South Florida and R. Parkman, Children's Hospital of Los Angeles) and are 'mast cell-like' in that they contain granules that can be shown to cytotoxic.

The nature of the cytotoxicity has yet to be elucidated. M. Slavin described the suppressor cells induced after TLI in rodents as thymus-independent, non-lymphoid, large and immature, and suggested that the effect of TLI is to cause a reversion of the animal's lymphoid system to the neonatal state. Examining the thymus after TLI, Strober found both histological and functional abnormalities in the medulla (failure of cell proliferation after exposure to interleukin-1 and a marked drop in interleukin-2 production when cultured with the lectin concanavalin A), which he explained by the premature departure of cells from the thymus.

There seemed to be some consensus that the suppressor cells in many experimental models are large granular lymphocyte-like cells and K. Singhal (University of Western Ontario) showed that such cells, isolated from normal bone marrow, can regulate *in vitro* immunoglobulin M responses to several non-transplantation antigens. These cells release a heat-stable, low molecular mass glycolipid that is suppressive *in vitro* if antigen is present from an early stage. The view was expressed that the role of natural suppressor cells is primarily to assist with the functional deletion of self-reactive clones, but they could also have great significance in protecting the mammalian fetus against destruction by maternal responses. The study by B. Roser (Institute of Animal Physiology, Babraham, UK) on neonatally induced tolerance was, however, a salutory reminder that suppressor T lymphocytes, in this case anti-idiotypic cells acting together with another, as yet unidentified, cell can be critically important in transplantation tolerance.

In summary, the conference left no doubt that approaches towards human tolerance are varied; that the main objective remains both valid and desirable; and that we still have a long way to go. ☐

Leslie Brent is in the Department of Immunology, St Mary's Hospital Medical School, London W2 1PG, UK.

Visual information

Do computers need attention?

from Anya Hurlbert and Tomaso Poggio

DESPITE enormous progress in the past few decades, machine vision is still far from achieving the goal that human vision attains with such speed and reliability — in David Marr's words[1], to "know what is where by looking". Recent results on the physiology and psychophysics of visual attention, one example of which is reported on page 693 of this issue[2], accentuate the gap between machines and humans, and provide a first step to understanding why it is so large and what machines must learn to overcome it.

Paradoxically, what appear to be the simplest tasks for humans may be the most difficult for machines. Consider, for example, recognizing your mother in a sketch of her sitting in the kitchen. You could immediately and effortlessly locate her face, match it with your memory and pronounce it a good or a bad likeness. If the sketch were upside-down you could easily right it for a proper view. You would probably expend the most painstaking scrutiny in determining just which feature was slightly off, but even so, your final judgement would be quick. In contrast, a computer, using the most sophisticated face-recognition routine, would perform the task slowly and incompletely, because it would not know where to start. Given the location of the two eyes in a sketch cluttered with dark, round blobs, a computer could be programmed to search for the mouth, nose and chin at the appropriate distances and methodically to match each feature to a virtually identical image in its memory. But failing to find the eyes, the computer could not go on to recognize the face.

Object recognition

The difficulty of the face-recognition problem — and, more generally, object recognition — has called into question one of the main assumptions in the construction of a machine that sees as humans do. The assumption holds that the goal of the first stages in vision is solely to determine 'where' things are, that is, to transform the initial image, an array of intensity values, into a map of the scene which records the distance and orientation of each surface point relative to the viewer (the $2\frac{1}{2}-D$ sketch[1]). In machine vision the $2\frac{1}{2}-D$ sketch may serve to guide a mobile robot around an obstacle or to control its manipulations as it picks up a tool. But, like the raw image from which it is computed, the $2\frac{1}{2}-D$ sketch is itself simply a large array of numbers. Although it may contain preliminary information for object recognition by assigning a colour or texture to each surface point, it does not tell 'what' things are. The critical task in object recognition is therefore to find the object or its crucial part within an array of intensity values or distances. Until now, many attempts to elucidate object recognition (reviewed in refs 3,4) have assumed that the relevant object is already located and isolated in the image.

Unlike machines, humans are adept in spotting the salient features of an object. To understand the mechanisms underlying this ability, psychophysicists have investigated visual attention. Treisman[5] and Julesz[6] have demonstrated that humans are extremely efficient in detecting a part of an image that differs in a single aspect from its background. For example, a red dot pops out in a field of yellow dots, and the same happens for a horizontal line in a field of vertical lines (Fig.1). The time required to detect the unusual item is independent of the number of other items, implying that the search for it occurs in parallel across the entire field. The human visual system obviously possesses a fast, parallel mechanism which can direct attention to salient chunks of the image. (This is sometimes called 'preattentive'. Here, we consider it as part of the entire attention mechanism whose characteristics probably require more complex descriptions than 'serial' or 'parallel'.) Although the possible computational purposes of this mechanism have not been probed by psychophysical experiments, its potential role in object recognition seems critical. For example, in face recognition the attention mechanism may perform two essential steps: first, to locate 'blobs' which could be eyes; and second, to direct processing towards the blobs to verify that they are eyes and thereby to initiate recognition. The role of attention, therefore, may be not only to spotlight distinctive parts of the image but, more importantly, to segment the image into objects or parts of objects, a crucial first step in determining what things are.

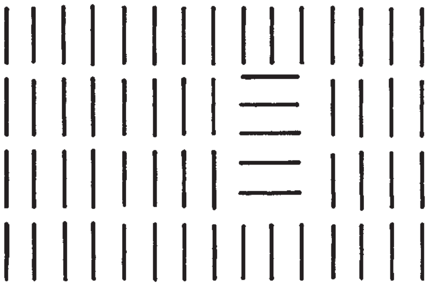

Fig 1. A patch of horizontal line segments pops out in a field of line vertical line segments.

Separable features

An important and still open question is: what are the features or primitives that drive attention? Likely candidates are separable features which, by definition, can be attended to selectively and are processed independently and in parallel. Pop-out and texture-discrimination experiments provide a test for separable features and so far have diagnosed colour, line orientation, line ends (terminators) and possibly crossings among the candidates.

Conjunction experiments test whether two or more separable features can combine to produce a higher-order primitive. For example, when a green T in a field of randomly mixed green Xs and brown Ts is the target, it does not pop out, and the time required to detect it increases linearly with the number of background items. Thus, the detection of a particular conjunction of colour and shape appears to require a search over each item in turn, across the entire field. Conjunction experiments thus reveal another aspect of the attention mechanism, a serial searchlight which appears to operate independently of eye movements and does for feature conjunctions what the parallel mechanism does for features.

Until recently, all conjunctions between known separable features had been shown to require the serial searchlight. The recent results of Nakayama and Silverman[7] reveal a surprising exception to this pattern. In pop-out experiments using fields of small rectangular patterns displayed on a colour television monitor, the authors demonstrate that binocular disparity and motion individually behave as separable features. The conjunction of motion and colour does not; the search for a pattern of blue upward-moving dots is slow and serial across a field of blue-downward patterns and red-upward patterns. Contrary to this trend, conjunctions of binocular disparity and either colour or motion behave as separable features: they are searched for in parallel. (The authors report that when the field splits into two planes, one in front of the other, the search for a conjunction amounts to a pop-out of the unusual item in one plane. Thus, we suggest that it may be possible, using a different kind of motion stimulus, to create separate planes of coherent motion and thereby induce a parallel search for motion-colour conjunctions.)

Experiments

The psychophysical studies on separable features coincide with the recent emphasis on functional localization in visual neurophysiology and neuroanatomy. It is tempting to draw an explicit connection between biology and psychophysics by equating different visual cortical areas with different feature maps; for example, calling area V4 the colour map, MT the motion map and V1 the orientation map. Psychophysics suggests that in a given feature map there exists at each spatial location a collection of neurones each tuned to a different value of the feature (for example, red, green or blue for colour). Although such an organization has not been demonstrated, evidence for segregation of functionally similar neurones in distinct cortical areas is steadily accumulating. From this point of view, the results of Nakayama and Silverman have interesting implications for neurones and feature maps: they preclude the existence of neurones tuned for both motion and colour; predict the existence of neurones tuned to a particular combination of binocular disparity and motion and of neurones tuned to disparity and colour; and suggest that feature maps are replicated at each of several disparity planes. The prediction of disparity–motion-tuned neurones is supported by the recent report[8] of similar neurones in cortical area MT.

Other recent work on visual neurophysiology puts the emphasis on a different aspect of attention. Rather than address computational questions such as 'what are the salient features?' and 'how does the attention mechanism work?' or the psychophysical question 'is feature processing parallel or serial?', the new class of physiological experiments on alert animals seeks to demonstrate the ways in which attention can modulate neuronal responses. In the course of such experiments, insights into the neural circuitry and anatomical location of the attention mechanism have emerged. For example, based on studies of attention-mediated modulation in the inferior parietal lobe (area 7), Mountcastle[9] proposed that neurones there are responsible for directing attention to visual targets.

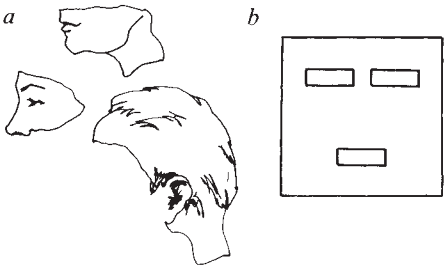

Fig. 2 *a*, each separate set of 'face' features is sufficient to suggest the hypothesis of a face. *b*, It is spatial relation between features and not the features themselves that cue recognition to the face of hypothesis.

More recent research demonstrates the effects of attention at other levels in the visual pathway. Moran and Desimone[10] have recently shown that, in the monkey, the response of a neurone in V4 or IT to a preferred stimulus (for example, a red horizontal bar) is dramatically reduced when the animal ignores it and instead attends to an ineffective stimulus (such as a green vertical bar) within the same receptive field (which, for IT neurones, may extend at least 12°). The response of the neurone to the preferred stimulus is unaffected when the attended stimulus is outside its receptive field. Thus, V4 and IT neurones are able to filter out an irrelevant stimulus when it competes with a relevant stimulus within the same receptive field. V1 neurones do not have this property, since the monkey cannot even perform the differential attention task when two stimuli are close enough to fit within a single receptive field in V1.

The psychophysical experiment in humans reported in this issue by Sagi and Julesz[2] provides an intriguing complement to the physiological results described above. The authors find that visual attention directed to a random location for an orientation-discrimination task enhances the detection of a test flash presented simultaneously within a certain radius of the target. The area of enhancement, which the authors conjecture to be the area covered by the searchlight of attention, varies from 1.5° at 2° eccentricity to about 3° at 4° eccentricity. Interestingly, these areas are likely to be larger than the average receptive-field sizes in V1.

The above results imply that attention to one region of an image involves both suppression of visual processing in irrelevant domains and enhancement of visual processing in relevant domains. Thus, attention may indeed be responsible for directing a processing focus to specific locations in the initial steps of recognition. Yet although biological research may have found the key to machine vision, it has yet to describe how it opens the lock.

Attention mechanism

Computational results suggest that the attention mechanism is even more complex and powerful than experiments have revealed. Consider again the face-recognition problem. Individual features such as eyes or the curved line of nose and mouth can by themselves lead to the hypothesis of a face (Fig.2*a*). In contrast (Fig.2*b*), features alone cannot be the only cue for recognition. The spatial relationship between the two eye tokens and the closed outer contour can also cue the face-recognition process. Ullman[11] argues that spatial relations must be computed by a mechanism similar to the serial searchlight of attention.

The unravelling of the full complexity of visual attention will clearly involve computational, psychophysical and physiological research. It will influence not only our understanding of visual perception but also the architecture and the control structure of machine vision systems. □

1. Marr, D. *Phil. Trans. R. Soc.* **B275**, 483 (1976).
2. Sagi, D. & Julesz, B. *Nature* **321**, 693 (1986).
3. Besl, P.J. & Jain, R. *Computing Surveys* **17**, 74 (1985).
4. Harmon, L.D., Kahn, M.K., Lasch, R. & Ramig, P.F. *Pattern Recognition* **13**, 97 (1979).
5. Treisman, A. in *Physical and Biological Processing of Images* (eds Braddick, O.J. & Sleigh, A.C.) 316 (Springer, Berlin, 1983).
6. Julesz, B. *Trends Neurosci.* **7**, 41 (1984).
7. Nakayama, K. & Silverman, G.H. *Nature* **320**, 264 (1986).
8. Van Essen, D.C. & Maunsell, J. *J. Neurophysiol.* **49**, 1148 (1983).
9. Lynch, J.C., Mountcastle, V.B., Talbot, W.H. & Yin, T.C.T. *J. Neurophysiol.* **40**, 362 (1977).
10. Moran, J. & Desimone, R. *Science* **229**, 782 (1985).
11. Ullman, S. *Cognition* **18**, 97 (1984).

Anya Hurlbert and Tomaso Poggio are in the Center for Biological Information Processing, Whitaker College, and at the Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.