

Elemental operations in vision

Pieter R. Roelfsema

The Netherlands Ophthalmic Research Institute, Meibergdreef 47, 1105 BA Amsterdam, the Netherlands; and Department of Experimental Neurophysiology, Centre for Neurogenomics and Cognitive Research, Vrije Universiteit Amsterdam, The Netherlands

Many cognitive tasks are serial in that they require several steps that have to be carried out in a sequence. Here, I outline a new theory of how these processing steps are implemented in vision by networks of neurons that span several areas of the cerebral cortex. The theory explains how individual neurons can contribute to the elementary processing steps, and also how several processing steps can be arranged in a sequence to form more complex visual routines, just as computer programs can be composed of sequences of instructions. Evidence in support of the theory comes from recent neurophysiological findings in monkeys engaged in complex visual tasks. It is likely that future work will allow these ideas to be generalized to other sensory modalities and cognitive functions.

Introduction

When we think about a problem we often spend several seconds introspectively going through a number of processing steps. In the past few years neuroscientists have started to investigate what happens in the cerebral cortex during such a 'step of thought'. This has revealed several important insights, including that networks of neurons spanning several areas of cortex jointly up- or downregulate their activity as they converge on a solution for that processing step. A number of recent studies have also started to investigate how multiple steps are executed in series to solve complex tasks. Thus, neuroscience is starting to shed light on the implementation of serial cognitive tasks, an issue that used to belong primarily to the domains of cognitive science, psychology and artificial intelligence research [1,2]. In this article, I will outline a theory in vision that (i) explains how networks of neurons cooperate in elementary processing steps, and (ii) provides insight in how these processing steps are joined to form longer visual routines. In my view, the description of cognitive processes as sequences of elementary processing steps provides a common language that is useful for neuroscientists, psychologists, and workers in artificial intelligence.

There is a wealth of information about how individual neurons in the various areas of the cerebral cortex and subcortical nuclei are tuned to sensory stimuli, motor responses, associations between stimuli and responses, and how this tuning depends on the input and output connections of these structures. Too often, this information seems disconnected from cognition. At first sight,

it seems difficult to understand how a neuron with its fixed pattern of connections could contribute to something as flexible as cognition. Alan Newell was aware of the large gap between the single neuron and cognition [2], and he argued that a meaningful connection can only be made by describing the nervous system at levels above the single neuron. He proposed new system levels that are intermediate between the individual neuron and cognition (Box 1). In my opinion, sufficient scientific progress has now been made to make a start on the neurophysiological definition of these new system levels.

The proposed levels describe the activity of networks of neurons. Fortunately, the exploration of the interactions between networks of neurons situated in different brain areas is beginning to be a fruitful area of research [3,4]. To define the new levels, I will turn to the work of Ullman in vision [5]. In Ullman's theory, the first level above the single neuron is that of 'elemental operators', which can be arranged into sequences to form 'visual routines' at the next higher level. This theory did not have an implementation when it was proposed, but here I will describe how elemental operations and visual routines can indeed be reliably observed as the evolution of neural activity across time in the visual cortex [6–8]. In the following, I will focus on visual tasks but it is likely that future work will allow the generalization of the present ideas to other sensory

Box 1. System levels

Alan Newell suggested that the flexibility required for cognition does not occur at the level of individual neurons, but that it emerges only at higher levels of description [2]. A computer analogy illustrates the argument. If we ask how an individual neuron contributes to a cognitive routine, this question is akin to asking how a single connection of a computer processor contributes to the execution of a high-level computer program. We can define several system levels intermediate between the wiring of the processor and the program. The first level above the hardware level is that of the processor's machine instructions. A next higher level is the syntax of the programming language. The commands provided by the programming language finally give rise to functions and routines, which in turn give rise to programs. At these higher levels new concepts emerge, such as variables and control loops. An important feature gained at the level of the programming language is what Newell called 'distal access'. Distal access means that it is possible for a process to retrieve information that is not locally available. In modern programming languages, distal access comes for free, as it is always possible to retrieve the value of relevant variables. However, this information is not always available at the processor level, as variables are only loaded if required, but the shifts of information taking place between processor and memory to store and retrieve variables are not visible at the level of the program. This implies that a single step of the program typically corresponds many more steps at the level of the processor.

Corresponding author: Roelfsema, P.R. (p.roelfsema@ioi.knaw.nl).

Available online 5 April 2005

modalities and higher cognitive functions, such as navigation, language and reasoning in specific domains [1].

The base representation

The first cortical process that occurs when we present a visual image to a subject is that many visual features in the image are extracted in parallel by specialized areas of the visual cortex. Early visual areas extract contours; higher visual areas extract more complex features, like form-features, patterns of motion, and so on. This initial phase is mainly determined by the selectivity of the feedforward connections that propagate activity from lower to higher cortical areas [9]. The ‘feedforward sweep’ proceeds very rapidly so that most of the basic features are available within 100 or 120 ms after stimulus appearance. It is convenient to call this initial pattern of neuronal activity distributed across the many visual areas the ‘base representation’ [5,10,11].

The base representation is a good start, but it does not solve all problems in vision. Some of these unsolved problems are caused by the distributed nature of the neuronal representations, implying that features of a single object, like its location and shape, are represented in different visual areas [12]. This poses ‘binding problems’ when there is more than one visual object in view [13,14]. If the image contains e.g. a dog on the left and a cat on the right, then the base representation encodes the features LEFT, RIGHT, DOG and CAT, but it may not distinguish the image from another one that contains a cat on the left and a dog on the right. The binding problem can be solved by the inclusion of conjunction detectors in the base representation, such as DOG_LEFT neurons that respond only to dogs on the left. These hardwired conjunctions have been called ‘base-groupings’ [10]. Although useful, base-grouping fails as the single strategy for binding, as there are many more conjunctions possible than neurons available to encode them [14]. Additional conjunctions have to be computed ‘on the fly’ if required by the task, and this is one of the tasks of elemental operators.

Elemental operators

In Ullman’s theory, elemental operations are the first system level above the level of the single neuron. Table 1 lists several operations useful for vision. There are two gross categories. ‘Binding operators’ establish groupings that are not computed in the base representation, and ‘maintenance operators’ store intermediate results for use at a later point in time. In addition, there are a few operators that are not so easily classified. The list of Table 1 is probably not exhaustive, and future research might well add some elemental operators and remove others.

Binding operators

Tracing. Figure 1a illustrates a task that requires one of the binding operators – tracing. The observer’s task is to identify all squares that are connected to the yellow square. The computation of connectedness in the base representation is difficult, if not impossible, because the visual cortex cannot reserve a neuron for every possible configuration of connected squares. The tracing operator’s

job is to compute connectedness [5,15]. It can be implemented within a single retinotopic area of the visual cortex (Figure 1b; implementations that use lower as well as higher visual areas are discussed in [10,16]). When the image appears, feedforward connections from the retina activate a subset of the neurons (circles coloured brown) that form the base representation.

The tracing operator uses horizontal connections, which exist between neurons that respond to neighboring squares, to spread an increase in activity. It requires that the enhanced activity only spreads between neurons of the base representation (i.e. only neurons with a square in their receptive field). According to this rule, horizontal connections between active neurons can spread the activity enhancement, and they are therefore enabled (thick lines joining dots in Figure 1b,c), and the other connections are disabled (thin lines). We refer to the set of enabled connections as ‘interaction skeleton’, which links neurons that respond to connected squares. Small changes in the image cause small changes in the set of enabled connections, but fundamentally change the set of neurons that are linked (compare Image 1 and Image 2 in Figure 1b). Linkage by the interaction skeleton is implicit and cannot influence processing because other neurons cannot ‘see’ the pattern of enabled connections. It is the task of the tracing operator to make this grouping explicit. At step 1, an activity enhancement begins to spread from the yellow square to all neurons that respond to squares connected to it (black neurons in Figure 1c), thereby labeling all neurons that respond to a single perceptual group. Making groupings explicit by the spread of activity has been called incremental grouping [10].

We trained macaque monkeys to carry out such a curve-tracing task and obtained evidence that this operator is indeed implemented in the visual cortex. Neurons in the primary visual cortex responding to the traced curve had stronger activity than neurons responding to other, task-irrelevant contours [17]. This response-enhancement was not observed during the initial visual response, 40 ms after the presentation of the array, but it evolved across time, beginning at about 150 ms at the initial contour segments, with more distant segments evidencing this

Table 1. List of elemental operators that can be used to build visual routines

Operator	Type	Implementation
Visual search	Binding	Label spreading
Cuing	Binding	Label spreading
Trace	Binding	Label spreading
Region filling ^a	Binding	Label spreading
Association ^b	Binding	Label spreading
Working memory	Maintenance	Persistent firing
Suppression ^c	Maintenance	Active/passive inhibition
Matching ^d	Other	Unknown
Motor act ^e	Output	Body movement

^aThe region-filling operator labels regions occupied by a single perceptual object, generalizing the curve-tracing operator to 2 dimensions.

^bLinkage of arbitrary features that co-occur repeatedly.

^cRemoval of the neuronal response enhancement from features that were relevant at some point in time but lost their relevance later.

^dDetects similarities and differences between stimuli [52].

^eThe target of an upcoming eye or limb movement is specified by the focus of response enhancement in areas tuned to locations [53]. Other elemental operators can supply the information that instructs such a movement.

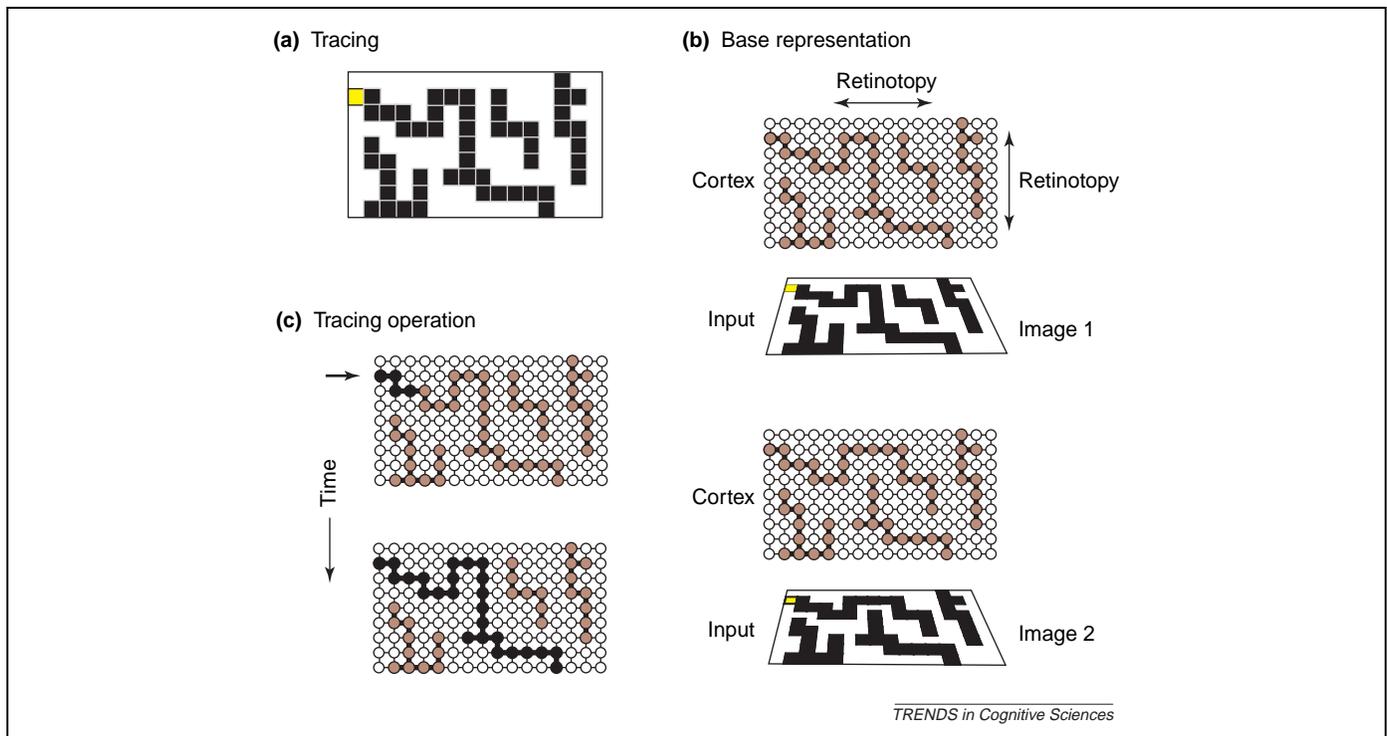


Figure 1. Implementation of the tracing operator. **(a)** Tracing task. The aim is to find all squares that are connected to the yellow one. **(b)** Tracing operator. Retinotopic area that contains a neuron for every square. Bottom-up input from the image activates some of the neurons (brown circles), and these cells form the base representation. Horizontal connections between activated neurons are enabled (thick lines), whereas the other connections are disabled (thin lines). The network of enabled connections is called interaction skeleton. A comparison of image 1 and 2 illustrates that different images enable different sets of horizontal connections. **(c)** If an enhancement of the neuronal responses (black circles) is injected at the representation of the yellow square, then it can spread through the network of enabled connections until all the squares that belong to a single object have been labelled.

response-enhancement at even later times, especially when the traced curve coursed near irrelevant curves [18]. Put another way, this response enhancement gradually spreads among the neurons that respond to connected squares until the whole object is eventually 'labelled' by an enhanced firing rate. When tracing is ready, the enhanced activity identifies all image elements that are connected to each other, and it thereby 'binds' them into a coherent object representation.

Visual search and cuing. Visual search and spatial cuing can also be viewed as binding operators. In search, the observer's task is to determine the location of a predefined shape (Figure 2a). Conversely, in cuing, a location is given and the task is to report the features of the object at that location. Figure 2 illustrates the key ideas behind neural network models of search and cuing [19–21], here implemented as an interaction between three visual areas. The first area contains neurons with large receptive fields (RFs) encoding colours irrespective of their location (Figure 2b: TI, or translational invariant). The second consists of neurons that encode the location of relevant objects irrespective of their features as a saliency map (SaM) [22], and the third area is an early visual area (EV) where neurons are tuned to conjunctions between colours and locations (a form of base-grouping).

The model assumes that neurons in EV tuned to a particular colour and retinotopic location are reciprocally connected to TI units that encode the same colour, and to SaM units that encode the same location. In Figure 2b, two objects are presented, and feedforward connections

activate cells in TI tuned to two colours, neurons in SaM tuned to two locations, and cells in EV tuned to the respective colour–location conjunctions (grey circles in Figure 2b) are neurons activated by feedforward connections). In this base representation, the internal skeleton of connections between the activated neurons is enabled (thick lines). A different input pattern activates different neurons in the base representation and thereby enables a different pattern of connections. Note that for both images, only neurons that encode features of the same object are linked. The pattern of enabled connections is implicit; that is, not visible to neurons in other areas. To make the groupings explicit, elemental operators have to spread a response enhancement through the network. If the task is to search for the yellow item, for example, then the activity of TI units tuned to yellow is facilitated (Figure 2c). The enhanced activity will spread through EV to neurons in SaM that encode the location of the yellow object. Conversely, if the task is to identify the colour at a given location, as in a spatial cuing task, then the activity enhancement is injected in SaM and spreads to neurons in TI encoding the appropriate colour (Figure 2d).

In this simple example little may seem to be gained by the incremental grouping process, as conjunctions between colours and locations are also available as base groupings in EV. Modeling studies demonstrated, however, that the scheme can be generalized to encode conjunctions between complex shapes and locations not encoded as base groupings (i.e. not repeated at every retinotopic location) [19,23]. Importantly, the proposed implementation is supported by neurophysiological data

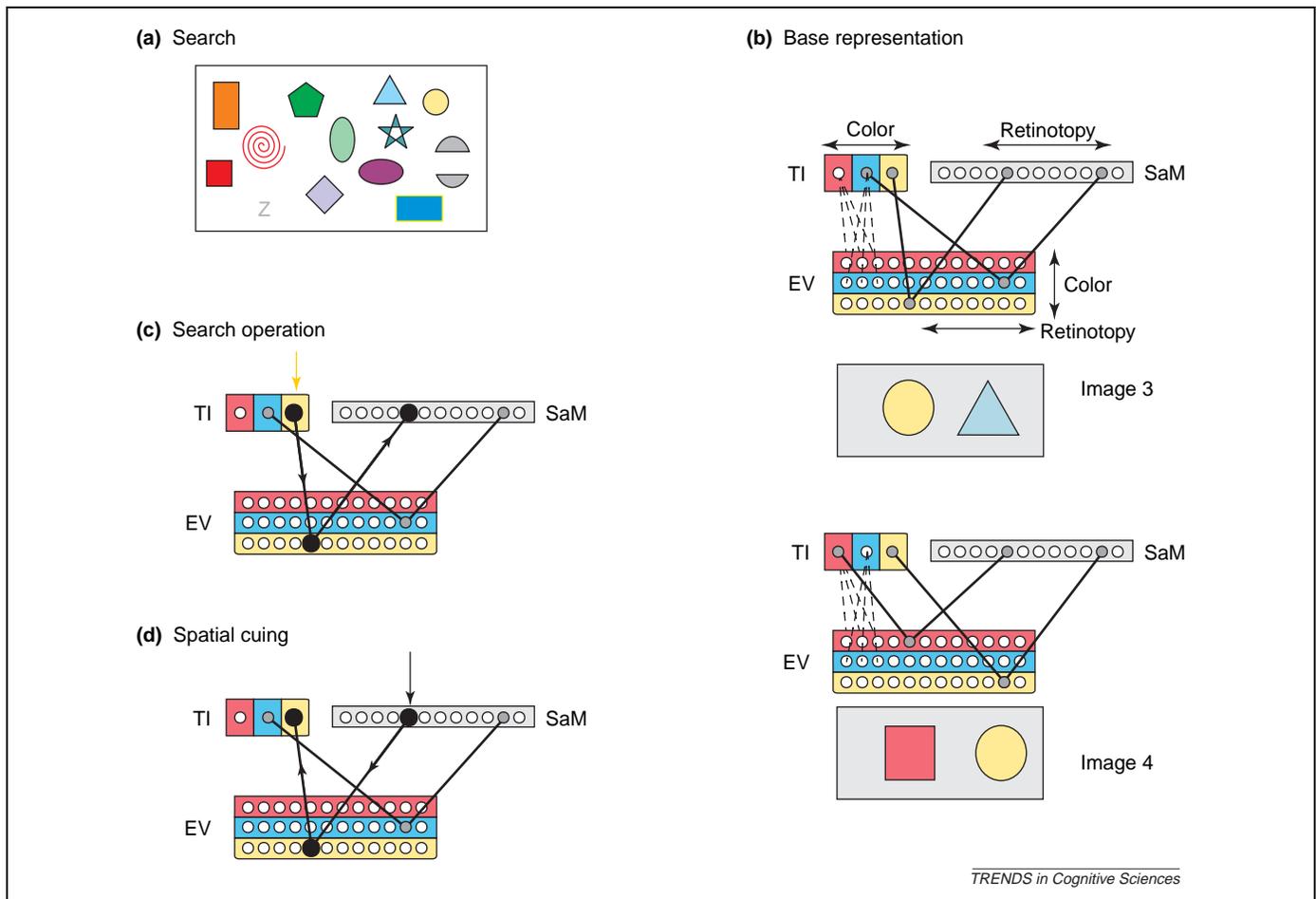


Figure 2. Implementation of search and cuing. **(a)** Search task. Where is the yellow object? **(b)** Models of visual search include multiple visual areas. TI is a colour-selective area with neurons that have a translationally invariant representation (large RFs). EV is an early visual area with colour selective neurons that have a small RF. SaM is a saliency map; an area with neurons tuned to spatial locations, but not selective for other visual features. Neurons in EV are reciprocally connected with neurons in TI with a similar colour tuning, and with neurons in SaM with a similar spatial tuning. Bottom-up input from the image activates a subset of the neurons (grey circles are the active neurons), and it thereby enables a subset of the connections (thick lines). Several disabled connections are shown as dashed lines. A comparison of image 3 and 4 illustrates that different images enable different sets of interareal connections. **(c)** During search for yellow, neurons tuned to yellow in TI increase their response. The enhanced activity spreads through EV to the neurons in SaM that encode the location where the yellow item can be found. **(d)** In a cuing task, neurons in SaM with a RF at the cued location enhance their response. This response enhancement spreads through EV to neurons in TI that encode the colour present at the cued location.

as neurons in early visual areas [24–27], areas that encode shapes [28] as well as in areas with saliency maps [29,30] enhance their response if they encode features of the target object during search and cuing.

Distal access. Binding operators provide access to all features of an object (distal access; see Box 1), in spite its distributed representation across many areas of visual cortex. The operators ensure that it does not matter whether an object is cued by location, shape, or one of its contours. Studies on the time required by binding operators consistently provide estimates between 100 and 300 ms [17,25,31]. Thus, as predicted by Newell [2], the time required to realize distal access is one order of magnitude longer than the time required for an interaction between neighboring cells (~ 10 ms) [32], and two orders of magnitude longer than an action potential (~ 1 ms). These neurophysiological findings also have clear-cut descriptions in psychology. Visual search, cuing, and also tracing involve shifts of visual attention [33–36]. Taken together then, descriptions from psychology (visual attention), neurophysiology (rate

enhancement), and artificial intelligence (distal access) have started to converge.

Maintenance operators

It is not enough to select an object, but the observer must be able to maintain the object in memory for future cognitive manipulations. Working memory is the key process that permits the temporary storage of information [37]. Features that are attended, that is, labelled by an enhanced response, have a higher probability of being remembered. In many areas of the visual and frontal cortex, the neuronal response enhancement for a to-be-remembered item persists when the visual target is removed from sight [8,38,39]. Thus, many of the areas that implement elemental operations can maintain the results of these computations during memory delays. It is also possible to execute an eye- or limb movement to maintain information across memory delays (see Box 2 for an example). Having listed several the elemental operations, I will now turn to the question of how they can be arranged in a sequence to solve more complex tasks.

Box 2. Reading a map

Imagine driving with your partner en route to a camping site. You have just passed a church and now you approach an unexpected fork in the road. You therefore ask your partner whether you should turn to the left or to the right. After a few seconds of looking at his or her map, your partner replies 'left'. To answer this question, he or she must find the road along which you are traveling, the final destination, and then determine which road will bring you to the destination.

When she directs her gaze to the map, feedforward connections rapidly extract various visual features (base representation in Fig. 1a). Thereafter, several elemental operators must be executed. Because your partner does not know where you are, she begins by conducting two visual searches, one for the camping-site to which you are headed and one for the church that you just passed. These searches can be conducted in either order; however, let's assume that she looks for the camping-site first. As a result, the location of the camping site is labelled with an enhanced neuronal response in areas with a map of the visual field (Fig. 1b). This location must be stored for future use, as your partner executes the second search for the church. She may put her finger on the map where the camping site was found ('2' in Fig. 1b), while looking for the church [45]. The second search produces the location of the church and labels it with an enhanced response (Fig. 1c).

The location of response enhancement can be used as the starting point for the subsequent trace operation. The trace operator labels the first part of the route, but it then encounters the junction. Some evidence suggests that participants only trace one curve at a time [51], I will therefore assume that your partner follows the right turn and eventually finds that this road is blocked (Fig. 1d). She identifies the cause of this blockade as a village near a river (the response enhancement at the road's end provides a spatial cue). She suppresses this part of the road to ensure that she will not take the erroneous turn twice (maintenance operator, Fig. 1e). After another tracing operation, she finds the correct route (Fig. 1f).

This example illustrates how a small set of operations can be used to solve a relatively complex task. At present, the explanatory power of this theory is limited because I did not specify how the appropriate operators get selected at each step of the routine. In cognitive theories, such as ACT [1] and SOAR [2], the order of these operators is controlled as items that are added or removed from working memory to bias the selection of subsequent operations. At the start of this task, for example, the content of working memory would bias the selection of search operators. The exploration of these control processes constitutes an important avenue for future research.

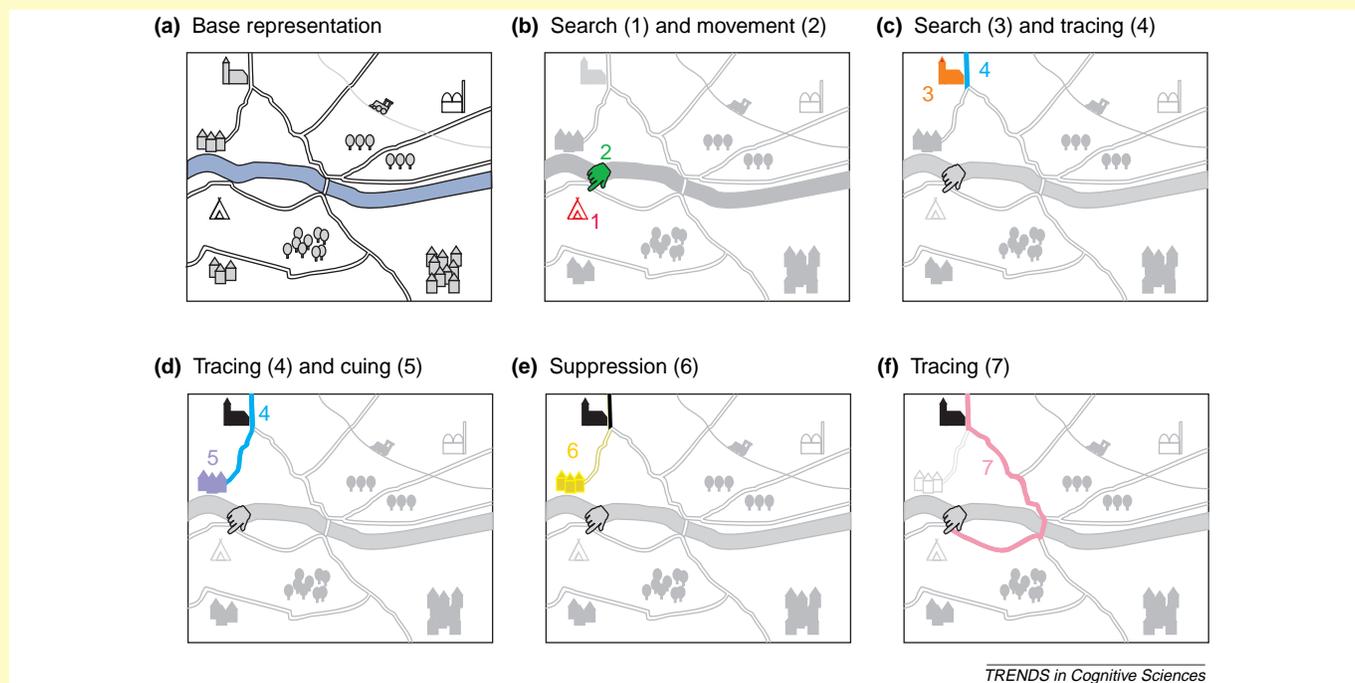


Figure 1. A sequence of elemental operations in map reading (see text for details).

Visual routines

Visual routines represent a system level higher than that of the elemental operators [5] (e.g. see Box 2). There are only a limited number of elemental operations; nonetheless, it is possible to construct a virtually unlimited set of visual routines as the operations can form many different sequences [5,10]. As each operation typically requires on the order of 100–300 ms to be executed, a routine can take a second or several seconds. This section addresses two issues regarding the organization and execution of routines: (1) how information is transferred from one elemental operator to the next, and (2) what mechanisms determine the order of operations.

To gain insight in these issues, we recently investigated a simple routine in the visual cortex of

macaque monkeys [7]. The task, illustrated in Figure 3a, required the monkey to fixate on a central grey marker (fixation point) that changed its colour (to either yellow or blue). The colour of the fixation point instructed the monkeys to search for a marker with the same colour. Thus, in the trial illustrated in Figure 3a, the monkey had to search for a blue marker. This marker represented the beginning of a target curve (T in Figure 3a) that the monkey had to trace to locate a larger circle at the other end that was the target for an eye movement.

This task can be solved by a routine composed of two operators, first search, then tracing. We measured neuronal activity in area V1 as the monkeys performed this task. The initial registration of the stimuli occurred at a latency of 39 ms (grey in Figure 3b); this corresponds to

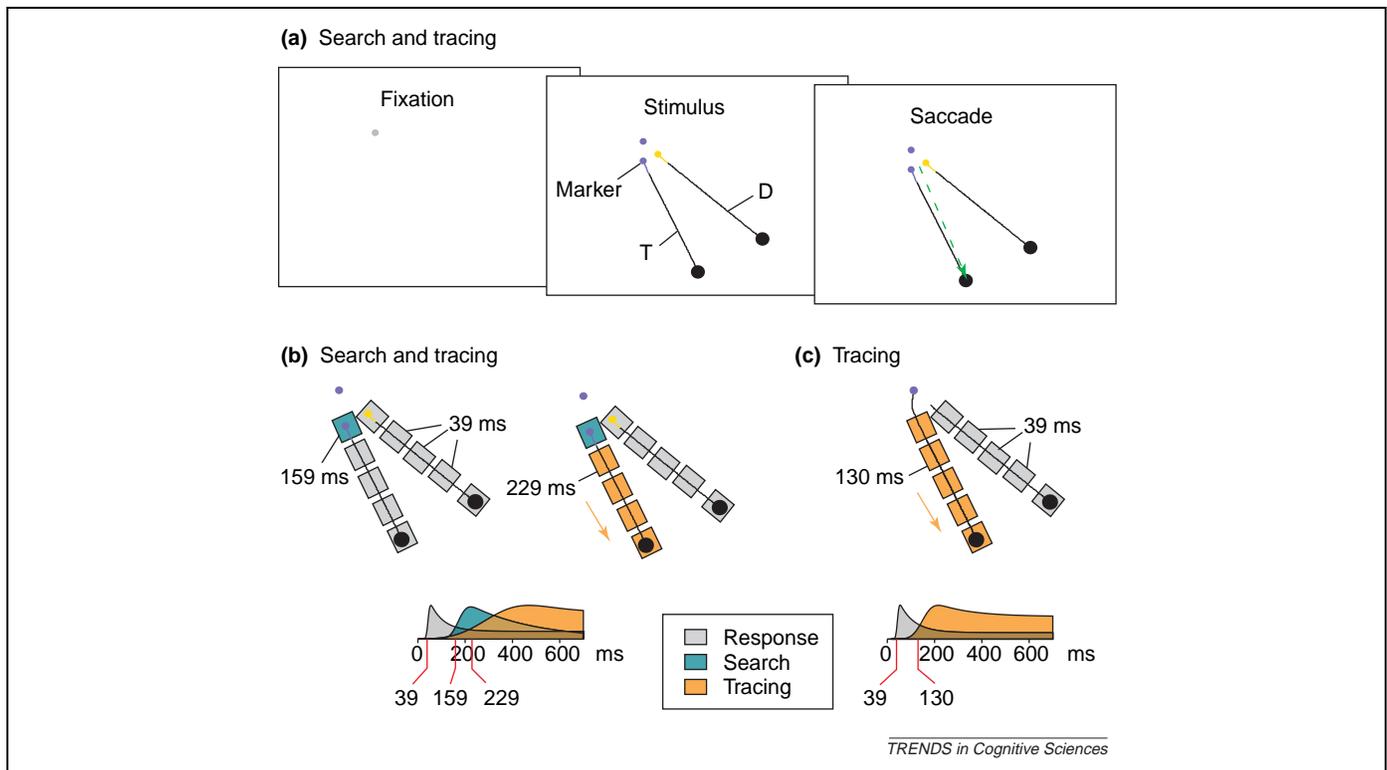


Figure 3. Neuronal responses in monkey area V1 during a visual routine. (a) The animal's task was to locate a marker that had the same colour as the fixation point (here blue) and to trace the curve (target curve, T) connected to that marker. The animal had to locate a larger circle at the other end of this curve, which was the target for an eye movement. The other curve (D) was a distractor. (b) Time course of neuronal activity in area V1. Thirty-nine ms after stimulus onset, all the neurons with RFs that fall on the curves are activated (grey rectangles). After 159 ms, neuronal responses to the blue colour are enhanced (green rectangle). This response enhancement can provide the starting point for the tracing operator that enhances responses evoked by the target curve at a latency of 229 ms (orange). Shown below are the average time course of the initial visual response (grey), the response enhancement due to search (green), and tracing (orange). (c) Simple task where the monkey does not have to search, as he can start to trace at the fixation point. The latency of visual response does not change (grey), but the response enhancement due to tracing now occurs after 130 ms, 100 ms earlier than in the more complex task.

what we called base representation. One hundred and twenty milliseconds later (159 ms after display onset), the neurons representing the cued colour in their receptive field enhanced their response (green in Figure 3b). This response enhancement was presumably caused by a colour-selective feedback signal from higher visual areas (see Figure 2c) that labelled the colour at the beginning of the target curve. After an additional 70 ms (229 ms after stimulus appearance), neuronal activity in V1 differentiated the target curve from the distractor curve, because neuronal responses evoked by the target curve were enhanced (orange in Figure 3b).

The monkeys also performed a simpler task that could be solved by just tracing (Figure 3c). The onset of the initial volley of activity in V1 was the same as in the complex task, but now the response enhancement caused by tracing occurred at 130 ms; that is, 100 ms earlier. Thus in the complex task, the search operator delayed the subsequent tracing operator, suggesting that these operators are executed serially.

Information transfer

In such a routine it is essential that information from one elemental operation is transferred to the next. The data of Figure 3 provide insight into how the visual cortex transfers information between elemental operators. The result of the search operator is a response enhancement of neurons in retinotopic areas (Figure 3, green RFs; see also

[25,40]), which can be used as starting point for the tracing operator. Thus, the enhanced response in retinotopic areas transfers the position information from the search operator to the tracing operator – this transfer is called 'call-by-focus'. Call-by-focus is a very general mechanism, as other variables, like shapes, colours, or motion directions can be transferred similarly, as a maintained response enhancement in areas tuned to these features [7]. The persistence of this response enhancement corresponds to the temporary storage of features in working memory [41–44] (Table 1).

Sequencing of operators

In the task of Figure 3a, the monkey first searched and then traced, in a fixed sequence. The order of operators may be more variable in other tasks, and subjects may even switch between strategies from trial to trial (e.g. [45]). Moreover, when the task is new, the participant may use trial and error to find the most efficient sequence of operations (Box 2). This raises the issue of control: what controls the sequencing of operators?

This question has, to my knowledge, not been answered satisfactorily at the neurophysiological level. By contrast, artificial intelligence models, such as ACT-R [1] and SOAR [2] are specifically designed to deal with issues of control. They make two assumptions: (i) the operations compete for control over behavior, and (ii) only one operation can be executed at one point in time. In the models, the operators'

competitive fitness depends on several factors. One factor is the current contents of working memory. Selected operators place their results in working memory, and they thereby influence the selection of subsequent operators. A second factor is the success of an operator's previous applications – the fitness of successful operators is increased, so that they are more likely to be selected when a similar problem reoccurs. This can eventually lead to the formation of fixed sequences called chunks [1,2], which is beneficial when sequences work well in particular situations.

Conclusions and future directions

A better understanding of control and chunking at the neurophysiological level would be of great value for our understanding of cognition generally. Some insight into issues of control comes from recent work on perceptual decisions and how they give rise to eye-movement commands (for reviews see [46,47]). This raises an important question for future research: do the processes that select elemental operators resemble those that select limb and eye-movements? If so, then insight into the formation of chunks might also have been gained, from experiments where monkeys were trained to carry out arm movements in fixed sequences. These studies revealed neurons in regions of motor cortex that are selective for entire movement sequences [48,49]. If this generalizes to visual routines, then the sequences of elemental operations might originate from areas outside the visual cortices; for instance, a recent study in the prefrontal cortex of monkeys has described neurons that encode abstract rules of how the monkey should process visual information [50]. An exciting prospect is to investigate how prefrontal neurons influence the selection of operators implemented in visual areas.

In conclusion, I have proposed that it is necessary to define new system levels that will bridge the gap between the individual neurons and cognition. As an example, I outlined an emerging theory of how networks of neurons spanning several areas of the cortex interact with each other to implement elementary operators in vision. I indicated how successive operators in vision are arranged in a sequence so that routines can be built that solve more complex tasks. Importantly, the elemental operators have correlates in psychology (shifts of attention), neurophysiology (firing rate enhancement), as well as in artificial intelligence (distal access), and the present theoretical framework could therefore foster direct communications between these fields. Indeed, the neurophysiological results reviewed here demonstrate that it is in fact possible to monitor routines in the visual cortex, and they testify that cognitive neurophysiology is entering an exciting era where it can start to test theories in artificial intelligence [5,45].

Acknowledgements

I thank Jillian Fecteau and Roos Houtkamp for their helpful comments on the manuscript. This work was supported by a HFSP Young Investigators grant.

References

- Anderson, J.R. and Lebiere, C. (1998) *The Atomic Components of Thought*, Erlbaum
- Newell, A. (1990) *Unified Theories of Cognition*, Harvard University Press
- Hupé, J.M. et al. (1998) Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature* 394, 784–787
- Moore, T. and Armstrong, K.M. (2003) Selective gating of visual signals by microstimulation of frontal cortex. *Nature* 421, 370–373
- Ullman, S. (1984) Visual routines. *Cognition* 18, 97–159
- Formisano, E. et al. (2002) Tracking the mind's image in the brain I: time-resolved fMRI during visuospatial working memory. *Neuron* 35, 185–194
- Roelfsema, P.R. et al. (2003) Subtask sequencing in the primary visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* 100, 5467–5472
- Rainer, G. et al. (1998) Selective representation of relevant information by neurons in the primate prefrontal cortex. *Nature* 393, 577–579
- Lamme, V.A.F. and Roelfsema, P.R. (2000) The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* 23, 571–579
- Roelfsema, P.R. et al. (2000) The implementation of visual routines. *Vision Res.* 40, 1385–1411
- Thorpe, S. et al. (1996) Speed of processing in the human visual system. *Nature* 381, 520–522
- Felleman, D.J. and Van Essen, D.C. (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* 1, 1–47
- Treisman, A.M. (1996) The binding problem. *Curr. Opin. Neurobiol.* 6, 171–178
- von der Malsburg, C. (1999) The what and why of binding: the modeler's perspective. *Neuron* 24, 95–104
- Jolicoeur, P. et al. (1986) Curve tracing: a possible basic operation in the perception of spatial relations. *Mem. Cogn.* 14, 129–140
- Edelman, S. (1987) Line connectivity algorithms for an asynchronous pyramid computer. *Comp. Vision Graphics Image Process.* 40, 169–187
- Roelfsema, P.R. et al. (1998) Object-based attention in the primary visual cortex of the macaque monkey. *Nature* 395, 376–381
- Roelfsema, P.R. and Spekreijse, H. (1999) Correlates of a gradual spread of attention over a traced curve in macaque area V1. *Soc. Neurosci. Abstr.* 25: 7.2
- van der Velde, F. and de Kamps, M. (2001) From knowing what to knowing where: Modeling object-based attention with feedback disinhibition of activation. *J. Cogn. Neurosci.* 13, 479–491
- Hamker, F.H. (2005) The reentry hypothesis: the putative interaction of the frontal eye field, ventrolateral prefrontal cortex, and areas V4, IT for attention and eye movement. *Cereb. Cortex* 15, 431–447
- Phaf, R.H. et al. (1990) SLAM: A connectionist network for attention in visual selection tasks. *Cogn. Psychol.* 22, 273–341
- Koch, C. and Ullman, S. (1985) Shifts in selective visual attention: towards the underlying neural circuitry. *Hum. Neurobiol.* 4, 219–227
- Fukushima, K. (1988) A neural network for visual pattern recognition. *IEEE Computer* 21, 65–75
- Desimone, R. and Duncan, J. (1995) Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222
- Chelazzi, L. et al. (2001) Responses of neurons in macaque area V4 during memory-guided visual search. *Cereb. Cortex* 11, 761–772
- Motter, B.C. (1994) Neural correlates of attentive selection for color or luminance in extrastriate area V4. *J. Neurosci.* 14, 2178–2189
- Treue, S. and Maunsell, J.H.R. (1999) Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas. *J. Neurosci.* 19, 7591–7602
- Chelazzi, L. et al. (1993) A neural basis for visual search in inferior temporal cortex. *Nature* 363, 345–347
- Schall, J.D. and Thompson, K.G. (1999) Neural selection and control of visually guided eye movements. *Annu. Rev. Neurosci.* 22, 241–259
- Colby, C.L. and Goldberg, M.E. (1999) Space and attention in parietal cortex. *Annu. Rev. Neurosci.* 22, 319–349
- Luck, S.J. et al. (1997) Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *J. Neurophysiol.* 77, 24–42
- Nowak, L.G. et al. (1997) Corticocortical connections between visual

- area 17 and 18a of the rat studies *in vitro*: spatial and temporal organisation of functional synaptic responses. *Exp. Brain Res.* 117, 219–241
- 33 Houtkamp, R. *et al.* (2003) A gradual spread of attention during mental curve tracing. *Percept. Psychophys.* 65, 1136–1144
- 34 Posner, M.I. (1980) Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25
- 35 Kim, M-S. and Cave, K.R. (1995) Spatial attention in visual search for features and feature conjunctions. *Psychol. Sci.* 6, 376–380
- 36 Treisman, A.M. and Gelade, G. (1980) A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136
- 37 Cowan, N. (2001) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav. Brain Sci.* 24, 87–185
- 38 Fuster, J.M. and Jervey, J.P. (1981) Inferotemporal neurons distinguish and retain behaviorally relevant features of visual stimuli. *Science* 212, 952–955
- 39 Funahashi, S. *et al.* (1993) Prefrontal neuronal activity in rhesus monkeys performing a delayed anti-saccade task. *Nature* 365, 753–756
- 40 Schall, J.D. and Hanes, D.P. (1993) Neural basis of saccade target selection in frontal eye field during visual search. *Nature* 366, 467–469
- 41 Funahashi, S. *et al.* (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* 61, 331–349
- 42 Miller, E.K. and Cohen, J.D. (2001) An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202
- 43 Supèr, H. *et al.* (2001) A neural correlate of working memory in the monkey primary visual cortex. *Science* 293, 120–124
- 44 Gnadt, J.W. and Andersen, R.A. (1988) Memory related motor planning activity in posterior parietal cortex of macaque. *Exp. Brain Res.* 70, 216–220
- 45 Ballard, D.H. *et al.* (1997) Deictic codes for the embodiment of cognition. *Behav. Brain Sci.* 20, 723–767
- 46 Gold, J.I. and Shadlen, M.N. (2001) Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* 5, 10–16
- 47 Glimcher, P.W. (2003) The neurobiology of visual-saccadic decision making. *Annu. Rev. Neurosci.* 26, 133–179
- 48 Tanji, J. and Shima, K. (1994) Role for supplementary motor area cells in planning several movements ahead. *Nature* 371, 413–416
- 49 Bullock, D. (2004) Adaptive neural models of queuing and timing in fluent action. *Trends Cogn. Sci.* 8, 426–432
- 50 Wallis, J.D. *et al.* (2001) Single neurons in prefrontal cortex encode abstract rules. *Nature* 411, 953–956
- 51 Jolicoeur, P. (1988) Curve tracing operations and the perception of spatial relations. In *Computational Processes in Human Vision: An Interdisciplinary Perspective* (Pylyshyn, Z., ed.), pp. 133–168, Ablex
- 52 Miller, E.K. and Desimone, R. (1994) Parallel neuronal mechanisms for short-term memory. *Science* 263, 520–522
- 53 Colby, C.L. (1998) Action-oriented spatial reference frames in cortex. *Neuron* 20, 15–24

ESF–EMBO Symposia on Three-Dimensional Sensory & Motor Space Probabilistic Mechanisms of Learning and Development in Sensorimotor Systems

Sant Feliu de Guixols (Costa Brava), Spain, 8–13 October 2005
Chair: Daniel Wolpert (University College London, UK)

This meeting will cover neurophysiological, molecular, psychological, neuroimaging and computational approaches to sensorimotor development and learning. One emphasis of the meeting will be to explore an evolving idea in the field; that the statistics of both the motor and sensory apparatus as well as statistics of the outside world determine through probabilistic mechanisms the processes underlying development and learning.

Further information available at <http://www.esf.org/conferences/lc05110>

Contact: European Science Foundation • ESF Research Conferences Unit • 1 quai Lezay-Marnésia
BP 90015 • 67080 Strasbourg Cedex • France

Tel. +33 (0)388 76 71 35 • Fax +33 (0)388 36 69 87 • E-mail: conferences@esf.org
<http://www.esf.org/conferences>