matsubara et al., 1984, 1985, found both anatomically and behaviorally that neurons linked in clusters are sensitive to the same slope. Local links within one area were always reciprocal; links between areas often were not (see also Hendrickson, 1985). Maunsell and Schiller, 1984,

found transient and sustained responses to gratings in V1 at 28 and 38 msec, and sustained responses in MT at 40 and in V4 at 49 msec. They concluded that faster magnocellular Y-neurons fire MT (which processes motion). ??This is a nice use of precise timings - any others, e.g. to IT and STS??

Tusa et al., 1981, examining 13 visual areas in cat, found a variety of different mappings of the image. Only V1 has a full projection of the retinal image. Moving up the hierarchy, other areas tend to have smaller projections that often do not go all the way to the periphery. Some are 2-dimensional, some map the field into 1-dimensional projections. ??say which? V2 has only 50° of the field, and at least 5 areas have limited vertical but extensive horizontal representations. Several areas have larger representations of the upper field than of the lower. Three areas have magnified representations of the peripheral field; the rest magnify, and appear to be specialized for, the center.

There appear to be two parallel streams of processing branching out of V1. One, going through MT and MST to 7a, is primarily involved with motion and spatial orientation. The other includes V4, VP, and IT; it chiefly handles shape and color. In addition, IT itself contains several areas, and links to other temporal lobe areas.

The Pathway Involved with Motion and Disparity

Although many cells in V1 and V2 are sensitive to speed as well as to slope, direction, and binocular disparity, ??percents? a higher percent of cells in MT (Maunsell and Van Essen, 1983a, b, c) are sensitive to motion, its direction and speed, and to disparity. These cells link to Layer 4b of V1, which has many motion-sensitive cells. Their receptive fields are roughly two orders of magnitude larger than V1 cells'. They show little or no response to color. Cells in MST are sensitive to direction and have very large receptive fields. ??how large? Some cells respond differently to motion when it is the object's or the perceiver's.

The Pathway Involved with Shape and Color

V1 has many cells sensitive to slope, length, width, direction, speed, binocular disparity, and color. V4 has a high percent of color-sensitive cells (Zeki, 1978a), though different researchers' estimates range from 100% to only a few. Zeki, 1983, found they are (possibly as the result of global contexts) sensitive to the perceived, not the sensed, color. ??where are angles, curves, etc; how are color, shape combined?? Many cells in V2 and VP are also sensitive to orientation and disparity. Von der Heydt et al., 1984, found cells in V2 (but not V1) responsive to illusory contours.

It seems likely that V1 is sensitive to more complex patterns than straight sloped edges, and it is necessary that the secondary visual areas process more complex patterns, including curves, textures, angles, and a variety of still more complex characteristics of objects. To quote Van Essen and Maunsell (1983, p. 574): "Several important stages of form analysis must take place between V1, where selectivity for a variety of relatively simple stimulus parameters is established, and IT and other temporal lobe regions, where some cells appear to be selective for highly complex features contained with [sic] stimuli such as a hand or a face. However, the nature of the processing that occurs at each of the stages beyond V1 remains a major enigma. The finding of orientation-sensitive cells in V2, V4 and VP and of cells sensitive for length and width is evidence that these areas convey information about form and pattern, but it seems likely that they should also be involved in higher levels of analysis, involving transformations of information that are qualitatively different from those occurring in striate cortex."

A major goal of the present model is to suggest and examine processes that might span the gap between simple edges and faces, hands, and other complex objects. Edges, colors, motion, binocular depth, texture - these are all at a simple level. Computer programs have been written to detect them, but these programs are far from recognizing complex real-world objects. How are they best combined together, building to higher levels? The running model can compare, for example, a system that simply links edges with ones that build more complex curves and shapes, and/or combine edge, color, and texture information.

Temporal Lobe and Amygdala Areas Sensitive to Complex Objects

The temporal lobe (classically considered an association area primarily involved with audition, language, and polysensory integration) has several areas of primary importance for visual recognition. These include IT (Gross et al., 1972, 1984; Rolls et al., 1977; Desimone et al., 1984), STS's middle and anterior parts (Rolls, 1984; Perrett et al., 1982, 1984), and STP (superior temporal polysensory) areas (Bruce et al., 1981). IT links to STS and STP, and also to the orbitofrontal cortex, and to the amygdala and hippocampus. Rolls, 1984, found the amygdala (to which

both IT and STS link) to be rich in cells sensitive to faces or to different aspects of face.

Gross et al., 1984, reported that most cells in IT have large (roughly 25°) receptive fields that include the center, and are selective for shape, color or texture, singly or in combination. In one sample, 13% of IT neurons did not respond to visual stimuli, 14% responded too weakly to study, and 30% responded non-selectively to 16 different test objects. The remaining 43% responded differentially to different objects. Of these, 36% were critically sensitive to shape, 15% to color, 7% to texture. Rolls et al., 1977, found that most cells did not respond to unique objects, but appeared to be selective for features like texture, color, shape. A small number (7%) responded to either faces or hands. Additional studies found cells sensitive to line drawings (square, cross, triangle, circle - Sato et al., 1980), curves (Fourier descriptors - Schwartz et al., 1983), and other complex stimuli (Richmond et al., 1983).

Larger percentages (10%, 20%, and in one of Rolls' samples 77%) of neurons in selective areas of STS fire selectively to faces, to different complex aspects of faces, and to different faces. Perrett et al., 1982, found 48 of 497 neurons responded to faces, but much less or not at all to other complex objects, geometrical designs, or gratings. Perrett et al., 1984, estimated from more extensive examinations that 3-5% of STS neurons respond to faces. Rolls, 1984, reported that 12% of amygdala neurons responded to visual stimuli, about 25% of these responding primarily to faces, and selectively to different faces. Rolls, 1985, found, for example, a neuron that responded to only one person, but stopped responding when the hairline was elevated. Gross et al., 1984, suggested that cells may be arranged in columns that concentrate on faces. Response latency ranged from about 70 to 200 msec. [do different types of cells have different latencies? What goes on between the 40 and 49 msec needed to fire neurons in MT and V4 and the 70-200 msec needed to fire some of these neurons?] [These researchers appear to feel that these are not exclusively areas for recognizing faces, but recognize a variety of complex objects. The suggestion is that whole networks of neurons are involved in the recognition of a complex object, and that some of the cells highly selective for particular faces may be more involved with symbolic, emotional and attentive reactions to the face after it has been recognized.]

Thus a number of studies by Gross, Perrett, Rolls, Bruce, Desimone, and their associates have found a small but significant percent of neurons that fire to one or another of the following: one individual face (over different expressions), any face, profiles, full face, and major parts of a face, like the two eyes. [more details? It is unclear exactly where recognition takes place, but this is a wealth of important new information.]

Human beings with lesions in both hemispheres of the temporal lobe suffer from Prosopagnosia (Damasio et al., 1982, Damasio, 1985). This is usually considered to be the inability to recognize the faces of familiar people, or even one's own face, with vision otherwise relatively normal. (Bender and Feldman, 1970, suggest that people with visual agnosias usually have other deficits.) But Damasio presents extensive evidence that other objects where fine distinctions are important, e.g., a particular make of car, a particular style of dress or shoe, a particular bird (when viewed by a bird watcher), are similarly indistinguishable. Thus this would appear to be a general inability to make use of fine details and subtle interactions needed to discriminate among similar objects at very high levels of recognition. These patients' otherwise relatively normal perception suggests that information gathered at any and every level is sent via parallel links to other levels, including higher levels.

The Structure Being Modeled, and Other Related Visual and Non-Visual Areas

The subset of the visual system that will be modeled includes the retina, lateral geniculate, striate cortex (V1), the parts of the temporal lobe involved with object recognition (IT, STS), and most of the known secondary visual areas (e.g., V2, V3, V4, MT) in between. This is considered to be the major system involved with the perceptual recognition of complex objects. Certainly the results of Gross, Perrett, Rolls, and their colleagues indicate that very high levels of recognition take place in the temporal lobe. However, since there are many links from other visual areas involved in other aspects of vision, they may be involved with object perception as well.

The areas involved in perception of motion and spatial relations (chiefly V2 and MT) that culminate in the parietal lobe may well be involved with shape and object perception. [evidence?] It is hard to see how one can be accomplished without the other. The simplest hypothesis is that evolution, to parallelize and to keep from crowding, diverged out to different regions, and then

converged back. In any case, these areas will also be used as needed. Their results will be merged into STS, a richly linked area where object recognition is known to take place.

Other than these areas sensitive to motion and spatial relations, the rest of the cortex will not be addressed in this first model. This includes other areas in the temporal and parietal lobes involved in combining information from several different sense modalities, linking perception with language and semantic memory, and handling and controlling perceptual-motor interactions, using kinesthetic feedback. In addition, areas of the prefrontal cortex are also known to respond to visual inputs; it is here that percepts get involved with higher mental processes.

The retina has major projections other than those through the LGN to the visual cortex, in particular via the superior colliculus and the pulvinar. These are involved in non-perceptual processes, chiefly the control of eye movement, pupil dilation, and focusing.??check? But Bruce et al., 1986, found that they can also be involved in perception of motion and form.

Alerting the organism, focusing attention, and the related processes that appear to be handled by the reticular activating system, may also play a crucial role in perception. The internal focusing of attention in the serial processing of a scene that Julesz, 1984, Hillyard, 1985, and others have discovered takes roughly 50-300 msec is an example of a useful perceptual process that might be modeled in an expanded system. [Note that this is too slow to play a role when the entire recognition process takes only 70-200 msec.]

The chief sub-cortical area directly involved with higher perceptual processes appears to be the amygdala. It is hypothesised that here is where drive, motivation, and emotion link with perception. These can be ignored, by simply assuming the system is operating under a cool constant drive to process information. Thus the amygdala and related systems will not be included.

It is hoped that it will be possible in later expanded models to begin to incorporate these systems, along with differential actions, both local and global, of the neurotransmitters.

Discussion and Summary

This is a large undertaking; even the first version of the model will probably need further simplifications. But it seems necessary to model an entire system, one that performs difficult tasks. Much remains unknown; but a great deal is now known. Incorporating this into a model can clarify alternative possibilities, constrain each part, suggest new experimental probes to fill in missing information, and make it possible to evaluate how well the system actually works.

Necessary Simplifications, and Possible Future Extensions

The first running program will be over-simple. Columns will have to be coded as functions that might be executed by plausible circuits, rather than actual models of the detailed circuitry. But a running program will give a firm framework for examining, and a good testbed for comparing, alternatives for the important parts. It should provide a basis on which to build, fleshing out sub-structures, comparing different parts, and changing parts as needed. Several major types of improvements and extensions should help to improve the model:

Key substructures, and in particular the microcircuitry of the columns, should be elaborated.

These are major research problems that depend on new empirical evidence interacting with tests of hypotheses as to plausible mechanisms.

The flow of information through the network of visual areas, and the transformations effected by these areas, should, as new information is gathered about living systems, be made more detailed and more brainlike.

The system that recognizes patterned objects should be combined with the system for motion and spatial orientation, to model all aspects of visual perception. Mechanisms should be added that focus attention (e.g., to notice and describe details, or to assess the import of percepts).

The visual system might fruitfully be combined with other closely related systems, for other senses, language and semantic memory, and motivation and drive. Finally, it might be embedded in a model for perceptual-motor learning.

The Need for Precise Working Models, and Criteria for Evaluating Them

Without a working model it is impossible to evaluate a theory about something so complex as perceptual recognition. It is not sufficient to describe the living system in an imprecise way, make some suggestions about plausible mechanisms that fit some aspects of the system, and show that these mechanisms can be described mathematically, or programmed. Nor should we wait until the total system is fully understood, by purely empirical means. Theories and models are essential to posing good questions and directing observation.

The following are proposed as the crucial requirements for a model:

A. *It should work, and well, at difficult problems.* Only this demonstrates that it is consistent and has real consequences, and makes possible precise comparisons between models.

B. *It should work within overall constraints of time, space, and complexity.* It must be as fast as the brain. Its components should not be implausibly large, numerous, or complex.

C. *It should work as the brain does - to the extent that this is known, and allowing for variations of the sort we see between and within species.* Where the brain's structures and processes are not known, the model should be as brain-like as possible. At the least it should not do things brains cannot, or do not, do.

D. *We should be guided by simplicity, elegance, generality, and efficiency* in the absence of firm knowledge about the brain.

E. *The model should predict new observations and experimental results.* It should be precise and clear enough in its structure and behavior to be open to experimental test, and to comparisons with other models.

F. *The model should be useful as a test-bed for evaluating and comparing possible mechanisms for specific parts of the total system.*

Summary Specification of the Proposed Model

The model outlined in this paper attempts to do what the living visual system does when it recognizes complex objects (e.g., hands, faces, parts of face) within 70 to 200 msec. This means that recognition of complex real-world objects must be achieved within roughly 20 to 200 steps.

This small number of steps must operate within the following constraints:

1. The retinal network achieves binocular maps of contrasting center-surround spots; these

are sent to the striate cortex via the lateral geniculate.

2. The striate cortex achieves maps of simple sloped edges and complex stopped edges, and possibly other higher-level features as well. These are stored in separate but complexly inter-linked columns that are arranged in hypercolumns and modules.

3. Processing continues through twenty or so visual areas, linked in a shallow hierarchy. Each appears to be specialized to a certain extent.

4. Many of the temporal lobe's individual neurons are each sensitive to different complex objects, or sets, or subparts, of objects.

5. Individual cells, microcircuits, columns, modules, and areas, are all linked in regular ways; these should be modeled as realistically as possible.

6. The total structure is built from large parallel layered arrays of probabilistic threshold elements (neurons) linked in microcircuits that converge and diverge information; processing is distributed throughout.

The structures and processes in each area are extremely complex, and only partially mapped out. They will be modeled with mechanisms that are as realistic and plausible as possible, but simplified to the point where they are tractable. Those general structural features of the brain that are known will be modeled. These include its massively parallel application of micro-modular neuron-executed processes, and constant cycling through diverging processes that disseminate information as appropriate (e.g., for global contexts and constancies), and converging processes that give successively larger receptive fields, and also combine, compare, and at least occasionally choose.

References

- Allman, J. M. and Kaas, J. H., The organization of the second visual area (VII) in the owl monkey: a second order transformation of the visual hemifield, *Brain Res.*, 1974, 76, 247-265.
- Allman, J. M. and Kaas, J. H., The dorsomedial cortical visual area: a third tier area in the occipital lobe of the owl monkey (*Aotus trivirgatus*), *Brain Res.*, 1975, 100, 473-487.
- Allman, J., Miezen, F., and McGuinness, E., Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparisons in visual neurons, *Ann. Rev. Neurosci.*, 1985, 8, 407-430.
- Arbib, M., *The Metaphorical Brain*, New York: Wiley, 1972.
- Bajcsy, R. K. and Rosenthal, D. A., Visual and conceptual focus of attention. In: *Structured Computer Vision: Machine Perception Through Hierarchical Computation Structures* (S. Tanimoto and A. Klinger, Eds.), New York: Academic Press, 1980, 133-149.
- Ballard, D. H., Cortical connections and parallel processing: Structure and function, *Behavioral and Brain Sci.*, 1986, 9, 67-120.
- Batcher, K. E., Design of a massively parallel processor, *IEEE Trans. Computers*, 1980, 29, 836-840.
- Beaulieu, C. and Colonnier, M., A comparison of the number of neurons in seven cortical areas of cat, *Soc. Neurosci. Abstr.*, 1984, 10, 473.
- Bender, M. B. and Feldman, M., The so-called "visual agnosias," *Brain*, 1972, 95, 173-186.
- Bernmond, J.-C., Delorme, C. and Quisquater, J.-J., Strategies for interconnection networks: some methods from graph theory, *Manuscript M58*, Phillips Research Lab., Brussels, 1983.
- Beurle, R. L., Properties of a mass of cells capable of regenerating pulses, *Philos. Trans. R. Soc. Lond. (Biol.)*, 1956, 240, 55-94.
- Bishop, P. O., Coombs, J. S., and Henry, G. H., Responses to visual contours: Spatio-temporal aspects of excitation in the receptive field of simple striate neurones, *J. Physiol.*, 1971, 219, 625-657. (a)
- Bishop, P. O., Coombs, J. S., and Henry, G. H., Interaction effects of visual contours on the discharge frequency of simple striate neurones *J. Physiol.*, 1971, 219, 659-687. (b)
- Bishop, P. O., Coombs, J. S., and Henry, G. H., Receptive fields of simple cells in the cat striate cortex, *J. Physiol.*, 1973, 231, 31-60.
- Blasdel, G. G. and Salama, G., Voltage-sensitive dyes reveal a modular organization in monkey striate cortex, *Nature*, 1986, 321, 579-585.
- Braitenberg, V., *On the Texture of Brains*, New York: Springer, 1977.
- Braitenberg, V., Charting the visual cortex. In: *Cerebral Cortex: Vol. 3. Visual Cortex* (A. Peters and E. G. Jones, Eds.), New York: Plenum, 1985, 379-413.
- Bruce, C. J., Desimone, R., and Gross, C. G., Visual properties of neurons in a polysensory area in the superior temporal sulcus of the macaque, *J. Neurophys.*, 1981, 46, 369-384.
- Bruce, C. J., Desimone, R., and Gross, C. G., Both striate cortex and superior colliculus contribute to visual properties of neurons in superior temporal polysensory areas of macaque monkey, *J. Neurophys.*, 1986, 55, 1057-1075.
- Bullier, J. and Henry, G. H., Laminar distribution of first order neurons and afferent terminals in cat striate cortex, *J. Neurophys.*, 1979, 42, 1271-1281.
- Cantoni, V., Ferretti, S., Levialdi, S. and Maloberti, F., A pyramid project using integrated technology. In: *Integrated Technology for Image Processing* (S. Levialdi, Ed.), London: Academic Press, 1985, 121-133.
- Cooper, L. N., Distributed memory in the central nervous system: possible test of assumptions in visual cortex. In: *The Organization of the Cerebral Cortex* (F. O. Schmitt, F. G. Worden, G. Adelman, and S. G. Dennis, Eds.), Cambridge: MIT Press, 1981, 479-503.
- Craik, K. J. W., *The Nature of Explanation*, Cambridge: Cambridge Univ. Press, 1943.
- Creutzfeldt, O. D., Garey, L. J., Kuroda, R., and Wolff, J.-R., The distribution of degenerating axons after small lesions in the intact and isolated visual cortex of the cat, *Exp. Brain Res.*, 1977, 27, 419-440.

- Crowther, W., Goodhue, J., Starr, E., Thomas, R., Milliken, W. and Blackadar, T., Performance measurements on a 128-node Butterfly parallel processor, *Proc. Int. Conf. Par. Proc.*, 1985, 531-540.
- Cynader, M., Matsubara, J. A., and Swindale, N. V., Functional topography in areas 17 and 18 of cat visual cortex, *Soc. Neurosci. Abstr.*, 1984, 10, 800.
- Damasio, A. R., Prosopagnosia, *Trends in Neurosci.*, 1985, 8, 132-135.
- Damasio, A. R., Damasio, H., and Van Hoesen, G. W., Prosopagnosia: Anatomical basis and behavioral mechanisms, *Neurol.*, 1982, 32, 331-341.
- Desimone, R., Albright, T. D., Gross, C. G., and Bruce, C. J., Stimulus selective properties of inferior temporal neurons in the macaque, *J. Neurosci.*, 1984, 2051-2062.
- Desimone, R., Schein, S. J., Moran, J. and Ungerleider, L. G., Contour, color and shape analysis beyond the striate cortex, *Vision Res.*, 1985, 25, 441-452.
- Despain, A. M. and Patterson, D. A., X-tree: a tree structured multi-processor computer architecture, *Proc. Fifth Ann. Symp. on Computer Arch.*, 1978, 144-151.
- Deutsch, J. A., *Models of the Nervous System*, New York: Wiley, 1967.
- Dodwell, P. C., *Visual pattern Recognition*, New York: Holt, 1970.
- Duff, M. J. B., CLIP4: a large scale integrated circuit array parallel processor, *Proc. 4th Int. Joint Conf. on Pat. Recog.*, 1976, 4, 728-733.
- Eccles, J. C., *The Neurophysiological Basis of Mind*, London: Oxford Univ. Press, 1952.
- Eccles, J. C., Recognition memory and consciousness. In: *Pattern Recognition Mechanisms* (C. Chagas, R. Gattas, and C. G. Gross, Ed.), Vatican City: Pontif. Acad. Sci., 1985, 309-332.
- Edelman, G. M., Group selection and phasic reentrant signaling: A theory of higher brain function. In: *The Neurosciences: Fourth Study Program* (F. O. Schmitt and F. G. Worden, Eds.), Cambridge: MIT Press, 1979, 1115-1139.
- Feldman, J. A., Dynamic connections in neural networks, *Biol. Cyber.*, 1982, 46, 27-39.
- Feldman, J. A., Four frames suffice: A provisional model of vision and space, *Behavioral and Brain Sci.*, 1985, 8, 265-289.
- Ferster, D. and LeVay, S., The axonal arborization of lateral geniculate neurons in the striate cortex of the cat, *J. Comp. Neurol.*, 1978, 182, 923-944.
- Fisken, R. A., Garey, L. J., and Powell, T. P. S., The intrinsic association and commissural connections of area 17 of the visual cortex, *Philos. Trans. R. Soc. Lond. (Biol.)*, 1975, 272, 487-536.
- Fox, G. C. and Otto, S. W., Algorithms for concurrent processors, *Physics Today*, 1984, AR-370, May, 13-20.
- Fukushima, K., Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position, *Biol. Cyber.*, 1980, 36, 193-202.
- Gatter, K. C. and Powell, T. P. S., The intrinsic connections of the cortex of area 4 of the monkey, *Brain*, 1978, 101, 513-541.
- Gerstein, G. L., Multi-electrode recordings of responses to patterned auditory stimuli??. In: *Complex Systems* (H. Haken, Ed.), Berlin: Springer, 1986.
- Gerstein, G. L., Perkel, D. H., and Dayhoff, J. E., Cooperative firing activity in simultaneously recorded populations of neurons: detection and measurement, *J. Neurosci.*, 1985, 5, 881-889.
- Gilbert, C. D., Laminar differences in receptive field properties in cat primary visual cortex, *J. Physiol.*, 1977, 268, 391-421.
- Gilbert, C. D., Microcircuitry of the visual cortex, *Ann. Rev. Neurosci.*, 1983, 6, 217-247.
- Gilbert, C. D., Horizontal integration in the neocortex, *Trends in Neurosci.*, 1985, 8, 160-165.
- Gilbert, C. D. and Wiesel, T. N., Clustered intrinsic connections in cat visual cortex, *J. Neurosci.*, 1983, 3, 1116-1133.
- Goodman, J. R. and Sequin, C. H., Hypertree: a multiprocessor interconnection topology, *IEEE Trans. Computers*, 1981, C-30, 923-933.
- Gross, C. G., Rocha-Miranda, C. E., and Bender, D. B., Visual receptive fields of neurons in inferotemporal cortex of the macaque, *J. Neurophys.*, 1972, 35, 96-111.
- Gross, C. G., Desimone, R., Albright, T. D., and Schwartz, E. L., Inferior temporal cortex as a visual integration area. In: *Cortical Integration* (F. Reinoso-Suarez and C. Ajmone-Marsan, Eds.), New York: Raven, 1984.
- Grossberg, S., Outline of a theory of brightness, color, and form perception. In: *Trends in Mathematical Psychology* (E. Degreef and J. van Bugenhout, Eds.), Amsterdam: North-Holland, 1984.
- Hammond, P. and MacKay, D. M., Differential responsiveness of simple and complex cells in cat striate cortex to visual texture, *Exp. Brain res.*, 1977, 30, 275-296.
- Hanson, A. R. and Riseman, E. M., Visions: A Computer System for Interpreting Scenes. In: *Computer Vision Systems* (A. R. Hanson and E. M. Riseman, Eds.), New York: Academic Press, 1978, 303-334.
- Hebb, D. O., *The Organization of Behavior*, New York: Wiley, 1949.
- Hendrickson, A. E., Dots, stripes and columns in monkey visual cortex, *Trends in Neurosci.*, 1985, 8, 400-405.
- Henry, G. H., Dreher, B. and Bishop, P. O., Orientation specificity of cells in cat striate cortex, *J. Neurophys.*, 1974, 37, 1394-1409.
- Hillyard, S. A., Electrophysiology of human selective attention, *Trends in Neurosci.*, 1985, 8, 453-456.
- Hoffman, A. J. and Singleton, R. R., On Moore graphs with diameter 2 and 3, *IBM J. Res. Devel.*, 1960, 4, 497-504.
- Hopfield, J. J., Neural networks and physical systems with emergent collective computational abilities, *Proc. Nat. Acad. Sci.*, 1982, 79, 2554-2558.
- Hubel, D. H., Exploration of the primary visual cortex, 1955-78, *Nature*, 1982, 299, 515-524.
- Hubel, D. H. and Livingstone, M. S., Complex-unoriented cells in a sub-region of primate area 18, *Nature*, 1985, 315, 325-327.
- Hubel, D. H. and Wiesel, T. N., Receptive fields of single neurones in the cat's striate cortex, *J. Physiol.*, 1959, 148, 574-591.
- Hubel, D. H. and Wiesel, T. N., Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, *J. Physiol.*, 1962, 160, 106-154.
- Hubel, D. H. and Wiesel, T. N., Receptive fields and functional architecture of monkey striate cortex, *J. Physiol.*, 1968, 195, 215-243.
- Hubel, D. H. and Wiesel, T. N., Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey, *J. Comp. Neurol.*, 1972, 146, 421-450.
- Hubel, D. H. and Wiesel, T. N., Ferrier Lecture: Functional architecture of macaque monkey visual cortex, *Philos. Trans. R. Soc. Lond. (Biol.)*, 1977, 198, 1-59.
- Imase, M. and Itoh, M., Design to minimize diameter on building-block network, *IEEE Trans. Computers*, 1981, 30, 439-442.
- Jones, E. G. and Powell, T. P. S., An anatomical study of converging sensory pathways within the cerebral cortex of monkey, *Brain*, 1970, 93, 793-820.
- Julesz, B., Textons, the elements of texture perception, and their interactions, *Nature*, 1981, 290, 91-97.
- Julesz, B., A brief outline of the texton theory of human vision, *Trends in Neurosci.*, 1984, 7, 41-45.
- Kennedy, H. and Bullier, J., A double-labeling investigation of the afferent connectivity to cortical areas V1 and V2 of the macaque monkey, *J. Neurosci.*, 1985, 5, 2815-2830.
- Kohonen, T., clustering, taxonomy and topological maps of patterns, *Proc. Int. Joint Conf. on Pat. Recog.*, 1982, 6, 114-128.
- Konorski, J., *Integrative Activity in the Brain*, Chicago: Univ. of Chicago Press, 1967.
- Kruger, J., Multi-electrode recording of patterned responses to visual stimuli??. In: *Complex Systems* (H. Haken, Ed.), Berlin: Springer, 1986.
- Kuck, D. J., Lawrie, D., Cytron, R., Sameh, A. and Gajski, D., The architecture and programming of the Cedar system, *Proc. 1983 LASL Workshop on Vector and Par. Proc.*, 1983.
- Kuffler, S. W., Discharge patterns and functional organization of the mammalian retina, *J. Neurophys.*, 1953, 16, 37-68.

- Kuffler, S. W., Nicholls, J. G. and Martin, A. R., *From Neuron to Brain*, Sunderland, Mass.: Sinauer, 1984.
- LeVay, S., Connolly, M., Houde, J., and Van Essen, D. C., The complete pattern of ocular dominance stripes in the striate cortex and visual field of the macaque monkey, *J. Neurosci.*, 1985, 5, 486-501.
- Leventhal, A. G., Retinal design of the striate cortex. In: *Models of Visual Cortex* (D. Rose and V. G. Dobson, Eds.), Chichester: Wiley, 1985, 380-389.
- Levick, W. R. and Thibos, L. N., Analysis of orientation bias in cat retina, *J. Physiol.*, 1982, 329, 243-261.
- Levine, M. D. and Leemet, J., A method for nonpurposive picture segmentation, *Proc. Int. Joint Conf. on Pat. Recog.*, 1976, 3, 494-498.
- Li, Z. N. and Uhr, L., Pyramidal Algorithms for Analysis of House Images, *Proc. Intersoc. Conf. on Artif. Intell. Applics.*, 1985, (a)
- Li, Z. N. and Uhr, L., Comparative Timings for a Neuron Recognition Program on Serial and Pyramid Computers, *Proc. Workshop on Computer Arch. for Pat. Anal. and Image Data Base Management*, IEEE Computer Society Press, 1985, 99-106. (b)
- Li, Z. N. and Uhr, L., A pyramidal approach for the recognition of neurons using key features, *Pat. Recog.*, 1986, 19, 55-62.
- Livingstone, M. S. and Hubel, D. H., Anatomy and physiology of a color system in the primate visual cortex, *J. Neurosci.*, 1984, 4, 309-356. (a)
- Livingstone, M. S. and Hubel, D. H., Specificity of intrinsic connections in primate visual cortex, *J. Neurosci.*, 1984, 4, 2830-2835. (b)
- Lougheed, R. M. and McCubbery, D. L., Multi-processor architectures for machine vision and image analysis, *Proc. Int. Conf. Par. Proc.*, 1985, 493-497.
- MacKay, D. M., Toward an information-flow model of human behavior, *Brit. J. Psychol.*, 1956, 47, 30-43.
- MacKay, D. M., The significance of 'feature sensitivity,' In: *Models of Visual Cortex* (D. Rose and V. G. Dobson, Eds.), Chichester: Wiley, 1985, 47-53.
- Malpelli, J. G., Activity of cells in area 17 of the cat in absence of input from layer A of lateral geniculate nucleus, *J. Neurophys.*, 1983, 49, 595-610.
- Marr, D., *Vision*, San Francisco: Freeman, 1982.
- Martin, K. A. C., Neuronal circuits in cat striate cortex. In: *Cerebral Cortex: Vol. 2. Functional Properties of Cortical Cells* (E. G. Jones and A. Peters, Eds.), New York: Plenum, 1984, 241-284.
- Matsubara, J. A., Cynader, M., and Swindale, N. Y., Physiological correlates of the anatomical connections within and between areas 17 and 18 of cat visual cortex, *Soc. Neurosci. Abstr.*, 1984, 10, 800.
- Matsubara, J. A., Cynader, M., Swindale, N. Y., and Stryker, M. P., Intrinsic projections within visual cortex: Evidence for orientation-specific local connections, *Proc. Nat. Acad. Sci.*, 1985, 82, 935-939.
- Maunsell, J. H. R., and Van Essen, D. C., The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey, *J. Neurosci.*, 1983, 3, 2563-2586. (a)
- Maunsell, J. H. R., and Van Essen, D. C., Functional properties of neurons in middle temporal visual area of macaque monkey: I Selectivity for direction, speed and orientation, *J. Neurophys.*, 49, 1983, 1127-1147. (b)
- Maunsell, J. H. R., and Van Essen, D. C., Functional properties of neurons in middle temporal visual area of macaque monkey: II Binocular interaction and sensitivity to binocular disparity, *J. Neurophys.*, 49, 1148-1167. (c)
- Maunsell, J. H. R., and Schiller, P. H., Evidence for the segregation of parvo- and magnocellular channels in the visual cortex of macaque monkey, *Abst. Soc. Neurosci.*, 1984, 10, 510.
- Maxwell, N., Methodological problems of neuroscience. In: *Models of Visual Cortex* (D. Rose and V. G. Dobson, Eds.), Chichester: Wiley, 1985, 11-21.
- McCulloch, W. S. and Pitts, W., A logical calculus of the ideas imminent in nervous activity, *Bull. Math. Biophys.*, 1943, 5, 115-133.
- Michael, C. R., Columnar organization of color cells in monkey striate cortex, *J. Neurophys.*, 1981, 46, 587-604.
- Michael, C. R., Serial processing of color in the monkey's striate cortex. In: *Models of Visual Cortex* (D. Rose and V. G. Dobson, Eds.), Chichester: Wiley, 1985, 301-309.
- Mishkin, M., Ungerleider, L. G., and Macko, K. A., Object vision and spatial vision: two cortical pathways, *Trends in Neurosci.*, 1983, 6, 414-417.
- Nevatia, R., Image segmentation. In: *Handbook of Pat. Recog. and image Processing* (T. Y. Young and K.-S. Fu, Eds.), New York: Academic, 1986, 215-231.
- Neveu, C. F., Dyer, C. R. and Chin, R. T., Object recognition using Hough pyramids, *Proc. IEEE Conf. Computer Vision and Pat. Recog.*, 1985, 328-333.
- Perrett, D. I., Rolls, E. T., and Caan, W., Visual neurones responsive to faces in the monkey temporal cortex, *Exp. Brain Res.*, 1982, 47, 329-342.
- Orban, G. A., *Neuronal Operations in the Visual Cortex*, Berlin: Springer, 1984.
- Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., and Jeeves, M. A., Neurones responsive to faces in the temporal cortex: studies of functional organization, sensitivity to identity and relation to perception, *Human Neurobiol.*, 1984, 3, 197-208.
- Peterson, J. C., Tuazon, J. O., Lieberman, D. and Pniel, M., The Mark III Hypercube-ensemble concurrent computer, *Proc. Int. Conf. Par. Proc.*, 1985, 71-73.
- Powell, T. P. S., Certain aspects of the intrinsic organization of the cerebral cortex. In: *Brain Mechanisms and Perceptual Awareness* (O. Pompeiano and C. Ajmone Marsan, Eds.), New York: Raven, 1981, 1-19.
- Rashevsky, N., *Math. Biophys.*, Chicago: Univ. of Chicago Press, 1938. (3d Edition, 1960).
- Reddaway, S. F., DAP - a flexible number cruncher, *Proc. 1978 LASL Workshop on Vector and Par. Proc.*, Los Alamos, 1978, 233-234.
- Richmond, B. J., Wurtz, R. H., and Sato, T., Visual responses of inferior temporal neurons in awake rhesus monkey, *J. Neurophys.*, 1983, 50, 1415-1432.
- Riganati, J. P. and Schneck, P. B., Supercomputing, *Computer*, 1984, 17, 97-113.
- Rockland, K. S., Lattice-like intrinsic neural connections in macaque prestriate visual cortex, *Neurosci. Abstr.*, 1983, 9, 476.
- Rockland, K. S. and Lund, J. S., Intrinsic laminar lattice connections in primate visual cortex, *J. Comp. Neurol.*, 1983, 216, 303-318.
- Rockland, K. S. and Pandya, D. N., Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey, *Brain Res.*, 1979, 179, 3-20.
- Rockel, A. J., Hiorns, R. W., and Powell, T. P. S., The basic uniformity in structure of the neocortex, *Brain*, 1980, 103, 221-244.
- Rodieck, R. W., Visual pathways, *Ann. Rev. Neurosci.*, 1979, 2, 193-225.
- Rolls, E. T., Neurons in the cortex of the temporal lobe and in the amygdala of the monkey with responses selective for faces, *Human Neurobiol.*, 1984, 3, 209-222.
- Rolls, E. T., Neuronal activity in relation to the recognition of stimuli in the primate. In: *Pattern Recognition Mechanisms* (C. Chagas, R. Gattas, and C. G. Gross, Eds.), Vatican City: Pontif. Acad. Sci., 1985, 203-213.
- Rolls, E. T., Judge, S. J., and Sanghera, M. K., Activity of neurones in the inferotemporal cortex of the alert monkey, *Brain Res.*, 1977, 130, 229-238.
- Rosenblatt, F., *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*, Washington: Spartan, 1962.
- Rosenfeld, A., Pyramids: multiresolution image analysis, *Proc. Third Scand. Conf. on Image Anal.*, July, 1983, 23-28.
- Rosenfeld, A. (Ed.), *Multiresolution Image Processing and Analysis*, New York: Springer-Verlag, 1984.
- Rumelhart, D. E., Hinton, G. E. and Williams, R. J., Learning internal representations by error propagation, In: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 1: Foundations* (D. E. Rumelhart and J. L. McClelland, Eds.), Cambridge: MIT Press, 1986.
- Sato, T., Kawamura, T., and Iwai, E., Responsiveness of inferotemporal single units to visual pattern stimuli in monkeys performing discrimina-

- tion, *Exp. Brain Res.*, 1980, 38, 313-319.
- Schaefer, D. H., A pyramid of MPP processing elements - experiences and plans, *Proc. 18th Int. Conf. on Syst. Sci.*, Honolulu, 1985.
- Schmitt, L., The use of a network representation of visual knowledge in a hierarchically structured vision system, *Unpubl. Ph.D. Diss.*, Comp. Sci. Dept., University of Wisconsin, 1981.
- Schwartz, E. L., Desimone, R., Albright, T. D. and Gross, C. G., Shape recognition and inferior temporal neurons, *Proc. Nat. Acad. Sci.*, 1983, 80, 5776-5778.
- Selfridge, O. G., Pattern recognition and modern computers, *Proc. West. Joint Comp. Conf.*, 1955, 7, 91-93.
- Seltzer, B. and Pandya, D. N., Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey, *Brain Res.*, 1978, 149, 1-24.
- Shaw, D. W., Organization and operation of a massively parallel machine. In: *Computers and Technology* (G. Rabat, Ed.), Amsterdam: North-Holland, 1986.
- Shipp, S. and Zeki, S., Segregation of pathways leading from area V2 to areas V4 and V5 of macaque monkey visual cortex, *Nature*, 1985, 315, 322-325.
- Siegel, H. J., *Interconnection Networks for Large Scale Parallel Processing*. Lexington, Mass: Lexington, 1984.
- Sillito, A. M., The contribution of inhibitory mechanisms to the receptive field properties of neurones in the striate cortex of the cat, *J. Physiol.*, 1975, 250, 305-329.
- Sillito, A. M., Functional considerations of the operation of GABAergic inhibitory processes in the visual cortex. In: *Cerebral Cortex: Vol. 2. Functional Properties of Cortical Cells* (E. G. Jones and A. Peters, Eds.), New York: Plenum, 1984, 91-117.
- Sillito, A. M., Salt, T. E., and Kemp, J. A., Modulatory and inhibitory processes in the visual cortex, *Vision Res.*, 1985, 25, 375-381.
- Sloan, K., A system for world model driven recognition of natural scenes, *Unpubl. Ph.D. Diss.*, Comp. Sci. Dept., Univ. of Pennsylvania, 1977.
- Sterling, P., Microcircuitry of the cat retina, *Ann. Rev. Neurosci.*, 1983, 6, 149-185.
- Sterling, P., Freed, M., and Smith, R. G., Microcircuitry and functional architecture of the cat retina, *Trends in Neurosci.*, 1986, 9, 186-194.
- Stone, J. and Dreher, B., Projections of X- and Y-cells of the cat's lateral geniculate nucleus to areas 17 and 18 of visual cortex, *J. Neurophys.*, 1973, 36, 551-567.
- Stone, J., Dreher, B., and Leventhal, A., Hierarchical and parallel mechanisms in the organization of visual cortex, *Brain Res. Revs.*, 1979, 180, 345-394.
- Swindale, N. V. and Cynader, M. S., Local interactions determine orientation and direction selectivity of neurons in cat visual cortex, *Soc. Neurosci. Abstr.*, 1984, 10, 471.
- Szentogathai, J., Functional anatomy of the visual centers as cues for pattern recognition concepts. In: *Pattern Recognition Mechanisms* (C. Chagas, R. Gattas, and C. G. Gross, Ed.), Vatican City: Pontif. Acad. Sci., 1985, 39-52.
- Szentogathai, J. and Arbib, M. A., Conceptual models of neural organization, *Neurosci. Res. Program Bull.*, 1974, 12, 307-510.
- Tanimoto, S. L., Regular Hierarchical Image and Processing Structures in Machine Vision. In: *Computer Vision Systems*, (A. R. Hanson and E. M. Riseman, Eds.), New York: Academic Press, 1978, 165-174.
- Tanimoto, S. L., A pyramidal approach to parallel processing, *Proc. 10th Int. Symp. on Computer Arch.*, Stockholm, 1983, 372-378.
- Tanimoto, S. L., A hierarchical cellular logic for pyramid computers, *J. Parallel and Distributed Computing*, 1984, 1, 105-132.
- Tanimoto, S. L., An approach to the iconic/symbolic interface. In: *Integrated Technology for Image Processing* (S. Levialdi, Ed.), London: Academic Press, 1985, 31-38.
- Tanimoto, S. L. and Klinger, A., (Eds.), *Structured Computer Vision: Machine Perception Through Hierarchical Computation Structures*, New York: Academic Press, 1980.
- Toyama, K., Maekawa, K., and Takeda, T., Convergence of retinal inputs onto visual cortical cells. I. A study of the cells monosynaptically excited from the lateral geniculate body, *Brain Res.*, 1977, 137, 207-220.
- Ts'o, D. Y., Gilbert, C. D. and Wiesel, T. N., Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis, *J. Neurosci.*, 1986, 6, 1160-1170.
- Tusa, R. J., Palmer, L. A., and Rosenquist, A. C., Multiple cortical areas: visual field topology in the cat. In: *Cortical Sensory Areas: Vol II: Multiple Visual Areas* (C. Woolsey, Ed.), Clifton N. J.: Humana, 1981, 1981, 1-31.
- Uhr, L., Layered "recognition cone" networks that preprocess, classify, and describe. *IEEE Trans. on Computers*, 1972, 21, 758-768.
- Uhr, L., A model of form perception and scene description, *Comput. Sci. Dept. Tech. Report #231*, University of Wisconsin, 1974.
- Uhr, L., "Recognition Cones" that perceive and describe scenes that move and change over time. *Proc. 3rd Int. Joint Conf. on Pat. Recog.*, San Diego, 1976.
- Uhr, L., "Recognition cones" and some test results; the imminent arrival of well-structured parallel-serial computers; positions, and positions on positions. In: *Computer Vision Systems* (A. Hanson and E. Riseman Eds.), New York: Academic Press, 1978, pp. 363-372.
- Uhr, L., Chapter 1: Parallel-serial variable resolution perceptual systems. In: *Structured Computer Vision* (S. Tanimoto and A. Klinger Eds.), New York: Academic Press, 1980.
- Uhr, L., *Algorithm-Structured Computer Arrays and Networks: Architectures and Processes for Images, Percepts, Models, Information*. New York: Academic Press, 1984.
- Uhr, L., Pyramid Multi-Computers, and Extensions and Augmentations, In: *Algorithmically Specialized Parallel Computers*, (L. Snyder, L. H. Jamieson, D. B. Gannon, and H. J. Siegel, Eds.), New York: Academic Press, 1985, 177-186.
- Uhr, L., Constructing multi-level multi-computer networks. In: *Evaluating Multi-Computers for Image Processing* (L. Uhr, K. Preston, S. Levialdi, and M. J. B. Duff Eds.), New York: Academic Press, 1986.
- Uhr, L., *Multi-Computer Architectures for Artificial Intelligence*, New York: Wiley, 1986 (in press).
- Uhr, L. and Douglass, R., A parallel-serial recognition cone system for perception *Pat. Recog.*, 1979, 11, 29-40.
- Uttley, A. M., Conditional probability computing in a nervous system. In: *Mechanization of Thought Processes*, London: Her Majesty's Stationery Office, 1959, 119-132.
- Van Essen, D. C., Visual areas of the mammalian cerebral cortex, *Ann. Rev. Neurosci.*, 1979, 2, 227-263.
- Van Essen, D. C., Functional organization of primate visual cortex. In: *Cerebral Cortex: Vol. 3. Visual Cortex* (A. Peters and E. G. Jones, Eds.), New York: Plenum, 1985, 259-329.
- Van Essen, D. C. and Maunsell, J. H. R., Hierarchical organization and functional streams in the visual cortex, *Trends in Neurosci.*, 1983, 6, 570-575.
- Von der Heydt, R., Peterhans, E., and Baumgartner, G., illusory contours and cortical neuron responses, *Science*, 1984, 224, 1260-1262.
- von der Malsburg, C., Self-organization of orientation sensitive cells in the striate cortex, *Kybernetik*, 1973, 14, 85-100.
- Weller, R. E. and Kaas, J. H., Cortical and subcortical connections of visual cortex in primates. In: *Cortical Sensory Areas: Vol II: Multiple Visual Areas* (C. Woolsey, Ed.), Clifton N. J.: Humana, 1981, 121-133.
- White, E. L. and Hersch, S. M., A quantitative study of thalamocortical and other synapses involving the apical dendrites of corticothalamic projection cells in mouse Smi cortex, *J. Neurocytol.* 1982, 11, 137-157.
- Williams, D. R., Seeing through the photoreceptor mosaic, *Trends in Neurosci.*, 1986, 9, 193-198.
- Winfield, D. A., Gatter, K. C., and Powell, T. P. S., An electron microscope study of the types and proportions of neurons in the cortex of the motor and visual areas of the cat and rat, *Brain*, 1980, 103, 245-258.
- Young, J. Z., *A Model of the Brain*. London: Oxford Univ. Press, 1964.
- Zeki, S. M., Uniformity and diversity of structure and function in rhesus monkey prestriate visual cortex, *J. Physiol.*, 1978, 277, 273-290. (a)

Zeki, S. M., Functional specialization in the visual cortex of the rhesus monkey, *Nature*, 1978, 273, 423-428. (b)

Zeki, S. M., Color coding in the cerebral cortex: the reaction of cells in monkey visual cortex to wavelengths and colours, *Neurosci.*, 1983, 9, 741-765.

Table 1. Some of the Major Facts about Object Recognition, and the Model's Way.

Facts About Visual Perception	The Model
Neurons: probabilistic threshold functions	Same, but simpler
Synapses: ?? msec electrical, 1 ms chemical	1 ms per threshold neuronal op
Conduction: slow analog, fast digital	1 ms per threshold neuronal op
Neurotransmitters, neuropeptides	Different types of global controls
Input to 2-D retina (3000x3000)	2-D retinal array input (64x64 to 512x512)
6M shape-sensitive cones, three colors	Input and process 3 colors
100M light, motion-sensitive rods	Input and process grey-scale
many more processors packed in the center	Regular mapping
Cone->bipolar->ganglion (1000x1000),	Various multi-layer methods
horizontal, amacrine gather, converge,	converging to get spots,
resulting in surrounded spot detectors	(including some that simulate closely).
Ganglia->LGN->V1	Links as in brain,
(diverge in neurons, converge in columns)	but smaller size, fewer
6 layers in cortex grey matter,	Several converge-diverge layers,
arranged in columns, and hypercolumns	searching for good structures that give desired behavior
Cortico-cortical white matter	Additional direct links
V1 Columns have neurons firing to spot,	Several combinations of
to simple sloped edge,	known structs and functs
to complex same-sloped edge	??use sizes, more detail of column
Layers, neuron types, etc.??	Modeled where known functions
Hypercolumn rows of columns,	Set of slope detectors
each tuned at 10^0 different slope	10^0 or other apart
each eye has adjacent row	each eye has adjacent row
Total 50 hypercolumns; 20 in 3^0	Regular mapping (50x50 in toto)
Same slope linked between hypercolumns	Same slope linked between hypercolumns
Different mappings in different areas	Constant resolution
Blobs interspersed, and Module structure	Blobs interspersed, and Module structure
Major path for Form, Color	Successively higher-level transforms
Major path for Motion, Spatial orientation	Successively higher-level transforms
Some pathways: V1 -> (V2, V3, MT)	Diverge to parallel paths,
V2 -> (V3, VP, V4)	as indicated by brain
V3 -> (V4, MT)	"
V4 -> (IT, POa, 8)	"
MT -> (V1, V2, V3, V4, VP, MST)	"
IT and STS sensitive to face, faces, etc.	Highest-level transforms should do the same
Neurons are part of networks	Transforms use networks

Rough, Incomplete Sketch of the Overall Structure:

retina[cones->(horizontals,bipolars)->(amacrine,ganglia)]->LGN->
->V1[L4->(L5,L3,L2),L6->LGN]->V2->V3->V4->(IT,MT)->STS

Aspects of the Larger System That Will Not Be Included in the Model

parietal, temporal combine information; superior colliculus, pulvinar, etc., move eye;
areas for combining senses, perceptual-motor control; prefrontal; higher mental processes;
amygdala and drive, motivation, emotion; reticular activating system;
??Separate table: Speeds, sizes, percents, loops, hierarch depth, etc.

