Cognitive and neural mechanisms of visual search

Steven J Luck

University of California, San Diego, USA

The visual search paradigm is a useful tool for studying the perception of multiple-element stimulus arrays, and recent studies have identified a number of factors that influence search performance. Of special importance is the selection process, and recent neurophysiological experiments have shown that the suppression of information from irrelevant objects plays a crucial role in the selection of visual search targets.

Current Opinion in Neurobiology 1994, 4:183–188

Introduction

For reasons of simplicity, psychologists and neuroscientists have commonly studied visual processing by measuring responses to stimuli presented individually on a blank background. Much natural visual processing involves the perception of objects embedded in large, complex scenes, however, and the perception of such multiple-element stimulus arrays involves computational issues that are absent when stimuli are presented in isolation [1–3]. As a result, vision researchers have increasingly employed visual search tasks in which subjects view arrays containing multiple items and must search for a predefined target stimulus.

Much of the recent interest in visual search follows the publication of Treisman's influential 'feature integration theory' [4]. This theory postulates that vision occurs in two main stages: a pre-attentive stage in which the retinal image is decomposed into spatially organized maps of elementary features; and an attentive stage in which the features at individual spatial locations are conjoined into integrated object percepts. Without attention, the features of an object cannot be bound together and 'illusory conjunctions' may occur (e.g. a red X and a blue O may be perceived as a blue X and a red O). Because binding is by definition unnecessary for individual features, however, simple features may be detected without attention. This theory has been supported by numerous visual search experiments that measured target–detection latency as a function of the number of items in the search arrays (the 'set size'): for conjunction targets, reaction time typically increases linearly as a function of the set size (serial search), whereas reaction time is typically constant across set sizes for feature targets (parallel search).

Although feature integration theory has provided an important framework for studying visual search, recent psychophysical studies have questioned some of the claims of this theory and have shown that there are several additional factors that may influence visual search performance. This review will begin by discussing some of these behavioral studies and will then focus on recent studies of the neural substrates of visual search.

Statistical factors and grouping

When subjects must make several independent decisions on a single trial, the probability of an error increases as the number of decisions increases, simply because increasing the number of decisions leads to more opportunities for making an error [5]. As a result, target detection performance would be expected to vary with set size in visual search tasks for purely statistical reasons, even in the absence of perceptual limitations. This 'statistical model' was tested in a recent study using feature search tasks by Palmer and colleagues [6••], who found that increases in set size led to reliable decreases in accuracy, and showed that these effects could be entirely accounted for by the statistical model (see also [7•]). Consistent with the proposal that attention is not required for the detection of simple features [4], no evidence was found for limited perceptual resources [6••]. The finding of a statistical explanation for set-size effects raises questions, however, about many previous studies in which set-size effects were attributed to perceptual limitations, including the conjunction search experiments that have been used to support the feature integration theory. Additional studies employing complex visual search tasks are required to determine the extent to which set-size effects in such tasks result from limited perceptual resources, as was previously assumed, or from statistical factors.

The statistical model typically assumes that independent target/non-target decisions are made for each item in a search array, but this assumption may be violated in certain cases. In particular, if the items comprising a
search array can be segregated at an early level into a small number of groups, it may then be possible to apply decision processes to entire groups, thereby reducing the number of decisions required to detect a target and increasing search efficiency. This line of reasoning has been central to several recent models of visual search, and has been supported by a number of psychophysical experiments (8,9,10,11). Evidence for grouping effects at an early stage of processing has also been obtained in a recent electrophysiological study, which demonstrated that responses in primary visual cortex may depend on whether or not the stimulus within a neuron's receptive field matches stimuli presented simultaneously outside the receptive field (12).

**Nature of the searched representation**

The initial studies of Treisman and colleagues (4,13) demonstrated that parallel search was possible for simple feature targets, suggesting that high-level detection processes have access to low-level feature information (e.g. at the level of primary visual cortex). Several recent studies have indicated, however, that parallel search processes may operate on somewhat higher-level representations. For example, parallel search performance has been observed for targets that differ from the distractors in terms of their apparent source of illumination or their three-dimensional interpretation (14,15). Similarly, search performance in stereoscopic displays may vary according to the perceived depth and occlusion of the stimulus items (16). In particular, changes in perceived occlusion created by very subtle feature changes can cause a dramatic shift from fast, parallel detection performance to slow, serial performance. Results such as these suggest that visual search operates at a level of representation that encodes relatively complex surface properties rather than primitive features.

**Models of binding**

When multiple objects are presented simultaneously, neurons that code the constituent features of these objects are activated concurrently; veridical perception requires a mechanism for binding together the features that belong to each individual object in order to avoid miscombining the features of different objects. Perhaps the simplest way of solving this 'binding problem' is to select a single zone of space for analysis by suppressing information from all other locations (3). Such a mechanism would be strictly serial, however, and would work poorly for overlapping objects. A more flexible selection mechanism might label all of the neurons that correspond to a given object with a specific 'tag' and treat features with the same tag as parts of the same object (17,18). Such a tagging process would lead to correct binding while allowing some degree of parallel processing and separation of spatially overlapping objects. One candidate mechanism that has been proposed for tagging related features is temporally correlated firing among neurons that code for the same object, and several models using this mechanism have been proposed (17,19,20,21,22,23). The role of oscillations and correlations in the primate visual system is currently controversial (24-26), however, and other alternative models for attentional selection are being explored, such as dynamic information routing (27).

**Electrophysiological studies of visual search**

Many models of visual search propose that the visual system selects one object (or a group of related objects) at a time for identification, proceeding in an order determined by the similarity between each object and a target template stored in memory; when an object has been selected, its neural representation is presumed to remain active while the representations of other competing objects are suppressed (9,24,28,29,30). Neurophysiological support for this selection-and-filtering scheme has been provided recently by single-unit studies in monkeys and event-related potential (ERP) studies in humans.

In one study of inferotemporal neurons in macaque monkeys (31), an initial cue stimulus was followed by an array of 2-5 test items, and the monkeys were required to make a saccade to an item in the test array that matched the cue stimulus (see Fig. 1). The stimuli were chosen such that one stimulus was effective in driving the neuron being recorded and the others were ineffective; a neuron's response to a test array, therefore, reflected the processing of the one item that matched the stimulus preferences of the neuron. In the period following the cue stimulus, firing rates were found to be maintained at a higher level when the cue stimulus was an effective stimulus for the neuron compared to when the cue was an ineffective stimulus. This prolonged activation of neurons that code the features of the cue stimulus may reflect a target template that is used to select the target item from the distractor items in the test array.

When the test array was subsequently presented, the initial response was the same whether or not the one effective stimulus within the array matched the cue. Beginning approximately 175 ms after the test array appeared, however, the response became suppressed if the effective stimulus was a non-target, but was maintained at a high level if the effective stimulus was the target for the subsequent saccade. Thus, neurons in inferotemporal cortex exhibit both template-like activity following the cue stimulus and suppression of non-target information following the test array. Suppression effects have also been reported recently in extrastriate area V4 (L. Chelazzi, unpublished data; see also 32), in the frontal eye fields (33), and in the superior colliculus (34). In addition, a scalp-recorded field po-
tential (the 'N2pc' component), which exhibits properties quite similar to the single-unit attention effects observed in macaque inferotemporal cortex, has been observed in human subjects [35].

As described above, psychophysical and modeling studies have suggested that the primary role of selective attention in visual search may be to suppress information from irrelevant stimuli, which may be useful both for solving the binding problem [3] and for mitigating the statistical consequences of performing several concurrent decisions [6**,7*]. The congruent finding of suppressive attention effects in several macaque visual areas thus provides an important point of convergence among studies of attentional selection that have utilized a variety of methodologies.

Suppression effects have also been observed in an ERP study of visual search in humans, in which subjects searched for a uniquely colored target item embedded in an array of 15 distractor items [36*]. Task-irrelevant 'probe' squares were presented at the location of the target item or a distractor item, and the sensory ERP waves elicited by the probe were used as an index of perceptual processing at the probed location (see Fig. 2). When the target was present in the array but the probe was flashed at a distractor location, the initial probe-elicited sensory response (the 'P1' wave) was suppressed compared to trials on which the target was absent from the array and the same distractor location was probed ([36*]; see also [37,38]). No enhancement was observed for probes presented at the target location, however, indicating that this effect was primarily suppressive. Interestingly, this suppression effect was observed when subjects were required to report the shape of the uniquely colored target but was absent when a simple feature detection task was used [36*]. This finding is consistent with feature integration theory's proposal that the detection of simple features does not require the same attentional processes as more complex discriminations [4]. However, an enhancement of a different sensory response (the 'N1' wave) was observed at the target location in both the feature and conjunction tasks, indicating that some forms of attention are used even for the detection of simple features [36*].

Neuropsychological studies of visual search

Patients with right hemisphere lesions often suffer an acute period of neglect during which they ignore stimulation from the left side of space, and this may be followed by a less severe chronic state in which they show 'extinction' to stimuli on the compromised side when stimuli are also presented concurrently on the intact side. Several studies of these patients have recently shown that reaction times for targets in the compromised field increase dramatically as the number of stimuli in the intact field is increased, whereas increasing the number of items in the compromised field has a much smaller effect on the search rate [39,40]. This pat-
Fig. 2. Stimuli and results from a visual search study that utilized probe-evoked ERPs to assess sensory processing [35]. In this study, the search arrays consisted of 14 grey items and two colored items that were selected at random from a set of four colors; one of these four colors was the target for a given trial block. (a) In a conjunction discrimination task, subjects reported the shape of the target item, if present. (b) In a feature detection task, subjects simply reported the presence or absence of the target color. Sensory processing at the target and distractor locations was assessed by recording the ERP evoked by a task-irrelevant ‘probe’ square that was flashed either at the target location, at the location of a distractor item on the opposite side of the array from the target, or at the location of a distractor on target-absent trials. (c) For the conjunction discrimination task, the probe-evoked P1 wave (a positive peak at 100 ms post-stimulus) was found to be suppressed for probes presented at distractor locations (opposite target) on target-present trials compared to target-absent trials, but no enhancement was observed for probes presented at the target location. The N1 wave (a negative peak at 170 ms post-stimulus), in contrast, was larger for probes presented at the target location compared to target-absent trials, but no additional suppression was observed for probes presented on the opposite side of the array from the target. Thus, one sensory process exhibited facilitation at the target location, whereas another exhibited suppression at distractor locations. (d) The P1 suppression effect was eliminated in the feature detection task, but the N1 facilitation effect remained; this suggests that although the attentional requirements of feature detection and conjunction discrimination differ, some attentional processes are used for both tasks.

The cognitive processes underlying visual search performance have been studied extensively over the last several years, and continued study is leading to refinements in the understanding of these processes. The past year has also seen a substantial increase in studies of the neural substrates of visual search, with notable progress in the areas of neural network modeling, neurophysiology, and neuropsychology. In particular, these studies have provided converging evidence that the filtering of irrelevant information from distractor items plays a crucial role in visual search.

Conclusions
Acknowledgements

Supported by ONR contract N00014-89-J-1806 and by grants from
NIMH (MH-25594), NIH (NS-17778), and the Human Frontier Sci-
ence Program. Thanks to L. Chelazzi, C. Moore, and S. Hillyard for
helpful comments on the manuscript.

References and recommended reading

Papers of particular interest, published within the annual period of
review, have been highlighted as:
• of special interest
•• of outstanding interest
1. Feldman JA: Connectionist Models and Parallelism in High
Level Vision. In Human and Machine Vision, vol II. Edited by
2. Hinton GE, McClelland JL, Rumelhart DE: Distributed
Representations. In Parallel Distributed Processing, vol 1.
Edited by Rumelhart DE, McClelland JL. Cambridge, MA:
3. Mozer MC: The Perception of Multiple Objects. Cambridge,
4. Treisman AM, Gelade G: A Feature-Integration Theory of
5. Lappin JS, Uttal WR: Does Prior Knowledge Facilitate the
Detection of Visual Targets in Random Noise? Percep Psych-
6. Palmer J, Ames CT, Lindsey DT: Measuring the Effect of
Attention on Simple Visual Search. J Exp Psychol [Hum Percep]

This paper contrasts the role of statistical factors and perceptual lim-
itations in visual search, and finds that the effects of set size on ac-
curacy in a feature search task can be accounted for by statistical
factors alone.
7. Shiu L, Pashler H: Negligible Effect of Spatial Precuing on
• Identification of Single Digits. J Exp Psychol [Hum Percep],
in press.

Presents a number of spatial cueing experiments in which the ef-
ccts of attention on target detection performance can be explained
in terms of statistical factors rather than limitations in perceptual pro-
cessing.
8. Duncan J, Humphreys GW: Visual Search and Stimulus Sim-
• Search and Attentional Engagement. J Exp Psychol [Hum Percep]

Describes an alternative model of selective attention and visual
search that explains many of the results that have been used to sup-
sport feature integration theory.
10. Humphreys GW, Muller HJ: Search via Recursive Rejection
(SERR): a Connectionist Model of Visual Search. Cogn Psy-

Describes a neural network model that relies heavily on grouping
processes and accounts for a large number of visual search results.
12. Knierim JJ, Van Essen DC: Neuronal Responses to Static
• Texture Patterns in Area V1 of the Alert Macaque Monkey.

In this study, responses were measured from neurons in primary
visual cortex in response to multiple-element stimulus arrays. Con-
sistent with visual search experiments showing that unique features
are effortlessly detected, the responses of many primary visual cor-
tex neurons depended on whether the stimulus within the receptive
field matched stimuli presented outside the receptive field.
13. Treisman A, Souther J: Search Asymmetry: a Diagnostic for
Preattentive Processing of Separable Features. J Exp Psychol
14. AKS DJ, Enns JT: Visual Search for Direction of Shading is
• Influenced by Apparent Depth. Percept Psychophys 1992,
52:63-74.

Provides evidence that three-dimensional stimulus attributes, such as
direction of illumination, influence the rate of visual search.
15. Kleffner DA, Ramachandran VS: On the Perception of

Provides psychophysical evidence that three-dimensional shading plays
an important role in texture segregation and visual search.
16. He ZJ, Nakayama K: Surfaces Versus Features in Visual

This psychophysical study examined the role of occlusion and com-
petition in visual search, and found that small changes in simple fea-
tures that produced significant changes in perceived surface proper-
ties also produced substantial changes in search rate. The authors
concluded that search operates on a surface-level representation rather
than on a feature-level representation.
17. Mozer MC, Zemel RS, Behrmann M, Williams CK: Learning
• to Segment Images using Dynamic Feature Binding. Neural

Describes a neural network model that performs object identifica-
tion with a system that provides common 'tags' for the features that
comprise a given object.
18. Lumer ED, Huberman BA: Binding Hierarchies: a Basis for
Dynamic Perceptual Grouping. Neural Computation 1992,
4:341-355.
19. Crick F, Koch C: Towards a Neurobiological Theory of Con-
Systems-Level Proposal for the Neural Substrates of Recall
21. Hummel JE, Biederman I: Dynamic Binding in a Neural Net-

Describes a neural network model in which correlations among pro-
cessing units are used in object identification.
22. Von der Malsburg C, Schneider W: A Neural Cocktail-Party
23. Niebur E, Koch C, Rosin C: An Oscillation-Based Model
• for the Neuronal Basis of Attention. Vision Res 1993,
18:2789-2802.

Presents a neural network model of attention in which oscillations are
used to bind together the features that comprise individual ob-
jects.
24. Koch C: Computational Approaches to Cognition: the
25. Young MP, Tanaka K, Yamane S: On Oscillating Neuronal
Responses in the Visual Cortex of the Monkey. J Neurophy-
26. Toovey JM, Rolls ET: Oscillatory Activity is Not Evident in
the Primate Temporal Visual Cortex with Static Activity.
27. Olshausen BA, Anderson CH, Van Essen DC: A Neurophys-
iological Model of Visual Attention and Invariant Pattern
Recognition Based on Dynamic Routing of Information. J

Describes a biologically plausible neural network model of visual
processing that integrates spatial selective attention, translation in-
variance, and size scaling into a single scheme.
28. Labeige D: Computational and Anatomical Models of Se-
lective Attention in Object Identification. In The Cognitive
Neurosciences. Edited by Gazzea MN. Cambridge, MA: MIT
Press, in press.

A comprehensive review of the computational principles and pos-
sible neural substrates of attentional selection in visual search and
related paradigms.


Describes the activity of neurons in macaque inferior temporal cortex during a visual search task. The results indicate that these neurons are involved in maintaining an attentional template and in suppressing information arising from unattended objects.


Provides evidence for attentional modulation of single-unit activity in areas V1, V2, and V4 of macaque visual cortex, and shows that this modulation is heavily influenced by the presence of competing distractor stimuli.


Describes a visual search study in which responses were recorded from neurons in the frontal eye fields. These neurons exhibited a suppression of activity when non-target stimuli fell within their response fields and the target stimulus was adjacent to the response field.


Presents data from neurons in the superior colliculus that show activity that is predictive of the direction of a saccade toward a visual search target.


An electrophysiological study of visual search in human subjects using the ERP technique. An attentional suppression effect is reported at distractor locations when subjects identify conjunction targets but not when they simply detect feature targets; a facilitation effect is reported at the locations of both feature and conjunction targets.


This study examined visual search performance in neglect patients and found evidence for an attentional deficit that was manifested in object-centered coordinates.


Provides evidence that split-brain patients can scan the left and right hemispheres independently, and that both the left and right hemispheres contain the necessary neural circuitry for performing visual search.

SJ Luck, Department of Neurosciences 0608, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093-0608, USA.