Attention and eye movements during contour mapping
Khayat, P.S.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Subtask Sequencing in the Primary Visual Cortex

ABSTRACT

Complex visual tasks can usually be decomposed into a number of simpler subtasks. Whether such subtasks are solved serially or in parallel is subject to considerable debate. Here we investigate how subtasks are coordinated in time by recording from the primary visual cortex (area V1) of macaque monkeys. The animals were trained to perform both a simple and a composite task. In the simple task, they had to mentally trace a target curve while ignoring a distractor curve. Neuronal responses in area V1 to the target curve were enhanced relative to responses to the distractor curve, 130 ms after stimulus appearance. In the composite task, the monkeys searched for a coloured marker, and traced a curve that was attached to this marker. In an initial phase of the trials neuronal responses reflected visual search, and the response enhancement due to curve tracing now occurred after 230 ms, 100 ms later than in the simple task. We conclude that subtasks of the composite task are carried out in a structured and sequential manner that can be monitored in the primary visual cortex.

1. INTRODUCTION

Visual processing is usually subdivided into an early pre-attentive, and a later attentive stage (1,2). Pre-attentive processing presumably corresponds to the rapid (within 100 ms) activation of feature selective neurons in the many visual cortical areas (3,4). However, many tasks depend on more elaborate processing capabilities that are attributed to visual attention. One example is visual search (2, 5), where the task is to localise a target item with a known attribute (shape, colour or other feature). During visual search, neurons in the visual cortex that encode the attribute that is searched enhance their activity (6-8), and this causes an activity increase of neurons that encode the target item’s location (9-14). At a psychological level of description, attention is directed to the location of the target item (15). Another task that requires attention is curve tracing (16) where two or more curves are displayed, and the subject has to identify all contour segments that belong to one of them. Attention is directed to the traced curve (17), and neuronal responses in the visual cortex evoked by this curve are enhanced (18). The modulation of firing rates caused by curve tracing and by visual search does not occur during the initial visual responses evoked by stimulus onset, but rather after a variable delay of 120 to 200 ms (7-10, 18) that is related to the subject’s reaction time (14). These results, taken together, indicate that the enhancement of neuronal responses in the visual cortex reflects the dynamics of attention shifts that are imposed by the task demands.

Complex visual tasks may be decomposed into subtasks, and require the application of multiple attentional operators to the visual input (19-22). It is, however, unclear how subtasks are coordinated in time (23). Some theories hold that subtasks are usually solved in parallel (24,25), whereas others hold that they are solved serially, and that the subject’s reaction time is the sum of times required by all subtasks (19,26-28). Here we set out to resolve this debate by a direct measurement of the time course of attentional operations in the visual cortex. We will compare the time course of the attentional response enhancement in area V1 between a
simple task that requires only curve tracing, and a composite task that requires visual search as well as curve tracing.

2. METHODS

Recording technique. Experiments were performed with two macaque monkeys using standard procedures that have been described (18). In brief, operations were performed under aseptic conditions and general anesthesia, which was induced with ketamine (15 mg/kg i.m.), and maintained after intubation by ventilating with a mixture of 70% N₂O and 30% O₂, supplemented with 0.8% isoflurane, fentanyl (0.005 mg/kg i.v.), and midazolam (0.5 mg/kg-h i.v.). The animals recovered for at least 21 days before training was resumed. All procedures complied with the NIH Guide for the Care and Use of Laboratory Animals (National Institute of Health, Bethesda, MD), and were approved by the institutional animal care and use committee of the Royal Netherlands Academy of Arts and Sciences. In a first operation a head holder was implanted, and a gold ring was inserted under the conjunctiva of one eye for the measurement of eye position. Multi-unit recordings were obtained from electrodes that were chronically implanted in area V1 in a separate operation (40-65 teflon coated platinum-iridium wires per hemisphere), and positioned 1-2 mm below the cortical surface.

Behavioural tasks. The monkeys were trained to perform two tasks. In both tasks, a trial was started as soon as the monkey's eye position was within a 1°x1° square window centred on a 0.3° fixation point. The stimulus appeared after an interval of 300 ms (Fig. 1). The first task was a simple curve tracing task. The monkey had to trace a target curve (T in Fig. 1A) that was connected to the fixation point to locate a circle at its other end, and to make a single eye movement to this circle. There also was an other, distractor curve (D in Fig. 1A) that was not connected to the fixation point and that could be ignored. The curves were composed of two or three third order polynomials (bezier curves), joined end to end, and were grey (17 cd·m⁻²) on a black background. The colour of the fixation point was blue on half of the trials, and yellow on the other half, but this colour was irrelevant. The second task was a composite task that required both visual search and curve tracing (Fig. 1B). A trial began when the monkey fixated a grey fixation point. After a brief fixation delay (300 ms), the colour of the fixation point changed to blue or yellow, and at the same time two curves and two coloured markers appeared on the screen. The monkey had to search for a marker with the same colour as the fixation point. This marker indicated the beginning of the target curve. The animal had to trace this curve to its other end to locate a larger circle that was the target for an eye movement (eccentricity of this circle ranged from 5 to 9.5 deg). The monkey was allowed to make only a single eye movement from the fixation point to this circle. The left marker was yellow on half of trials and blue on the other trials, and the right marker had the other colour. The various colour combinations resulted in a total of 4 (standard task, Fig. 3) or 8 (experiment with additional contour segment, Fig. 4) stimuli that were randomly interleaved. Reaction times were measured in psychophysical experiments, and in these experiments the monkeys could...
initiate their saccade immediately after presentation of the stimulus. In the experiments in which we recorded neuronal activity, however, the animals had to maintain their gaze on the fixation point for 700 ms after stimulus appearance. Then the fixation point disappeared, which was the cue to initiate a saccade. The two tasks were interleaved in blocks of 60-72 trials, and each session comprised 3 to 9 blocks per task.

**Figure 1.** Behavioural tasks. (A) Curve tracing task. The monkeys mentally traced a curve connected to the fixation point to locate a circle at the other end of this curve. A single eye movement had to be made to this circle. T, target curve. D, distractor curve. (B) In the composite task, the colour of the fixation point was initially grey, and changed to either yellow or blue upon stimulus appearance. The monkeys had to locate a marker with the same colour that indicated the start of the target curve. They had to mentally trace this curve to its other end to locate a circle that was the target for an eye movement.

**Recording and data analysis.** Recordings with a sufficient signal to noise ratio were obtained from about 50% of the chronically implanted wires. For these recording sites, receptive field (RF) dimensions were measured as described previously (29). RF eccentricity ranged from 1.7° to 6.4°. Median RF size was 1.0 deg². In the composite task (Fig. 1B), the receptive fields always fell on a grey segment of the target or distractor curve. Only correct trials were included in the analysis. Responses at individual recording sites to the various stimuli were normalised to the average peak response, a procedure that preserves differences in the peak response between stimuli (18). Population responses were computed by averaging across the normalised responses at different recording sites. The latency of a response enhancement was determined by fitting a curve to the difference between the response to the target and distractor curve. It was (arbitrarily) defined as the time that the fitted function reached 33% of its maximal value. The results that are reported here do not depend on the exact value of this arbitrary criterion. Appendix B provides a complete description of the method used to determine the latency of the visual response and the latency of attentional response modulation, as well as the method used to analyze the significance of differences in latencies between tasks, which is based on a Monte-Carlo procedure.
3. RESULTS

Behavioural performance. A first, psychophysical experiment compared reaction times between tasks. The average reaction times in the simple curve tracing task for monkeys B and G were 251 and 222 ms, respectively (performance was better than 99% correct) (Fig. 2A). In the composite task, the reaction time of monkey B increased by 141 ms (P<10⁻⁶, U-test), and his performance was reduced to 94% correct. The increase in reaction time of monkey G was 48 ms (P<10⁻⁶, U-test), a measure that presumably underestimates the additional processing time required for this task, since his performance dropped to 78%. Fast responses of monkey G were less accurate than slow responses, as if he responded on percentage of trials before he had solved the task (Fig. 2B).

![Figure 2](image-url)

Figure 2. Behavioural performance. (A) Reaction times in the simple (hatched bars) and composite task (black bars). Error bars indicate standard deviation. (B) Dependence of performance on reaction time in the composite task. For each monkey saccadic reaction times were divided between three equally large classes; fast, intermediate and slow. Abscissa, average reaction time in these categories. Ordinate, performance. Dashed line, chance level. Error bars, standard error of the mean.

Recordings from area V1. We then studied neuronal activity in area V1. During recording sessions, the monkeys had to maintain their gaze on the fixation point for 700 ms after the appearance of the stimulus. Figure 3A illustrates the location of the receptive field of a group of V1 neurons (multi-unit activity) relative to the stimuli of the simple curve tracing task. The receptive field was either on the target curve, which connected the fixation point to the circle that was the target for an eye movement, or on the distractor curve, but the contour segment inside the receptive field was identical for all stimuli. Activity was pooled across stimuli (with a different colour of the fixation point) for which the receptive field was on the target curve, and across stimuli for which it was on the distractor curve. Responses to the target curve (grey in Fig. 3C) were stronger than responses to the distractor curve (black) (P<10⁻⁶, U-test). Note
that the response to the target curve gradually deviates from the response to the distractor curve. To measure the latency of the response enhancement, the response to the distractor curve was subtracted from the response to the target curve (Fig. 3C, lower panel), and a curve was fitted to the response difference. The latency was defined as the moment that the response enhancement reached 33% of its maximal value, and it was equal to \(122 \pm 2\) ms (estimate \(\pm\) sd).

In the composite task, the neurons responded to the same contour segment that was placed inside their receptive field (Fig. 3B). In this task, the neurons also enhanced their response if their receptive field fell on the target curve (\(P<10^{-6}\), U-test). However, now the latency of this response enhancement increased to \(264 \pm 5\) ms (Fig. 3D), which was significantly longer than the latency in the simple task (\(P<0.001\)). Similar results were obtained across a population of 34 recording sites in two monkeys (Fig. 3E,F). The target curve evoked stronger responses than the distractor curve, in both tasks (paired t-test, \(P<2.10^{-6}\) in each task). At the population level, the latency of the response enhancement in the simple task was \(130 \pm 2\) ms, and it increased to \(229 \pm 12\) ms in the composite task (\(P<0.001\)). Thus, if the monkey also has to perform a visual search, the response enhancement to the target curve is delayed by about 100 ms. This additional delay is compatible with the average increase in reaction time of 90 ms (Fig. 2).
Previous studies suggested that visual search is implemented by a feature-selective top-down signal to retinotopic areas, which enhances the response of neurons that encode the target feature (5,7-9,21,30). To directly measure the time-course of this putative feedback signal in the composite task, unconfounded by curve tracing, a small change was made to the stimuli. An additional contour segment was added to the stimuli, and placed in the receptive field of a group of V1 neurons (Fig. 4A). This contour segment was irrelevant for the monkeys’ task, and had a colour that either matched or mismatched the target colour. Figure 4B illustrates the response of a group of neurons in area V1 to this irrelevant contour segment. Responses were averaged across the four stimuli for which the segment’s colour matched and across the four stimuli for which it mismatched the target colour. This dissociates the effects of colour matching from the neurons’ colour tuning, the position of the coloured markers, and the location of the target curve (31). Responses at this recording site were enhanced transiently if the colour in the receptive field matched the target colour (P<0.01, U-test, latency = 166±5 ms). Similar effects were observed across a population of 28 recording sites in area V1. Responses were stronger if the colour in the receptive field matched the target colour (P<0.001, paired t-test, N=28). This response enhancement was also significant at 8 individual recording sites (P<0.01, U-test at each site), but significant response reductions were not observed. Responses were averaged across these 8 sites to measure the latency of this effect, which was 159±6 ms (Fig. 4C). The response enhancement was also relatively transient at the population level, as if the colours lost their relevance during the later stages of a trial.

A control experiment investigated the responses to the irrelevant contour segment during the simple curve tracing task, which did not require color search (Fig. 4D). In this task, the strength of the neuronal responses did not depend on the relationship between the color of the fixation point and the color in the receptive field, neither at any of the 8 recording sites that exhibited a match-enhancement in the composite task (P>0.05, U-test at each site) (Fig. 4E), nor at the population level (P>0.05, paired t-test, N=28). This confirms that the response enhancement in the composite task reflects visual search, which appears to be implemented by a color-selective feedback signal to retinotopic areas (feature-based attention, see also refs 12,32).

Figure 3 (opposite). Temporal profile of response enhancement due to curve tracing in the simple and composite task (A,B) Location of the receptive field (grey rectangle) of a group of neurons in area V1 relative to the stimuli of the simple (A) and composite task (B). The receptive field was on the target [T] or distractor curve [D]. (C,D) Neuronal response to the target (grey) and distractor curve (black), during the simple (C) and composite task (D). The lower panels show the difference between the response to the target and distractor curve, and a curve that was fitted to this response difference, in order to measure the latency of the response enhancement. The latency (arrow) was defined as the time at which the response enhancement reached 33% of its maximum. Black bar, 95% confidence interval of the latency. (E,F) Population responses across 34 recording sites to stimuli of the simple (E) and composite task (F). The lower panels of E and F are scaled differently than the upper panels, to reveal the time course of modulation (see scale bar).
Figure 4 Feature-based attention during colour search. (A) An irrelevant contour segment was placed in the receptive field of a V1 recording site. This contour segment had the same colour as the fixation point (colour-match, upper stimuli) or the other colour (colour-nonmatch, lower stimuli). (B) At this recording site, responses to a matching colour (grey curve) are enhanced relative to responses to a nonmatching colour (black curve). (C) Population response pooled across 8 recording sites that had a significant response enhancement to the target colour. (D) In the simple task, the colour of the fixation point was irrelevant. (E) Population response of the same 8 recording sites during the simple task. Scale bar applies to the lower panels of c and e.

Eye position controls. The monkeys had to maintain fixation within a 1°x1° window during stimulus presentation. Nevertheless, small changes in eye position within the fixation window change the position of the relatively small V1 receptive fields on the curves, and may thereby influence response strength. Thus, systematic differences in fixation between conditions might contribute to the differences in response magnitude. This possibility was excluded in a stratification procedure that makes the eye position distributions identical for the conditions that are compared (18). Stratification did not influence the results, since latencies (124±3ms for the simple task, and 235±7ms for the composite task, after stratification) and significances (P<2·10^-6, paired t-test, for both tasks after stratification) of the response enhancements to the
target curve were virtually unchanged. Stratification also did not decrease the significance of the response enhancement to a matching colour of an irrelevant contour segment in the search phase of the composite task (Fig. 4) (P<0.0005, paired t-test, latency 148±10 ms, after stratification).

4. DISCUSSION

Our results go beyond existing neurophysiological data by showing how multiple attentional operators are coordinated in time. Figure 5 (right panels) shows best fitting curves to the average visual response (black), the response enhancement caused by curve tracing (white), and the response enhancement due to visual search (grey). In both the simple and the composite task, the first event in area VI is the visual response that is triggered at a latency of 39 ms. At a later point in time, the strength of the visual response is modulated by the task demands. If the task can be solved by just curve tracing, responses to the target curve are enhanced after 130 ms. However, in the composite task the response enhancement due to curve tracing is delayed by 100 ms, since a visual search that enhances neuronal responses to the target color at a latency of 159 ms is carried out before.

![Figure 5. Time-course of attentional processing in the two tasks. Left panels show the spatial distribution of the response enhancement, and right panels its time course, as determined by best fitting curves. In both tasks, the visual response has a latency of 39 ms (black area). In the simple task, responses to the target curve are enhanced at a latency of 130 ms (white). In the composite task, visual search first enhances responses to the target colour after 159 ms (grey). The location of this response enhancement can provide the starting position for the subsequent curve tracing operation, but this operation is delayed substantially.](image-url)
It is important to stress that these results do not indicate whether area V1 is actively involved in the solution of these tasks, or whether it passively reflects shifts of attention due to feedback from other areas. We have suggested previously how shifts of attention can be implemented by the propagation of rate enhancements within and between visual areas (21). The active participation of area V1 in this process may be beneficial, since V1 contains a high spatial resolution representation of the stimulus. During curve tracing, for example, perceptual grouping criteria such as collinearity and connectedness need to be evaluated, in order to identify all contour segments that belong to a single curve. Visual attention appears to spread across the target curve from attended segments to others that are related to them by these grouping criteria (17). At a physiological level of description, the responses of V1 neurons to all segments of the target curve are eventually enhanced (18,29). Neurons in area V1 with nearby receptive fields influence each other through horizontal connections (33,34). Neurons that are tuned to collinear orientations, and that therefore usually respond to segments of the same curve, have the strongest horizontal connections (35,36). Thus, these connections can propagate the response enhancement selectively, from neurons that respond to initial segments of a curve to neurons that respond to its successive segments. In area V1 receptive fields are small, and the response enhancement can therefore be propagated at a high spatial resolution. This may be important when the target curve comes close to other, distracting curves (21). A similar argument holds for visual search. If feature selective feedback enhances neuronal responses to a target item in area V1, its location is encoded with high precision, which may be important if the item is surrounded by nearby distractors.

The data summarized in Fig. 5 support a theory of Ullman (20), who suggested that complex visual tasks are solved by “visual routines” that are composed of a number of attentional operators, arranged in an orderly sequence. When a new stimulus is presented, feedforward connections rapidly (within 100 ms) propagate activity to the various visual areas. The features that are computed during this phase roughly correspond to those that can be detected pre-attentively in a variety of psychophysical tasks (1-5,21). However, vision is not done once this initial activation pattern across the visual areas has been established, since attentional operators are applied thereafter. We have previously suggested how attentional operators, such as visual search and curve tracing, can be implemented in the visual cortex by a modulation of the strength of neuronal responses (21). In Ullman’s theory, there is a limited library of these attentional operators, which loosely corresponds to the “instruction set” of the visual cortex. It is possible to compile a virtually unlimited set of routines from these operators, since they can be arranged to form many different sequences (20-22). The present data demonstrate that it is possible to monitor the progression of such a visual routine in the primary visual cortex.

The results also have implications for more general theories on performance in cognitive tasks, such as for example ACT-R (19) and Soar (28) that also decompose complex cognitive tasks into subtasks that are solved sequentially. These theories have been criticised for their assumption that the total reaction time of a subject equals the sum of the times occupied by the processes that are responsible for solving the subtasks (23), an assumption dating back to Donders (1868) (26). Opponents argue that subtasks can, in principle, also be solved in parallel, and that the total reaction time therefore need not be equal to the sum of times
required by the subprocesses (24,25). The present study is, to our knowledge, the first to directly measure the duration of the subprocesses in a composite task. The data show a temporal succession of visual search and curve tracing in the composite task, and thereby provide evidence against strictly parallel models. The logic of this task may seem to enforce such a succession of operators, since visual search has to supply the starting point of the curve tracing operator. We note, however, that search and tracing could in principle proceed in parallel, if the monkeys were able to trace both curves during visual search. In that case the end point of the target curve would be immediately evident upon completion of the search. The substantial increase in reaction times and the delayed enhancement of neuronal responses to the target curve in the composite task are incompatible with such a parallel model.

Importantly, the results are also not consistent with strictly serial models of subtask sequencing, since successive attentional operators exhibit considerable temporal overlap. The response enhancement caused by visual search, the first attentional operator in the composite task, decays only gradually during the later phase of a trial. We propose that the temporal overlap between successive attentional operators is important, since it allows the transfer of information from one operator to the next. In a computer program parameters are transferred from one subroutine to another by either copying them to another memory address (call by value), or by providing a pointer to the memory location where the relevant information can be found (call by reference). Persisting attentional rate modulations can fulfil an equivalent role in the visual cortex. This is illustrated by the composite task, where the first operator, visual search, results in a focus of response enhancement on a retinotopic map at locations where the target colour can be found (9-14) (green Fig. 5). This focus encodes the coordinates of the beginning of the target curve, and provides the starting position for the subsequent curve tracing operator. We call this "call by focus", since the curve tracing operator can retrieve the relevant information by being sensitive to the focus of the response enhancement on the retinotopic map. In different routines, other parameters such as motions, shapes or colours may need to be exchanged between operators (21,22). This can be implemented in an analogous fashion, by persisting rate enhancements in areas where neurons are tuned to the relevant features.

REFERENCES