Parallel visual computation

Dana H. Ballard*, Geoffrey E. Hinton† & Terrence J. Sejnowski‡

* Department of Computer Science, University of Rochester, Rochester, New York 14627, USA
† Department of Computer Science, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213, USA
‡ Department of Biophysics, The Johns Hopkins University, Baltimore, Maryland 21218, USA

The functional abilities and parallel architecture of the human visual system are a rich source of ideas about visual processing. Any visual task that we can perform quickly and effortlessly is likely to have a computational solution using a parallel algorithm. Recently, several such parallel algorithms have been found that exploit information implicit in an image to compute intrinsic properties of surfaces, such as surface orientation, reflectance and depth. These algorithms require a computational architecture that has similarities to that of visual cortex in primates.

UNTIL recently, the profound difficulty of vision was not fully appreciated, in part because we humans are so good at seeing. We can recognize objects in images very fast—within a few hundred milliseconds—and without appreciable effort. It therefore came as a surprise in the 1960s and the early 1970s when pioneering workers discovered that automated image interpretation requires an enormous amount of computation and that it is very hard to find simple features in real images that allow objects to be separated from one another and recognized[1,2].

Recent work in computer vision has been strongly influenced by two factors. The first is an increased awareness of the architecture of biological visual systems combined with a realization that parallel processing may be essential for competent vision systems because the serial architecture of conventional digital computers is too inefficient to deliver the massive amount of computation required[3–7]. The second factor is a much better understanding of how the domain of two-dimensional intensity arrays is related to the domain of three-dimensional objects. In an image of an object, each element in the intensity array depends on the illumination and on the 'intrinsic' properties of the three-dimensional surface patch being imaged, such as its orientiation relative to the viewer, its orientation relative to light sources, and its reflectance. When these optical constraints are combined with plausible assumptions about the nature of physical surfaces, it is possible to interpret the information that is encrypted in an intensity array in terms of intrinsic properties of three-dimensional surfaces[8–12].

This review discusses a parallel approach to visual computation and its relevance to visual processing in the cerebral cortex of primates.

Parallel architectures

The computational problems in vision are best understood at the early stage of visual processing where the inputs and the goal of the computation are known. To implement a computational theory, both the representation of the information and a well defined algorithm are needed. The hardware of a machine strongly affects the degree of difficulty in implementing and running an algorithm. The examples presented here were run on conventional digital computers simulating parallel machines, because the flexibility of a general-purpose computer was useful in exploring a variety of parallel architectures without making a commitment to any particular one.

Conventional digital computers use von Neumann architecture in which a single central processor performs sequential operations on data structures (including programs) stored in a general-purpose memory. Although central processors are very fast today, their sequentiality seriously limits the rate at which computations can be performed. The dramatic decrease in the cost of logic circuits has made parallel architectures attractive if the problem can be solved with a parallel algorithm.

In a parallel computer many separate processing units operate in parallel and exchange information through a communication network. The parallel architectures that are now being studied can be broadly classified according to the computational power of the processing units and the type of information that is exchanged between them[13,14]. The most complex, and the most difficult to analyse, is 'message-passing' parallelism, in which symbolic messages are passed between units that are as powerful as von Neumann machines[15].

A much simpler type of parallel architecture is based on processing units that exchange only approximate real values and can perform only simple arithmetic using limited memory. Models of neural networks are frequently of this 'value-passing' type, with the average firing rate of a neurone interpreted as the transmitted value. The power of this architecture lies in its ability to propagate information simultaneously along private links between units. However, the required connectivity and the way the transmitted values are used are different for each problem, so the flexibility of a general-purpose computer is lost and a formidable wiring problem is created. In this review we explore how value-passing parallelism can be applied in vision.

Relaxation

The goal of many problems in vision is to find the optimal interpretation of an image consistent with known optical and geometrical constraints. The difficulty is that an enormous number of constraints must be simultaneously satisfied. One way to solve such large optimization problems is to implement the constraints as excitatory and inhibitory links between processing units and to allow the units, whose values represent the physical properties of objects, to iteratively approach a self-consistent solution[16–23].

The classical problem of finding the shape of a soap film bounded by a non-planar wire hoop illustrates how a global solution can be achieved through successive local interactions. Imagine that the soap film is viewed from above and that its height is represented by a number in each cell of a two-dimensional array. The wire hoop fixes the heights around the edge, but an interior height is only constrained by being equal to the average of its neighbours. Regardless of the initial assignment of the interior heights, the correct shape of the soap film can be computed by iteratively assigning to each interior cell the average of its neighbours, a process called relaxation.

Just as the balance of forces acting on a piece of soap film act as physical constraints on the solution, so the optical constraints of image formation and the general nature of the surfaces causing the image restrict the plausible solutions that are consistent with the intensity array. The boundary condition for the computation is the raw input—that is, the whole intensity array—just as the height of the wire hoop is the boundary condition for the soap-film problem. The soap-film analogy is

0028-0836/83/440021—06$01.00

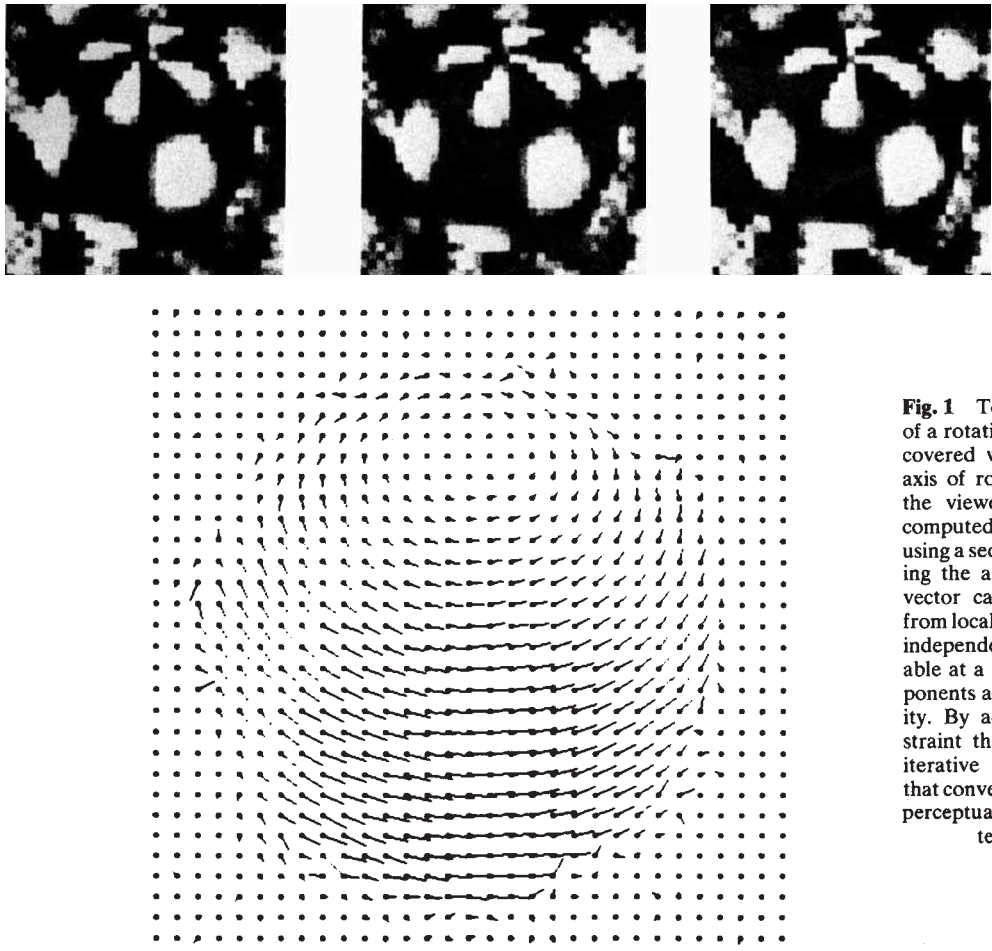

Fig. 1 Top, three successive frames of a rotating sphere whose surface is covered with a pattern and whose axis of rotation is inclined towards the viewer. Bottom, flow pattern computed by a relaxation algorithm using a sequence of 32 frames including the above three. Each velocity vector cannot be computed solely from local information since only one independent measurement is available at a given point, and two components are needed to specify velocity. By adding the additional constraint that the flow is smooth, an iterative algorithm can be derived that converges robustly to the correct perceptual interpretation[26]. (Courtesy B. K. P. Horn.)

helpful in conveying the general idea behind value-passing computation, but as we shall see, parallel visual computation can differ from the soap-film example in important ways.

Surface orientation

The goal of early visual processing is to recover intrinsic properties like surface orientation[6,24]. Such properties cannot be deduced from the local intensities alone because there are infinitely many possible combinations of orientation, reflectance and lighting that could explain any single intensity value. So, in addition to the optical constraints that relate image intensities to combinations of intrinsic parameters, the interpretation process must also use plausible assumptions about the world that constrain the intrinsic parameters at neighbouring locations. For example, the orientation and reflectance of a surface at one point strongly affect its *probable* orientation and reflectance at neighbouring points, though they do not definitely rule out any of the possibilities.

Because it must use constraints that are plausible but not definitive, the interpretation process in early vision has the form of a massive best-fit search whose objective is to find sets of intrinsic parameters that are compatible with the optical constraints and satisfy as well as possible the plausible assumptions that link intrinsic parameters at neighbouring points.

Horn's analysis of the recovery of surface orientation from the shading in an image is one of the earliest examples of how the physics of imaging can constrain the three-dimensional interpretation of an image[24,25]. If the reflectance function of the surface and the direction of illumination are known, then an iterative 'shape from shading' algorithm can fill in the most likely surface when given the surface shape along the boundary of the object (Fig. 2). The result of the computation is an intrinsic image; that is, a map of an intrinsic property (in this case surface orientation) in register with the image[6]. The same type of algorithm can also be used to compute the intrinsic image of local velocity[26], as shown in Fig. 1.

Stereo fusion

In addition to being able to recover the surface orientation from the shading of a bounded surface, the human visual system can recover the relative depth of surface markings by fusing binocular images. Julesz has demonstrated that when each eye is presented with the same pattern of random dots, except that one region in one eye is horizontally displaced, then this displaced region is perceived as a surface with a different depth from its surround[27]. Because all the dots look the same, it is difficult to decide which dot in one image corresponds to which dot in the other image, but this correspondence must be decided to perceive depth.

In the Marr–Poggio cooperative algorithm for stereopsis, each unit stands for a hypothesis about the correspondence of a particular pair of dots, and it therefore represents a patch of surface at a particular depth[28]. There are excitatory interactions between neighbouring units with the same depth to ensure continuity of surface, and inhibitory interactions between units that represent different depths at the same image location to ensure that depth assignments are unique; if the sum of all the inputs to a unit from the two images and from local interactions is above a threshold, the value of the unit is set to 1, and otherwise to 0. During the relaxation various combinations of depth assignments are tried and the network eventually 'locks' into a consistent solution in a way that resembles our perceptual experience when we fuse a random-dot stereogram[27].

There are important qualitative differences between the shape-from-shading problem and stereo fusion. Like the soap-film problem, the shape-from-shading algorithm searches a space that has the character of a single hill because each step brings the system closer to a unique solution, so that convergence is guaranteed. In contrast, the stereo-fusion algorithm has a search space that resembles a mountain range: even if the search can be made to climb towards progressively better interpretations, the final solution may be only a local optimum rather

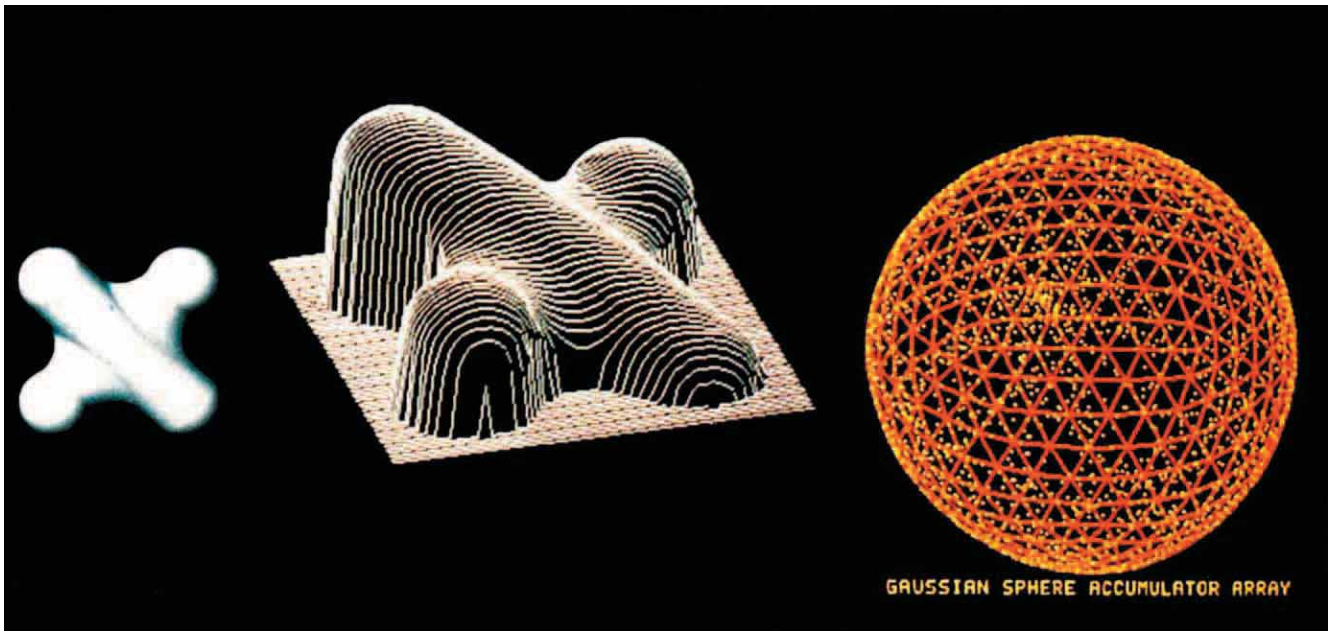

Fig. 2 Demonstration of an algorithm that generates both the shape of an object and the direction of illumination given only a shaded image of the object. The input to the algorithm is the image shown on the left. The output of a relaxation algorithm that was used to compute the shape is shown in the centre. During the computation the current estimate of the surface normals was used to compute an estimate for the direction of illumination in a two-dimensional space of direction units (right). The direction with highest activation in turn was used to obtain a better estimate of the surface normals, and the process was repeated until the estimates converged[60]. (Courtesy C. M. Brown.)

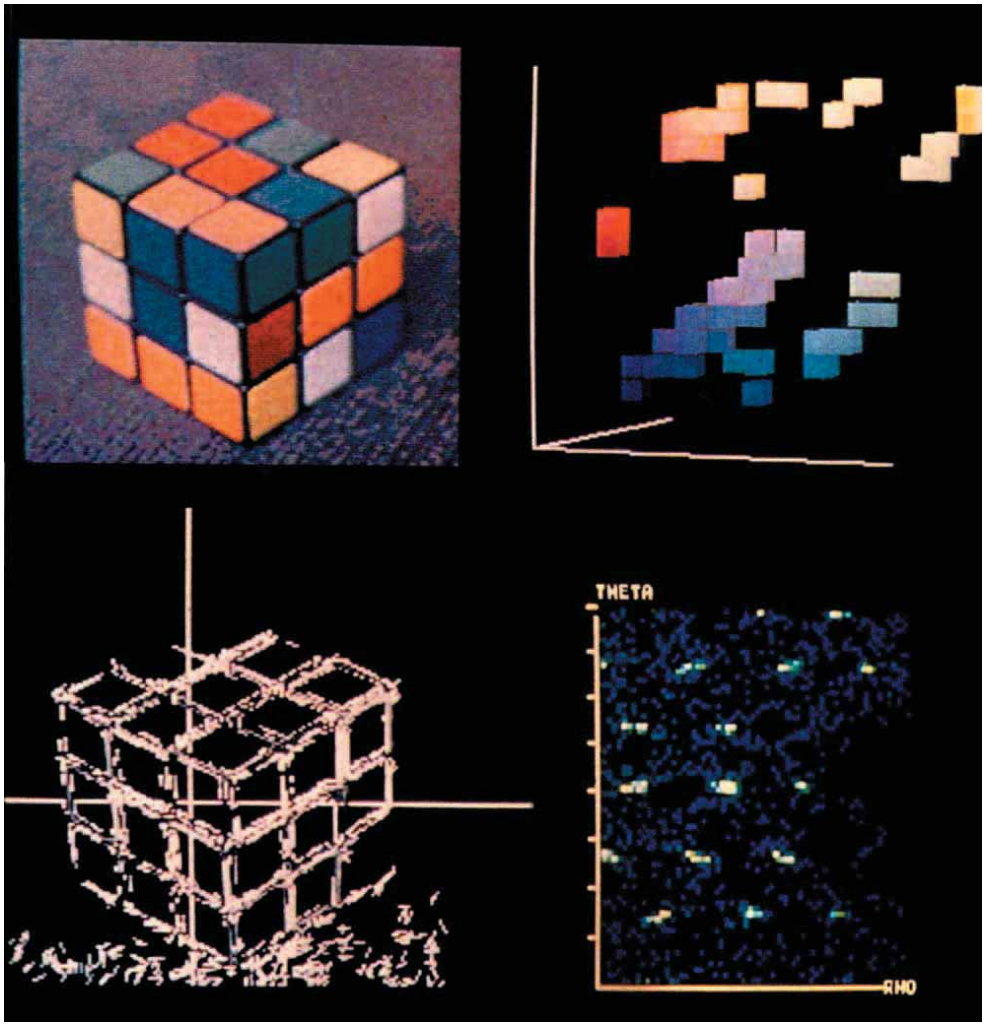

Fig. 3 The image of a Rubik cube (top left) mapped into two global feature spaces: colour space (top right) and line space (bottom right). In the colour space, a transformation was performed from each point in the colour image to a three-dimensional space of units whose axes represent three colour components. The colour of each unit is the colour it represents, and only the units with the highest activity levels are shown. Clusters of units are activated by all the faces of the Rubik cube with roughly the same colour value. Note, however, that the textured background activates many units in colour space. In the line space, local edges were first detected in the image by finding the position where the intensity was changing the fastest (bottom left). Note that many of the edges are incomplete and other edges are obscured by noise in the background. Long lines in the image can be detected by allowing colinear edges to support each other. This may be done by using another two-dimensional space (line space) in which each point represents an infinite line parameterized by the line's perpendicular distance from the origin (rho) and the angle with the x-axis made by the perpendicular to the line from the origin (theta). Each local edge unit provides input to all the global line units in line space with which it is consistent. Edge units may be inaccurate and incomplete, but lines in the image appear in line space as the units with the highest activation (white clusters)[58,67].

than a global one. This difference in the character of the search space is partly due to the fact that stereo fusion involves discrete choices between alternative correspondences for each dot, whereas the shape-from-shading computation involves gradual changes of continuous parameters, though in general there may be local optima even with continuous parameters.

The relative simplicity of the search space for the shape-from-shading algorithm disappears if it is generalized to cases where the locations of object boundaries are not known in advance and can only be decided by using assumptions about the smoothness of surfaces. In the more general problem, choices of values for continuous variables like the local surface orientation must interact with choices of values for discrete variables like the presence or absence of an edge. Search spaces that contain many local optima may well be unavoidable when designing algorithms that make both kinds of choices.

A search procedure based on statistical mechanics has recently been introduced that can escape from the local optima at which simple hill-climbing algorithms would be trapped[29,30]. If noise is added to the decision rule, so that each unit sometimes adopts a state which is less consistent with the current states of the other units, the system relaxes to different global configurations with different probabilities. After sufficient time the probability of finding it in a global configuration depends only on how consistent that configuration is, with more consistent configurations having higher probabilities. This kind of search is effective in situations in which the correct interpretation is much more consistent than any other, and this may well be typical of normal vision[31].

Representations

The performance of an algorithm depends on what information in the image is explicitly represented by the units and how that information is coded. In the Marr–Poggio algorithm the depth of a surface patch is represented by activating one of a set of units, each of which is 'tuned' to a different depth. This type of representation is qualitatively different from an analog representation in which the firing rate of a single unit directly encodes the value of a physical variable.

By dividing the range of a continuous variable into discrete intervals, each unit becomes, in effect, a hypothesis about the existence of a feature in the image, and the firing rate can then be used to encode the probability of the hypothesis being correct rather than the value of the physical parameter[32,33]. This type of 'interval encoding' makes it easier to implement searches in spaces that contain both continuous and discrete variables, because it treats continuous variables in the same way as discrete ones.

Several variables in the visual cortex are represented by broadly-tuned neural mosaics that seem to be a form of multi-dimensional interval encoding (Fig. 4). For example, disparity-sensitive neurones in primate visual cortex fall into three groups: those that are tuned to the plane of focus and those that respond to visual stimuli that are either convergent or divergent[34]. The same neurones also respond preferentially to oriented edges and line segments; these are broadly tuned over 20–40° and their range of sensitivities overlap considerably[35,36]. The broadness of the tuning need not imply imprecision, since we have, for example, only three types of cone photoreceptors with broadly-tuned spectral sensitivities, but we are nonetheless capable of fine colour discrimination.

An important practical issue that arises when implementing a relaxation algorithm is the spatial resolution of the units that represent intrinsic properties. With high resolution, fine irregularities may introduce noise, but with low resolution important discontinuities may be smoothed out. Experience has shown that having a range of resolutions is most effective[37–40], and experiments indicate that a range of resolutions is also used in the nervous system[35,41,42]. For example, in 1979 Marr and Poggio proposed an improved model of stereopsis that incorporated multiple spatial resolutions[43,44]. Matching between the right and left images was first performed on the largest channel, which has the least ambiguity, and this helped matches at progressively finer scales. The issue of exactly what primitive features should be used to perform the matching is not yet resolved[45–47].

Visual maps and feature spaces

Our visual system is capable of computing surface orientation from shading and depth from binocular images, but we do not yet know how these computations are organized. We do know that an important organizing principle in the early stages of visual processing is the orderly mapping and remapping of the visual field. Twelve maps of the visual field have been found in the primate cerebral cortex so far[48,49], and there are also many subcortical maps.

In retinotopic maps, neighbourhood relationships between nearby positions of the visual field are retained, while through convergence and local interactions the response properties of cells are modified. Already at the level of a retinal ganglion cell, light falling on surrounding positions inhibits the central response, a combination that is well known to enhance contrast information[50]. In cortical maps higher-order combinations are used to compute orientation and directional selectivity[36]. Extraction of these features has proven to be very effective in preprocessing images for computer vision.

Most of the neurones in extrastriate visual maps receive their major visual input either directly or indirectly from the striate cortex, the first area of cerebral cortex to receive visual input, and they have larger receptive fields and less topographic order than neurones in striate. The proportion of cells sensitive to a particular property is different in the various maps. For example, the cells in an area called MT appear to be particularly sensitive to the direction of stimulus motion[51-53] and their responses are selectively influenced by motion of the background[54] (Fig. 5). In the inferior temporal cortex the receptive fields of some neurones cover most of the visual field and their preferred stimulus is often not an oriented edge[55,69]. The functions of most of these visual maps remain obscure[56].

Parts of an object can often be identified as a whole by common properties, such as colinearity of edges, colour and direction of movement, even though the object is partially obscured by other objects. Barlow[57] has pointed out that an economical way to link up parts of an image having a common feature is to collect that information in a single location irrespective of where it came from in the image. Neighbouring locations in this new space represent neighbouring values along a non-spatial dimension rather than neighbouring positions in the image; two examples of such non-retinotopic feature spaces are shown in Fig. 3.

The first use of a transformation from an image to a non-retinotopic feature space was made by Hough[58], who patented a machine based on this principle in 1962. Hough found that noisy particle tracks in bubble chamber photographs could be efficiently analysed by transforming the data into a line space in which each point represented an infinite line with a particular slope and perpendicular distance from the origin, as explained in Fig. 3. Each detected edge 'votes' for all the lines in line-space that are consistent with it. Even though each piece of evidence is noisy, a line that is present in an image will receive many votes.

Hough transforms have been generalized in several ways. First, the feature space need not be simply related to the local features computed in intrinsic images[59]; second, the units in the feature space may themselves interact and even feed back to the intrinsic images[60]. For example, in the computation of surface orientation from shading, the direction of the light source had to be specified as a boundary condition. The two angles that specify the direction of illumination are global variables and can be computed from the intensity array and from a very approximate intrinsic image of the surface orientation. The new direction of illumination can then be used to refine the intrinsic image. The coupled algorithms converge to an intrinsic image of the object's surface and the direction of the light source[60], as shown in Fig. 2.

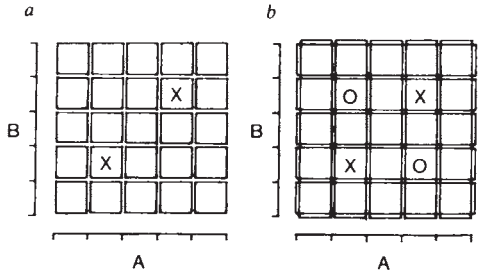

Fig. 4 Schematic drawings of two ways that a two-dimensional space of features (such as orientation angle and disparity) can be represented by populations of units. *a*, Each unit is assigned to a tile of the mosaic and jointly represents the intervals given by its position in the mosaic. *b*, Each unit is assigned to an interval of one variable and is insensitive to the other. If the features of two different objects are simultaneously present, then in representation *a* two units are activated, but in representation *b* four units are activated, as indicated by the Xs. Note that in the latter case there is an ambiguity in which pair of features go together since the same units would also be activated by 'ghost' objects whose features are indicated by the Os.

Although extrastriate maps have an overall hierarchical organization, forward projections are generally accompanied by reciprocal feedback connections. The function of these feedback connections is unknown, but one possibility is to enhance or 'focus attention' on the parts of the image having a common feature[57]. A more general function of feedback, suggested by Fig. 2, is to link two cortical areas for the purpose of joint interdependent computation.

Neural representations

The map of surface orientation is an important intermediate step in building a three-dimensional interpretation of an image. Evidence has not yet been found for neurones in the visual cortex sensitive to surface orientation, but the appropriate stimuli have probably not yet been used to look for them. This is not an easy task given the variety of cues that could be used to generate images of surfaces and the large number of visual areas. Although computer vision may offer suggestive hints, it should be emphasized that algorithms are not unique solutions to problems; the example of stereopsis, for which we have several working models, illustrates that further evidence from psychophysics as well as physiology will be necessary to constrain possible explanations[22,47].

Enough neurones exist in the visual cortex to represent explicitly many features of natural images to which the human visual system is sensitive[31]. Beneath every square millimetre of cortex there are $\sim 10^5$ neurones; in the striate cortex this is about the size of a hypercolumn that contains all the machinery needed to analyse a single patch the size of one receptive field[42]. Most neurones in the visual cortex respond best only when two or more features are simultaneously present. Even though the number of neurones in the visual cortex seems comfortably large, there is a finite limit to the number of features that can be explicitly represented, and it is particularly costly to represent all possible conjunctions of properties. For example, if enough neurones were provided to represent all possible combinations of values of edge orientation, disparity, velocity, spatial frequency, colour, etc., the number required would increase exponentially with the number of dimensions. With a resolution of 10 intervals on each dimension, the 10^5 neurones in a hypercolumn could only cover a five-dimensional feature space[68].

One solution to the exponential explosion is to give up on making all conjunctions explicit and to represent different features in different maps[59]. The price of this compromise is a possible confusion in the correspondence between pairs of features in different maps that are simultaneously activated by several objects (Fig. 4). In fact, when coloured letters are briefly presented to human observers, incorrect matches between the colours and the letters are quite often reported in certain conditions, a phenomenon called illusory conjunctions[61].

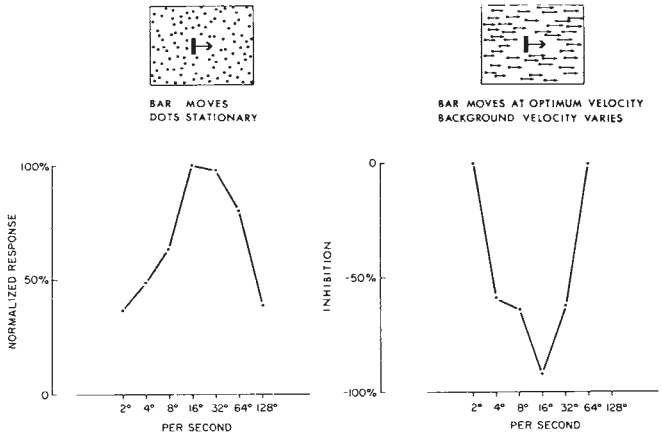

Fig. 5 Selective influence of a moving background of random dots on the response of a directionally-selective neurone in area MT of the owl monkey. The maximum firing rate occurs when a bar is moving in its preferred direction at its optimum velocity[54]. Left, firing rate (normalized to the maximum rate) to a bar moving against a stationary background as a function of velocity. Right, response of the bar moving at its optimum velocity as a function of the velocity of the background. The response is inhibited if the background moves in the same direction as the bar with the maximum inhibition occurring when the background moves at the same velocity as the bar. The area of the visual field influencing the response to the bar was about 20,000 square degrees, 60 times the area of the classical receptive field. In V4, another area of extrastriate cortex, the responses of some cells to coloured stimuli are strongly influenced by the colour of the surrounding illumination[70,71], and these long-range effects may be mediated by a recently identified colour pathway originating in striate cortex[72]. The discovery that the surround is as selectively tuned as the classical receptive field may lead to a better understanding of the algorithms used in the visual system to compute intrinsic images such as that of optical flow and reflectance. (Courtesy J. M. Allman.)

Another way to cover a feature space of high dimension with a limited number of units is to assign them to large, overlapping regions. A single point in feature space is then represented not by activity in a single unit, but by the joint activity of many units. Hinton[62] has recently shown that accurate information is not lost by these coarsely tuned units; the information has simply been distributed over a population of units. Moreover, far fewer units are required to achieve the same accuracy. For binary units that respond only when a feature falls within their region of feature space, the number of units required to achieve a given accuracy is proportional to $1/r^{(k-1)}$, where r is the radius of the region and k is the dimensionality of the space. Thus, the relative efficiency of coarse coding increases with the dimensionality of the space being covered.

Although coarse coding efficiently encodes single features, two nearby points in feature space that are simultaneously excited may be confused. The sparseness of the features that occur in any one image therefore sets an upper bound on the allowable coarseness of the coding.

Future directions

It is generally thought that at the later stages of visual processing the features of an object are represented relative to a frame of reference based on the object rather than in a frame relative to the retina[10]. By imposing an appropriate frame of reference on an object, and recoding features relative to this frame, the visual system can generate 'object-based' features that are independent of viewpoint and therefore represent the invariant shape of the object. Algorithms that use generalized Hough transforms in coarsely coded feature spaces[63,64] have been proposed for choosing the object-based frame and generating the object-based features. Thus, the same type of parallel relaxation algorithm that was used in early visual processing might also be useful for higher levels. This would be consistent with the similar cytoarchitectonic organization found throughout the cerebral cortex, from primary sensory to higher associational areas[65,66].

Despite the power of parallel computation, some problems in vision clearly require sequentiality. Consider, for example, the role of eye movements in building up a detailed internal representation of a complex object, such as a face. The integration of sequentially acquired information raises issues about the nature of visual working memory and the central control of processing. Parallel models of these aspects of vision are being explored but are beyond the scope of this review.

We have described a recent trend within computer vision towards parallel models and the relationship of these models to processing in visual cortex. As the parallel architectures in computer vision come to resemble biological visual systems more and more closely and as we gain more experience with these architectures, we may achieve new insights that will better help us to appreciate nature's designs.

We thank Francis Crick, Jerome Feldman, and Helen Sherk for helpful comments on the manuscript. This work was supported by grants from the System Development Foundation.

1. Roberts, L. G. in *Optical and Electro-Optical Information Processing* (eds Tippett, J. P. *et al.*) 159–197 (MIT Press, Cambridge, 1965).
2. Rosenfeld, A. *Picture Processing by Computer* (Academic, New York, 1969).
3. Barlow, H. B., Narasimhan, R. & Rosenfeld, A. *Science* **177,** 567–575 (1972).
4. Arbib, M. A. *Annl Biomed. Engng* **3,** 238–274 (1975).
5. Marr, D. in *Computer Vision Systems* (eds Hanson, A. R. & Riseman, E. M.) 61–80 (Academic, New York, 1978).
6. Barrow, H. G. & Tenenbaum, J. M. in *Computer Vision Systems* (eds Hanson, A. R. & Riseman, E. M.) 3–26 (Academic, New York, 1978).
7. Feldman, J. in *Parallel Models of Associative Memory* (eds Hinton, G. E. & Anderson, J. A.) 49–104 (Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1981).
8. Barrow, H. & Tenenbaum, J. *Proc. IEEE* **69,** 572–595 (1981).
9. Ballard, D. H. & Brown, C. M. *Computer Vision* (Prentice-Hall, Englewood Cliffs, New Jersey, 1982).
10. Marr, D. *Vision* (Freeman, San Francisco, 1982).
11. Brady, M. *Computing Surveys* **14,** 3–71 (1982).
12. Horn, B. K. P. in *Psychology of Computer Vision* (ed. Winston, P. H.) 115–155 (McGraw-Hill, New York, 1975).
13. Fahlman, S. E. *NETL: A System for Representing and Using Real-World Knowledge* (MIT Press, Cambridge, 1979).
14. Fahlman, S. E., Hinton, G. E. & Sejnowski, T. J. *Proc. AAAI Conf. Artificial Intelligence,* Washington DC 109–113 (Kauffman, Los Altos, 1983).
15. Hillis, W. D. *MIT A. I. Laboratory Memo* No. 646 (1981).
16. Hinton, G. E. thesis, Univ. Edinburgh (1977).
17. Ullman, S. *Computer Graphics Image Processing* **10,** 115–125 (1979).
18. Riseman, E. M. & Arbib, M. A. *Computer Graphics Image Processing* **6,** 221–276 (1977).
19. Rosenfeld, A., Hummel, R. & Zucker, S. W. *IEEE Trans. Systems, Man Cybernet.* **6,** 420 (1978).
20. Ballard, D. H. *Proc. 7th int. Joint Conf. Artificial Intelligence,* 1068–1078 (Kauffman, Los Altos, 1981).
21. Anderson, J. & Hinton, G. E. in *Parallel Models of Associative Memory* (eds Hinton, G. E. & Anderson, J. A.) 9–48 (Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1981).
22. Attneave, F. in *Organization and Representation in Perception* (ed. Beck, J.) 11–24 (Lawrence Erlbaum Associates, Hillsdale, New Jersey 1982).
23. Waltz in *Phychology of Computer Vision* (ed. Winston, P. H.) 19–91 (McGraw-Hill, New York, 1975).
24. Horn, B. K. P. *Artificial Intelligence* **8,** 201–231 (1977).
25. Ikeuchi, K. & Horn, B. K. P. *Artificial Intelligence* **17,** 141–184 (1981).
26. Horn, B. K. P. & Schunk, B. G. *Artificial Intelligence* **17,** 185–203 (1981).
27. Julesz, B. *Foundations of Cyclopean Vision* (University of Chicago Press, 1971).
28. Marr, D. & Poggio, T. *Science* **194,** 283–287 (1976).
29. Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. D. *Science* **220,** 671–680 (1983).
30. Hinton, G. E. & Sejnowski, T. J. *Proc. IEEE Conf. Computer Vision and Pattern Recognition,* 448–453 (IEEE Computer Science Press, Silver Spring, Maryland, 1983).
31. Gibson, J. J. *The Senses Considered as Perceptual Systems* (Houghton Mifflin, Boston, 1966).
32. Barlow, H. B. *Perception* **1,** 371–394 (1972).
33. Feldman, J. & Ballard, D. H. *Cognitive Sci.* **9,** 205–254 (1983).
34. Poggio, G. F. & Fischer, B. J. *J. Neurophysiol.* **40,** 1392–1405 (1978).
35. Hubel, D. H. & Wiesel, T. N. *J. Physiol., Lond.* **160,** 106–154 (1962).
36. Hubel, D. H. & Wiesel, T. N. *Proc. R. Soc.* **B198,** 1–59 (1977).
37. Rosenfeld, A. & Vanderbrug, G. J. *IEEE Trans. System, Man Cybernet.* **7,** 104–107 (1977).
38. Hanson, A. R. & Riesman, E. M. in *Computer Vision Systems* (eds Hanson, A. R. & Riseman, E. M.) 129–163 (Academic, New York, 1978).
39. O'Rourke, J. in *Proc. 7th int. Joint, Conf. Artificial Intelligence* Vol. 2, 737–739 (Kauffman, Los Altos, 1981).
40. Sloan, K. in *Proc. 7th int. Joint Conf. Artificial Intelligence* Vol. 2, 734–736 (Kauffman, Los Altos, 1981).
41. Julesz, B. & Schumer, R. A. *A. Rev. Psychol.* **32,** 575–627 (1981).
42. Hubel, D. H. & Wiesel, T. N. *J. comp. Neurol.* **158,** 295–305 (1974).
43. Marr, D. & Poggio, T. *Proc. R. Soc.* **B204,** 301–328 (1979).
44. Grimson, W. E. L. *From Image to Surface* (MIT Press, Cambridge, 1981).
45. Mayhew, J. E. W. & Frisby, J. P. *Artificial Intelligence* **17,** 349–385 (1981).
46. Baker, H. H. & Binford, T. O. in *Proc. 7th int. Joint Conf. Artificial Intelligence* Vol 2, 631–636 (Kauffman, Los Altos, 1981).
47. Mayhew, J. in *Physical and Biological Processing of Images* (eds Braddick, O. J. & Sleigh, A. C.) 204–216 (Springer, New York, 1983).
48. Allman, J. M., Baker, J. F., Newsome, W. T. & Petersen, S. E. in *Cortical Sensory Organization* Vol. 2 (ed. Woolsey, C. N.) 171–185 (Humana, Clifton, New Jersey, 1981).
49. Van Essen, D. C. *A. Rev. Neurosci.* **2,** 227–263 (1979).
50. Hartline, H. K. & Ratliff, F. *J. gen. Physiol.* **40,** 1357–1376 (1957).
51. Allman, J. M. & Kaas, J. H. *Science* **191,** 572–575 (1976).
52. Zeki, S. *Proc. R. Soc.* **B207,** 239–248 (1980).
53. Maunsell, J. H. R. & Van Essen, D. C. *J. Neurophysiol.* **49,** 1127–1167 (1983).
54. Allman, J. *Perception* (in the press).
55. Gross, C. G., Rocha-Miranda, C. E. & Bender, D. B. *J. Neurophysiol.* **35,** 96–111 (1972).
56. Cowey, A. in *Neurosciences: 4th Study Program* (eds Schmitt, F. O. & Worden, F. G.) 395–413 (MIT Press, Cambridge, 1979).
57. Barlow, H. B. *Proc. R. Soc.* **B212,** 1–34 (1981).
58. Hough, P. V. C. *US Patent 3,069,654* (1962).
59. Ballard, D. H. & Kimball, O. A. *University of Rochester Department of Computer Science Tech. Rep.* TR-70 (1981).
60. Brown, C. M., Ballard, D. H. & Kimball, O. A. *Proc. DARPA Image Understanding Workshop* (National Technical Information Service, Springfield, 1982).
61. Treisman, A. M. & Schmidt, H. *Cognitive Psychol.* **14,** 107–141 (1982).
62. Hinton, G. E. *Proc. 7th int. Joint Conf. Artificial Intelligence* Vol 2, 1088–1096 (Kauffman, Los Altos, 1981).
63. Hinton, G. E. in *Proc. 7th int. Joint Conf. Artificial Intelligence* Vol. 2, 683–685 (Kauffman, Los Altos, 1981).
64. Hrechanyk, L. H. & Ballard, D. H. *Proc. IEEE Workshop on Computer Vision,* 44–82 (IEEE Computer Science Press, Silver Spring, Maryland, 1982).
65. Mountcastle, V. in *Neurosciences: Study Program* (eds Schmitt, F. O. & Worden, F. G.) 21–42 (MIT Press, Cambridge, 1979).
66. Jones, E. G. in *The Organization of Cerebral Cortex* (eds Schmitt, F. O. *et al.*) (199–235 (MIT Press, Cambridge, 1981).
67. Duda, R. O. & Hart, P. E. *Commun. Ass. Computer Machinery* **15,** 11–15 (1972).
68. Ballard, D. H., & Coleman, P. C. in *Proc. Workshop on Vision, Brain & Cooperative Computation* (University of Massachusetts, Amherst, 1983).
69. Richmond, B. J., Wurtz, R. H. & Sato, T. *J. Neurophysiol.* (in the press).
70. Zeki, S. *Nature* **284,** 412–418 (1980).
71. Schein, S. J., Desimone, R. & de Monasterio, F. M. *Investig. Opthal. Visual Sci.* **24,** S107 (1983).
72. Livingstone, M. S. & Hubel, D. H. *Nature* **304,** 531–534 (1983).

ARTICLES

Turbulent dissipation and shear in thermohaline intrusions

Nordeen G. Larson & Michael C. Gregg

Applied Physics Laboratory and School of Oceanography, University of Washington, Seattle, Washington 98105, USA

Laterally coherent patches of turbulence were discovered on the upper and lower boundaries of thermohaline intrusions. The ratio ε/J_b is used to compare the relative importance of turbulence produced by Reynolds stress and that produced by the buoyancy flux of double diffusion. Velocity profiles show near-inertial motions through the intrusions.

FRONTAL regions in the ocean are characterized by the frequent occurrence of 5–30 m thick interleaving lenses of water, which are referred to as intrusions. The horizontal and vertical transports of heat and salt by intrusions may be major factors in the flux balances of fronts. Therefore, determination of the mechanisms by which the intrusions are formed and the rates of subsequent mixing is essential to understand frontal dynamics.

During the past year, we have used the Advanced Microstructure Profiler (AMP)[1] to take many sets of closely spaced profiles through intrusions in the Bahamas, in a warm-core Gulf Stream ring, and in the California Current. Velocity microstructure measurements were used to determine ε, the rate of viscous dissipation of turbulent kinetic energy. Observations of temperature–salinity (Ts) fine structure were used to estimate J_b, the buoyancy flux across interfaces believed to be sites of double

0028-0836/83/440026—07$01.00