Preemption Effects in Visual Search: Evidence for Low-Level Grouping

Ronald A. Rensink
Harvard University

James T. Enns
University of British Columbia

Experiments are presented showing that visual search for Mueller-Lyer (ML) stimuli is based on complete configurations, rather than component segments. Segments easily detected in isolation were difficult to detect when embedded in a configuration, indicating preemption by low-level groups. This pre-emption—which caused stimulus components to become inaccessible to rapid search—was an all-or-nothing effect, and so could serve as a powerful test of grouping. It is shown that these effects are unlikely to be due to blurring by simple spatial filters at early visual levels. It is proposed instead that they are due to more sophisticated processes that rapidly bind contour fragments into spatially-extended assemblies. These results support the view that rapid visual search cannot access the primitives formed at the earliest stages of visual processing; rather, it can access only higher-level, more ecologically-relevant structures.

The processes that underlie human vision are often divided into two fundamentally different classes: operations that are carried out in parallel over space, and operations that are not (e.g., Neisser, 1967; von Helmholtz, 1867/1962). For the most part, parallel processes are rapid (i.e., they occur within a few hundred milliseconds), effortless, and automatic (i.e., they cannot be affected by immediate changes in higher-level goals), whereas nonparallel processes are slower, more effortful, and nonautomatic. In its current embodiment, this dichotomy divides vision into an early preattentive and a subsequent attentive stage (e.g., Beck, 1982; Julesz, 1984; Treisman, 1988; Treisman, Cavanagh, Fischer, Ramachandran, & von der Heydt, 1990). The preattentive stage is thought to extract simple "features" in parallel across the visual field, while the attentive stage employs a variety of serial processes to assemble these features into coherent descriptions of the external world. Evidence for this view is based largely on studies of visual search, in which rapidly-detected items are believed to contain features that differ from those in their surroundings (see, e.g., Treisman, 1988).

Because of their simple nature, features are sometimes thought to include the set of "visual primitives", i.e., the properties directly obtained from the spatiotemporal filters at the earliest levels of processing. Strictly speaking, however, visual search data only provide evidence of structures that can be rapidly accessed to make a conscious report of target presence. These rapidly-accessed structures need not be visual primitives; indeed, some are complex scene-based properties obtained via rapid-interpretation processes (see, e.g., Enns & Rensink, 1990a, 1990b, 1991a). Conversely, there is also no *a priori* reason for visual primitives to be rapidly accessible.

Thus the following question arises: Are visual primitives in fact rapidly accessed? In this paper, we argue that they are not. To this end, we present three sets of experiments. The first (Experiments 1-3) shows that search for Mueller-Lyer (ML) stimuli is based on the complete configurations rather than on component line segments or junctions. It is shown that the components are *preempted* by the configurations into which they are grouped, becoming inaccessible to rapid search. The second set (Experiments 4-6) shows that this preemption cannot be accounted for by blurring at the earliest stages of vision—rather, the phenomenon must involve a more "active" kind of *low-level* grouping that takes place at levels below those easily accessible to search. The third set (Experiments 7-8) investigates the structures upon which this low-level grouping acts, and shows that it operates not only on luminance-defined lines, but also on virtual lines formed from dot arrays. It is proposed that—at least for ML stimuli—three specific processes are involved, each differing in its dependence on the orientation, contrast sign, and endpoint separation of the segments involved. A computational analysis is then sketched that accounts for the particular characteristics of these processes.

Primitive Elements of Vision

The notion that vision is based on a small set of primitive elements is an old one, and the quest for these primitives has extended over much of the history of experimental psychology (e.g., Bruner, Goodnow, & Austin, 1956; Hubel & Wiesel, 1962; Julesz, 1981; Mach, 1886/1959; Selfridge & Neisser, 1960; von Helmholtz, 1867/1962;). Almost all these

approaches assumed (although sometimes only implicitly) that primitives have four key characteristics: (i) they describe properties that are simple, i.e., are related in a direct way to the measurements made at the retina, (ii) they are formed by processes that act rapidly, typically within a few hundred milliseconds, (iii) they are formed by processes that operate in parallel across the visual field, the outputs creating a spatially-organized "map", and (iv) they are formed by processes that are automatic, i.e., that cannot be altered by sudden changes in higher-level goals.

The existence of such primitives has received considerable support from work in both physiology and psychophysics. Physiological studies show that the cortical areas most directly connected to the visual input (e.g., V1, V2, and MT) contain cells that respond preferentially to particular values of properties such as luminance, color, orientation, and motion (Hubel & Wiesel, 1962, 1977; Schiller, 1986; Zeki, 1992). These properties can be determined extremely quickly, with cell response latencies generally far less than 100 ms (see, e.g., Bishop, 1984). Also, the outputs of these cells often form retinotopic maps, each describing the distribution of the corresponding property in the visual field (see, e.g., Robson, 1980). For example, retinal ganglion cells are sensitive to local gradients in luminance and color; the maps at this "ganglion" level describe how these quantities are distributed over the image. Similar maps have been found at locations further along the visual stream, such as in thalamic LGN and cortical areas V1, V2 and V4 (Hubel & Wiesel, 1962, 1977; Zeki, 1992). In accord with their tentative identification as visual primitives, these properties are also determined by processes that are—apart from the possible inhibition of inputs and outputs (Haenny & Schiller, 1988; Moran & Desimone, 1985)—independent of higher-level control.

Psychophysical support for the notion of visual primitives rests largely on two different kinds of study. The first is concerned with the limits of spatial and color vision, investigating the thresholds for tasks such as hyperacuity, frequency discrimination and contrast discrimination. These studies show that much of threshold vision can be modelled by simple linear filters of various sizes and orientations, with sensitivity profiles similar to those of neurons at early processing levels (see, e.g., Graham, 1989, Kelly & Burbeck, 1984). These filters are believed to operate in parallel across the visual field, with characteristics that—apart from differences in scale—are much the same at all locations in foveal and perifoveal vision (Robson & Graham, 1981; Watson, 1987). Although higher-level expectations can facilitate performance, this appears to be due largely to the monitoring of selected filters (Yager, Kramer, Shaw, & Graham, 1984). Thus, if issues of processing time are disregarded, the outputs of these filters can be tentatively identified with the primitives found in the physiological studies.

A rather different type of study is often used to investigate issues of processing time. These studies focus on the deter-minants of effortless visual search and texture segmentation (e.g., Beck, 1982; Julesz, 1984; Treisman, 1988). In visual search, for example, the task is to determine as quickly as possible whether a given target item is present in a display among a set of nontargets (or "distractors"). If a target immediately "pops out" from the distractors (as, for example, a Q among O's), this is taken as evidence that the target has a

unique feature (in this case, a diagonal or vertical line). If search is more difficult, with detection time increasing in proportion to the number of items in the display (as for an O among Q's), the target is assumed to have a subset of the features contained in the distractors, with the absence of a feature not being signalled explicitly (Treisman & Souther, 1985). A continuum of search rates exists between the two types of search, with speed increasing as the targets and the distractors become more distinct from each other (Duncan & Humphreys, 1989; Treisman & Gormican, 1988).

These results suggest that the incoming image is analyzed rapidly (within a few hundred milliseconds) and automatically into a set of simple features—e.g., particular values of length, orientation—that are represented by set of "maps", each map describing the distribution of the associated feature across the visual field. The effortless "pop out" of a unique feature can be explained by the ease of finding the *presence* of activity somewhere in the relevant map; detecting the absence of a feature is more difficult because it involves small *differences* in the total activity, making it susceptible to noise (Treisman & Gormican, 1988). Another characteristic of such a process is that a unique *conjunction* of features is difficult to detect, since the spatial relations required for this are not available from the separated maps. Such relations are believed to be established by more effortful and time-consuming operations at a second level of processing that involves attention (e.g., Beck, 1982; Julesz, 1984; Treisman, 1988). The existence of separate "preattentive" and "attentive" levels has—naturally enough—prompted a tentative identification of the rapidly-accessible "features" with the "primitives" of the physiological and spatial-vision studies (e.g., Treisman, et al., 1990).

This identification can be maintained even in the face of recent findings that feature conjunctions *can* be rapidly detected when the individual features are sufficiently distinct from each other (Nakayama & Silverman, 1986; Treisman & Sato, 1990; Wolfe, Cave, & Franzel, 1989). In particular, the rapid detection of conjunctions can be explained by higher-level strategic control of the feature maps, possibly via an excitatory/inhibitory mechanism similar to that found at early cortical levels (Moran & Desimone, 1985). But although higher-level control can now be brought to bear on the search process, it only *selects* the particular feature maps to be considered—it does not control the *kinds* of properties they describe.

A more serious obstacle to identifying features with visual primitives stems from recent studies showing that texture perception and visual search can be based on properties such as lighting direction and surface slant (Enns & Rensink, 1990a, 1990b, 1991a; Kleffner & Ramachandran, 1992; Ramachandran, 1988). These properties are considerably more complex than simple image measurements: they are scene-based, obtainable only via image interpretation. Indeed, rapid search can even be based on part-whole associations (Wolfe, 1992a). Consequently, there exist rapidly-accessible structures that are not simple visual primitives.

In what follows, we examine the converse issue: Can the rapidly-constructed visual primitives always be rapidly accessed? Note that it is not enough that they can *sometimes* be rapidly accessed, for an isolated structure may be left unchanged under certain conditions, and so effectively act as

a primitive. Rather, the criterion is that primitives must *always* be accessible—they can never be *pre*empted by more complex structures they form (cf. Liberman & Mattingly, 1989). If such *pre*emption exists, it would indicate that search cannot rapidly "reach down" to the lowest levels of visual processing. This in turn would imply that the features of visual search do not correspond to visual primitives, but to structures formed at some higher level of processing.
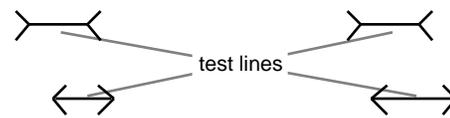
Segment Versus Overall Length

In order to determine whether simple visual primitives are always accessible to the mechanisms underlying rapid search, we begin by focusing on one particular property—line length. This property is obviously image-based, and can support both rapid visual search (e.g., Treisman & Gormican, 1988) and texture segmentation (e.g., Beck, 1982; Julesz, 1984).

It is possible to measure line length via simple spatial filters. Indeed, several models of texture segmentation use exactly such a mechanism to explain the sensitivity of texture perception to the shape of texture elements (e.g., Bergen & Adelson, 1988; Gurnsey & Browse, 1989; Malik & Perona, 1990). But measuring length in this way is only suitable for isolated straight-line segments—when lines contain a curve or bifurcation, something more sophisticated is needed. More generally, length is an extensive quantity, meaning that the measure assigned to the whole is the sum of the measures assigned to its parts (see, e.g., Carnap, 1966). Since these parts generally extend across space, length cannot be determined from an arbitrarily small neighborhood about an individual point (as can be done for quantities such as color). Instead, information must be collected from and assigned to an extended assembly of segments that have been grouped together on some basis or other.

The length estimate assigned to a line segment embedded in a configuration can therefore take on at least two possible values: (i) a *segment* length (the physical length of the isolated segment), or (ii) an *overall* length (the extent of the entire configuration). When a segment is isolated in the image, these values are equal; when it is embedded in a configuration, they may differ substantially. The issue here is to determine which (if any) of these estimates governs visual search[1]. If speeds depend on segment length, it indicates that visual search accesses only the component segments, with assemblies being formed at some later stage. If speed depends on overall length it indicates that the segments have been preempted by the assemblies before they could be accessed by the search mechanisms.

To determine which of these situations holds, we carried out a series of visual search experiments in which the items were ML configurations. These stimuli (see Figure 1), can introduce configural effects that are noticeable even with casual viewing (Coren & Girgus, 1978). In particular, the test line of the "wings-out" configuration in Figure 1a appears longer than that of the "wings-in" configuration, even though the two test lines have the same physical length. In Figure 1b, the "wings-in" test line is physically longer than the wings-out line, but the equal overall extents now cause the test lines to appear more equal in length. Given the strength of its effects at "attentive" levels, the

A. Different overall length   B. Different segment length

Figure 1. The classical Mueller-Lyer configuration. A: Test lines are equal in physical length, but the upper line appears longer. B: The upper test line is 50% shorter than the lower test line but appears to be more similar in length.

ML configuration may well induce configural effects at early levels. Note, however, that the effects at the two levels need not be the same.

Consider now a visual search experiment in which targets and distractors are ML configurations differing only in wing arrangement. Three types of test condition (shown in Figure 2) can be defined:

   a) *different-overall*: Target test lines are the same as the distractor test lines. Segment lengths are therefore the same; overall lengths are different.
   b) *different-segment*: Target test lines are 2/3 the length of distractor test lines. Overall lengths are the same; segment lengths are different.
   c) *different-combined*: Target test lines are 1/2 the length of distractor test lines. Both overall and segment lengths are different.

The influence of configural effects can be evaluated via the relative difficulty of search as a function of the central test line length, as shown in Figure 3. Because search speeds increase as targets and distractors become more distinct (Treisman & Gormican, 1988), and because the arrangement of the lines is constant across all conditions, the targets detected most quickly can be taken to differ most in the particular type of length accessed by search. Depending on whether grouping takes place, and whether preemption occurs, at least three different response patterns are possible:

   a) *assembly-based*: Search is fastest for stimuli with different overall length (i.e., different-overall and different-combined tests). This indicates that search is based entirely on assemblies of line segments, so that both grouping and preemption have occurred.
   b) *segment-based*: Search is fastest for stimuli with test lines of different physical length (i.e., different-segment and different-combined tests). This indicates that search is based only on component segments, so that neither grouping nor preemption have occurred.
   c) *junction-based*: Search is equally fast for all conditions. This could occur, for example, if search were based on the type of wing junction. It could also

---

[1] Other measures of length are also possible. For example, it is pos-sible to define an *apparent length* as the inferred length of the segment that has been "recovered" from the assembly. If the surrounding lines in the assembly introduce distortion into the recovery process, apparent length could differ from physical length. In such a case, the meaning of "apparent length" would correspond to the way it is used in studies of attentively-viewed length illusions. The experiments here vary only segment length and overall length, since these are easiest to determine.
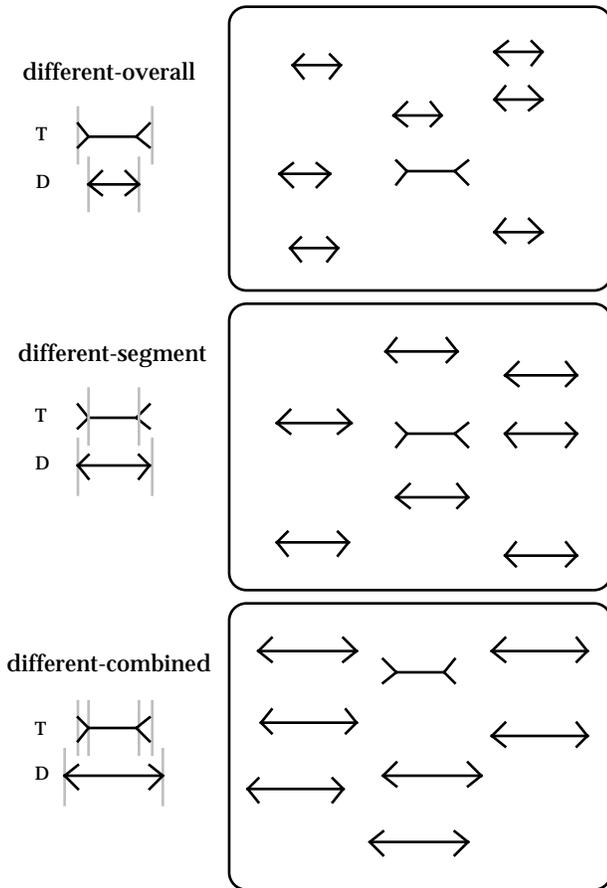
Figure 2. Examples of visual search displays involving Mueller-Lyer stimuli. The target item (T) is distinguished from the distractor item (D) by wing type. Different-overall test: T differs from D in overall length of the configuration. Different-segment test: T differs from D in the physical length of the test line (33% shorter). Different-combined test: T differs from D in both overall length of the configuration and physical length of the test line (50% shorter).
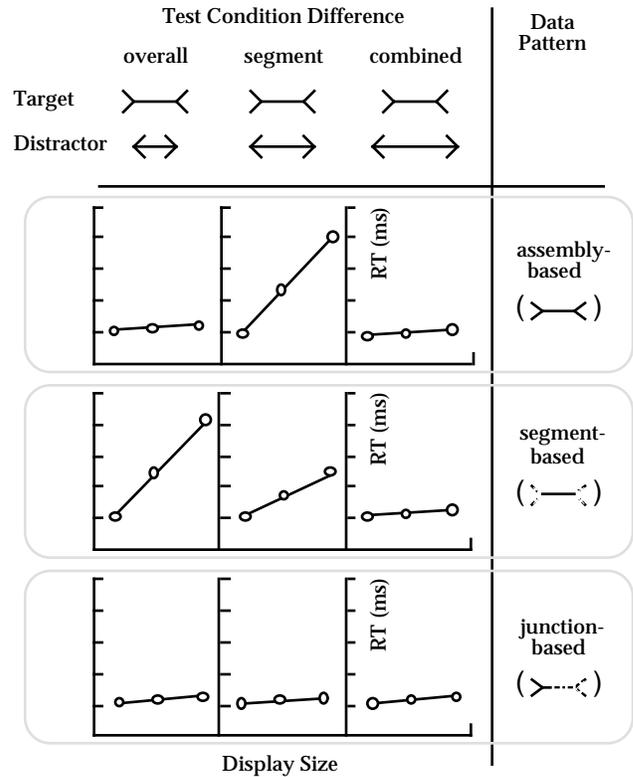

Figure 3. Three possible patterns of search speeds. Assembly-based: search is fastest for stimuli with test lines of different overall length (i.e., different-overall and different-combined conditions); this indicates that search is entirely based on the entire group, and so components have been preempted. Segment-based: search is fastest for stimuli with test-line segments of different length (i.e., different-segment and different-combined conditions); this indicates that search is based only on the components, and so grouping has not occurred. Junction-based: search is rapid for all conditions; this indicates that search is based on junction type, or on both assembly- and component-based properties.

occur if search were based on both assemblies and components (i.e., if grouping has occurred, but not preemption).

General Method

The methods used to test visual search were similar to others in the literature (e.g., Enns & Rensink, 1991a; Treisman & Gormican, 1988; Wolfe, Cave, & Franzel, 1989), with observers searching for a single target among a total of 2, 8 or 14 items. The target was present on half the trials, which were randomly distributed over the presentation sequence. A Macintosh computer generated the displays, controlled the experiments, and collected the data (Enns, Ochs, & Rensink, 1990). Ten observers participated in each condition of a given experiment. Each observer completed 3 sets of 60 test trials in each condition, following 20-30 practice trials that were not recorded.

Observers

All observers had normal or corrected-to-normal vision. Half had no previous experience in visual search or response time tasks, and half had extensive experience in these tasks. Because there was partial overlap in the subject sample between experiments (i.e., the experienced observers), all statistical analyses between experiments treated the observers conservatively as independent observations.

Stimuli and Procedure

In each set of experiments, target and distractor items were composed of line segments differing in their spatial arrangement, and sometimes also in the length of their central test lines (see Figures 4-12). In most conditions, one of the ML configurations (usually "wings-out") was the target item, while the other (usually "wings-in") was the distractor. Wings were always composed of lines oriented 45° away from the test line.

On each trial, items were distributed randomly on an imaginary 5 x 5 grid subtending 15.5° horizontally and 7° vertically; this distribution was constrained to keep item density the same for all displays. Each item was placed in a virtual cell measuring 3.1° x 1.4° and randomly shifted in position; to prevent contact with adjacent items it could not

come within 0.1° of the edge of the grid. Each trial began with a fixation symbol presented for 500 ms, followed by the display, which remained visible until the observer responded or 3 s had elapsed. The display was followed by a central feedback symbol (plus or minus sign), which served as the fixation point for the next trial. Target presence or absence was reported by pressing one of two response keys. Observers were instructed to maintain fixation as much as possible and to keep errors below 10% in each block of trials.

Data Analyses

Observers were quite accurate overall, with each making fewer than 10% errors on average. However, there were systematic differences in accuracy. Consistent with other reports, target-present trials led to more errors than target-absent trials (e.g., Enns & Rensink, 1991a; Klein & Farrell, 1989). Most important for present purposes, however, was the observation that errors tended to increase with reaction time (RT), which indicated that observers were not simply trading accuracy for speed.

Only the RT data from correct responses were used and these were analyzed in the same way for each experiment. First, simple regression lines were fit to the target-present and target-absent data for each observer. The estimated slopes of these lines were submitted to analyses of variance in which test condition (different-overall, different-segment, different-combined) and trial type (present, absent) were the effects of interest. Fisher's *LSD* tests (protected *t*-tests) were used to determine the reliability of slope differences. The reported *t*-tests, therefore, are tests of differences in RT slope, based on the pooled error variance and degrees of freedom from the overall analysis.

Experiments 1-3: Preemption of stimulus components

The first set of experiments examined whether the components of the ML stimuli are preempted by assemblies formed at levels below those rapidly accessible to visual search. If search speed is influenced by segment length, it indicates that search can rapidly access the individual components of the stimuli. But if search speed is influenced only by overall length, it indicates that the components are preempted, with search based entirely on assemblies.

Experiment 1

Experiment 1 comprised six conditions, shown in Figure 4. The first three (A-C) were different-overall, different-segment and different-combined tests, with the target a wings-out configuration having a test line of length 1.4°. The lengths of the distractor test lines were 1.4° (different-overall), 2.1° (different-segment), and 2.8° (different-combined). Complementary to these direct conditions, the other three (D-F) were reversed, with target and distractor items switched. Depending on the type of length measure rapidly accessible to search, at least three different sets of results could be expected.

First, if assembly-based estimates alone were accessible, an assembly-based pattern would be expected, with search fastest for the different-overall and different-combined tests (see Figure 3). Since long items are generally easier to find
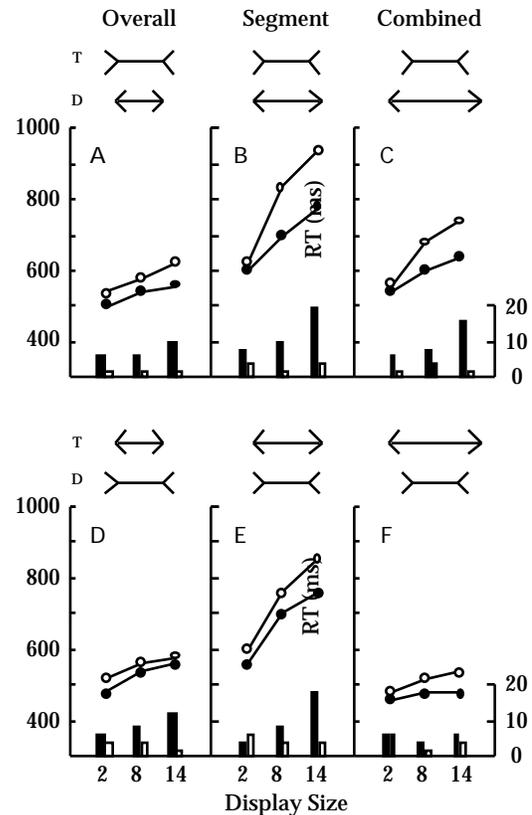

Figure 4. The stimuli used in Experiment 1, along with the mean RT and percentage error data. Conditions A-C tested the same target item (T) against distractor items (D) that varied in length (direct tests). Conditions D-F tested different-length target items against the same distractor items (reversed tests). Black symbols indicate target-present trials; white symbols target-absent trials.

among short than vice versa (Treisman & Gormican, 1988), search should also be faster for the direct than for the reversed different-overall test. This asymmetry should disappear for the different-segment tests, where targets and distractors have similar overall lengths, and be inverted for the different-combined tests, where the direct condition has a target shorter than the distractor.

Second, if segment-based estimates alone were accessible, a segment-based pattern would be expected, with search being faster for the different-segment test and fastest of all for the different-combined test. The asymmetry for length would not be found for the different-overall tests, since both direct and reverse conditions have the same test line lengths. It would, however, be found in the different-segment and different-combined tests, with the direct conditions being slower than their reversed counterparts.

Finally, if both segment- and assembly-based estimates were accessible, or if search could be based on the line junctions forming the wings, a junction-based pattern would be expected, with search equally fast in all test conditions.

The mean RTs from correct responses and percentage errors for the six conditions are shown in Figure 4. As can be seen from the graphs, search speeds followed a clear assembly-based pattern, showing that search is dominated by assembly-based estimates. Indeed, results confirm all predictions based on the preemption of component segments.

First, search was fast for both different-overall tests. In the direct case (Condition A), mean RT slopes were 5.2 ms/item for target-present trials and 7.1 ms/item for target-absent trials. In the reversed case (Condition D), the corresponding mean RT slopes were 7.8 and 5.5 ms/item. Also, the small difference observed between target-present trials in the direct and reversed different-overall tests (A vs. D), although not significant, [$t(18) = 1.04$], was consistent with the expected search asymmetry for length.

Second, search was far slower for both different-segment tests. In the direct case (Condition B), mean RT slopes weref 15.6 and 28.5 ms/item; this was significantly slower than for the different-overall test (Condition A) [$t(18) = 4.14$, $p < .01$ for target-present; $t(18) = 8.53$, $p < .01$ for target-absent]. In the reversed case (Condition E), mean RT slopes were 16.4 and 23.2 ms/item; this was significantly slower than the different-overall test (D) [$t(18) = 3.43$, $p < .01$ for target-present; $t(18) = 7.05$, $p < .01$ for target-absent]. Search speeds were much the same for the direct and reversed different-segment tests (B vs. E) [all $ps > .05$]. This lack of asymmetry was consistent with the targets and distractors having approximately the same length estimates—otherwise, length asymmetry (Treisman & Gormican, 1988) would result in different search speeds for the two conditions.

Third, search was again faster for both different-combined tests. In the direct case (Condition C), mean RT slopes were 10.2 and 12.9 ms/item; this was faster than the different-segment test (B) [$t(18) = 2.15$, $p < .05$ for target-present; $t(18) = 2.31$, $p < .05$ for target-absent]. In the reversed case (Condition F), mean RT slopes of 2.5 and 4.7 ms/item; this again was significantly faster than the different-segment test (E) [$t(18) = 5.54$, $p < .01$ for target-present; $t(18) = 7.37$, $p < .01$ for target-absent]. The inverted search asymmetry in the different-combined tests (C vs. F) was significant [$t(18) = 3.07$, $p < .01$ for target-present; $t(18) = 3.27$, $p < .01$ for target-absent].

Experiment 2

The results of Experiment 1 show that visual search is dominated by assemblies rather than segments, and thus indicate that grouping does occur. They do not, however, prove that estimates based on the components of the assemblies are completely inaccessible, for these may simply have been insufficiently distinct to support rapid search. Experiment 2 therefore examined several control conditions to determine the role played by these components.

In the first condition (A), items were isolated test lines. The first stimulus set (A1) contained a long target and short distractor, corresponding to the isolated test lines of the different-segment stimuli of Experiment 1E (i.e., Condition E of Experiment 1). The results (see Figure 5 and Table 1) were striking: search was far faster for the lines when isolated (mean RT slopes of 6.0 ms/item for target-present trials and 9.6 ms/item for target-absent trials) than when they were embedded in the ML configurations (mean RT slopes of 16.4 and 23.2 ms/item; $t(18) = 5.75$, $p < .01$ for target-present; $t(18) = 4.95$, $p < .01$ for target-absent).

Condition	Stimuli (1)	RT Slope Present	Absent	Stimuli (2)	RT Slope Present	Absent
A	T / D	6.0	9.6	T / D	10.6	20.0
B	T / D	7.3	10.3	T / D	12.8	18.8
C	T / D	11.9	20.1	T / D	4.6	5.3

Figure 5. The stimuli and mean RT slopes (in ms/item) in Experiment 2. T = target item; D= distractor item.

The second stimulus set (A2) switched the choice of target and distractor items. In accord with the search asymmetry for line length, this yielded somewhat slower search (10.6 and 20.0 ms/item; $t(18) = 3.54$, $p < .01$ for target-present; $t(18) = 8.01$, $p < .01$ for target-absent). Nonetheless, search for these isolated lines was still faster than when they were embedded in the corresponding ML configurations (Experiment 1B), although this was significant only on target-present trials (15.6 and 28.5 ms/item; $t(18) = 2.26$, $p < .05$ for target-present; $t(18) = 1.52$ for target-absent).

The next two conditions examined the influence of the wing junctions. Junctions are not generally considered to be capable of supporting rapid search, since they involve only a spatial arrangement of the lines (Beck, 1982; Julesz, 1984; Treisman, 1988). But recent work has shown that at least some kinds of line arrangements *can* be rapidly detected (Duncan & Humphreys, 1989; Elder & Zucker, 1992; Enns & Rensink, 1991a). Because the junction types in the targets and in the distractors differed from each other and were held constant throughout Experiment 1, observers might have used this cue to speed up search.

Condition B examined the individual junctions of Experiment 1. To provide a robust comparison between single- and double-junction search, the stimuli were the left halves of the targets and the right halves of the distractors. These junctions are minimally distinct from one another, differing only in the relative position of the test line (see Figure 5). Thus, if observers in Experiment 1 used junction differences to aid search, this particular choice would be the "worst case": if search was relatively fast for these stimuli, it would be even faster for a more discriminable pair of junctions.

In the first stimulus set (B1), all stems were 0.7 ° long, so that the items were halves of the different-overall stimuli of Experiment 1. The results (see Figure 5 and Table 1) show that search for these isolated junctions was relatively fast (mean RT slopes of 7.3 and 10.3 ms/item). This is much the same speed as for the corresponding "complete" configurations of Experiment 1A and 1D [both $ps > .05$]. The second stimulus set (B2) increased the length of the

Table 1
Summary of Data in Experiment 2

Condition	Target	Mean reaction time (ms) Display Size 2	8	14	Mean % Correct Display Size 2	8	14
A1	Present	532	566	620	91	92	82
	Absent	562	633	685	97	93	94
A2	Present	556	659	708	95	97	93
	Absent	636	776	861	96	95	97
B1	Present	496	560	570	95	94	82
	Absent	512	582	629	97	99	97
B2	Present	581	661	708	99	89	82
	Absent	602	717	843	97	98	99
C1	Present	499	591	643	91	92	93
	Absent	538	664	779	98	95	90
C2	Present	474	505	512	98	93	90
	Absent	472	554	593	95	93	96

distractor stem so that items had the same overall length (1.1^O); they were now halves of the different-segment stimuli of Experiment 1. As seen from Figure 5 and Table 1, search slowed down [mean RT slopes of 12.8 and 18.8 ms/item; $t(18) = 4.24$, $p < .01$ for target-present; $t(18) = 6.55$, $p < .01$ for target-absent], although it still tended to be faster than for the "complete" configurations of Experiments 1B and 1E [$ps > .05$].

To determine the effect of the particular test-line lengths chosen, the next condition (C) used stimuli with shorter test lines. The first stimulus set (C1) was the same as that of B1, but with test lines only half as long. These show no evidence for junction type to dominate as the length estimates are reduced—instead of speeding up, search slowed down [mean RT slopes of 11.9 and 20.1 ms/item; $t(18) = 3.55$, $p < .01$ for target-present; $t(18) = 7.55$, $p < .01$ for target-absent]. This is consistent with search based entirely on the overall length of the junctions, which become more similar as the test lines become shorter (Duncan & Humphreys, 1989; Treisman & Gormican, 1988). If junction types are properties that facilitate search, their contribution here is virtually nonexistent. These junctions, however, still showed a trend toward faster search than that for the "complete" configurations of Experiments 1B [$t(18) = 2.05$, $p < .05$ (1-tailed) for target-present; $t(18) = 1.47$ for target-absent] and 1E [$t(18) = 2.36$, $p < .05$ for target-present; $t(18) = 0.89$ for target-absent].

The next stimulus set (C2) used the stimuli of C1, but with the distractor test lines adjusted so that all junctions had the same overall extent. In contrast to the results of Conditions B1/C1, speeds increased as the test lines became shorter [mean RT rates of 4.6 and 5.3 ms/item; $t(18) = 5.63$, $p < .01$ for target-present; $t(18) = 11.41$, $p < .01$ for target-absent]. This might be explained by a lessened influence of the length estimates, but in light of the earlier results it is more likely due to a difference in pixel intensity (i.e., the number of pixels within the bounds of the item differ for target and distractor): as the test lines become shorter, this difference increases, thereby speeding up search (Duncan & Humphreys, 1989; Treisman & Gormican, 1988).

Taken together, Conditions A-C show that when highly distinctive components, such test lines or wings, are combined into ML configurations, the resulting structures are more difficult to detect than the components alone. Most importantly, search for individual line segments was never slower than for the configurations in which they were embedded, and indeed, could even be considerably faster.

Experiment 3

The combining of easily detected components into structures that are difficult to detect can be attributed to one of two things. It could be due to *preemption*, which would make the components effectively invisible to rapid search. But it might also be due simply to the *complexity* of the stimuli, the variation of wings and test lines creating sufficient "noise" to slow down search (Duncan & Humphreys, 1989). To test this latter possibility, two more conditions were examined (Figure 6). Here, both wings-in and wings-out configurations were used for the target (a random choice being made for any target-present trial) and the distractors (equal proportions present in each display).

In the first condition (consistent overall), overall length was held constant while segment length varied. Targets had test line lengths of 1.4^O and 2.1^O, while the distractors had test line lengths of 0.9^O and 1.4^O. Thus, each target and distractor set contained items of dissimilar segment length but similar overall length (Figure 6). The difference in overall lengths between targets and distractors was set to that of Experiments 1A and 1D. If overall length alone
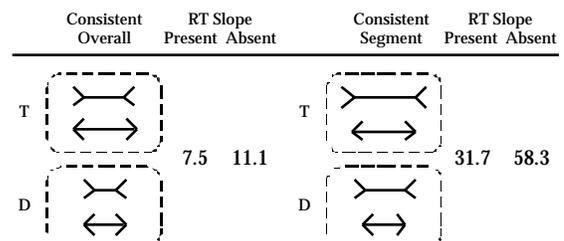

Figure 6. The stimuli and mean RT slopes (in ms/item) in Experiment 3. T = target item; D= distractor item.

governed search, rates in these two conditions would be similar. But if wing type or test line segments were also being used, the interference caused by variations in (and overlapping distributions of) target and distractor components would cause search to slow down considerably (Duncan & Humphreys, 1989; Treisman & Gormican, 1988).

Mean RTs and percentage errors are given in Table 2 and mean RT slopes in Figure 6. Contrary to the predictions of the "interference" hypothesis, the variations in wing types and test line lengths had no large effect: search remained easy, with mean RT slopes of 7.5 ms/item for target-present trials and 11.1 ms/item for target-absent trials. These speeds were not significantly different from the different-overall rates of Experiment 1 [for both $t(18) < 1$]. Interestingly, a comparison of these rates with those of Experiment 2A shows that a 3:2 difference in overall length has an influence similar to a 3:2 difference in segment length [both $ps > .05$], supporting the idea that only one type of length estimate is involved here.

The second condition (consistent-segment) examined the converse situation, in which segment lengths were held constant while overall length was varied. The test lines of the targets were 2.1^O long, and the test lines of the distractors 1.4^O long. If overall length were the only quantity that can be accessed, search would be far slower than for the different-overall tests of Experiment 1, since this quantity no longer reliably discriminates targets from distractors. But if segment length could be accessed, this effect would not be expected (and might even be reversed), since observers could use the lengths of the test lines as a consistent cue.

The results again show a dependence on overall length alone: mean RT slopes were 31.7 and 58.3 ms/item, far above those of the first condition [$t(9) = 4.88$, $p < .01$ for target-present; $t(9) = 9.52$, $p < .01$ for target-absent]. Taken together, the results of Experiments 1-3 show that—at least for ML stimuli—individual components are preempted by the assemblies they form, and that this preemption cannot be overridden, even to facilitate the task. This in turn implies the existence of a low-level grouping of segments into items. Note that this *intra-item* grouping is to be distinguished—at least in principle—from the *inter-item* grouping of items into spatially-extended regions (e.g., Duncan & Humphreys, 1989; Grossberg, Mingolla, & Ross, 1993).

Experiments 4-6: Existence of Active Grouping

The results of Experiments 1-3 demonstrate a clear preemption effect: visual search can rapidly access estimates based on complete line-segment assemblies, but cannot access more primitive estimates based on individual segments. As such, these results show that some form of intra-item grouping takes place at levels below those easily accessible to search. They do not, however, point to the mechanisms responsible.

In order to investigate the nature of this low-level grouping, we examined how search speed is affected by changes in stimulus properties. If speeds change in a relatively gradual way, this would indicate a "passive" grouping due to linear mechanisms such as simple spatial filters. But if speeds remain largely stable, with sudden transitions caused by small changes in particular properties, this would indicate the existence of more "active" grouping due to more sophisticated processes.

Experiment 4

One explanation of the grouping found in Experiments 1-3 is that it might be due to spatial filtering of the kind believed to exist at the earliest stages of visual processing (e.g., Ginsburg, 1986; Watt, 1988; Wilson & Gelb, 1984). These filters can remove various spatial frequencies to yield blurred versions of the original image; models based on such filters can explain many aspects of texture perception (e.g., Beck, Sutter, & Ivry, 1987; Bergen & Adelson, 1988; Gurnsey & Browse, 1989; Malik & Perona, 1990; Voorhees & Poggio, 1988). Although differing in details, these models all postulate a two-stage process that begins with an array of small-scale (high-resolution) filters operating directly on the image. The outputs of this stage are divided into two subsystems—one for positive changes in contrast (light edges) and one for negative changes (dark edges). The absolute values of these outputs are sent to a second stage of large-scale (low-resolution) filters; the lines of maximum change in these second-stage outputs are then taken to be the texture boundaries.

Thus, it might be that low-level assemblies are simply "blobs" formed by blurring the input at the earliest stage of visual processing. To test this hypothesis, Experiment 4 used the stimuli of Experiment 1, but with gaps of various sizes placed in the central test lines. To control for the number of pixels in the items, gaps were the same size in both the different-overall and the different-segment stimuli. If the effects are due to distortions induced by large-scale filters, neither the existence of small gaps nor changes in their location (i.e., in the center or at the end of the test line) should greatly affect search rates. If the effects are due to small-scale filters, both the size of the gaps and their locations should have a large impact.

Table 2.
Summary of Data in Experiment 3

Test	Target	Mean reaction time (ms)			Mean % Correct		
		Display Size			Display Size		
		2	8	14	2	8	14
Consistent-Overall	Present	535	591	626	95	96	92
	Absent	566	648	699	98	98	96
Consistent-Segment	Present	865	1138	1243	86	79	76
	Absent	1039	1450	1735	88	89	84

Condition		Different Overall	RT Slope Present	Absent	Different Segment	RT Slope Present	Absent
(1A,B)	T / D		5.2	7.1		15.6	28.5
A	T / D		3.1	5.0		13.2	26.5
B	T / D		3.1	9.7		10.8	29.0
C	T / D		4.4	7.9		11.6	24.2
D	T / D		19.5	29.5		29.6	57.0
E	T / D		16.2	31.5		25.4	53.1
F	T / D		16.7	28.9		24.4	55.1
G	T / D		14.0	28.5		25.4	46.3

Figure 7. The stimuli and mean RT slopes (in ms/item) in Experiment 4. For purposes of comparison, an upper row has been added to show the RT slopes of Conditions A and B of Experiment 1. T = target item; D = distractor item.

The stimuli are shown in Figure 7. To maintain compatibility with Experiment 1, the distance between the wing corners was set to 1.4^O for targets and different-overall distractors, and to 2.1^O for different-segment distractors. Since different-combined tests are not needed to determine whether grouping occurs, they were omitted from subsequent experiments.

In the first three conditions (A-C), gaps were placed in the center of the test lines, with sizes of 0.35^O, 0.70^O, and 1.05^O respectively. Search rates were measured as in Experiment 1. Table 3 provides a detailed summary of the results, with mean RT slopes given in Figure 7.

Inspection of these slopes reveals a dramatic pattern: speed was unaffected by the presence of a central gap, no matter what its size (Conditions A-C). Rates for the different test-line fragments did not differ significantly from one another [all ps > .05], a trend not true of the individual junctions (Experiment 2), indicating that the junctions were still being preempted. Speeds in both the different-overall and different-segment tests did not differ significantly from those for their contiguous counterparts in Experiment 1 [all ps > .05]. The influence of overall length was also maintained: search remained significantly slower in all corresponding different-segment tests [$t(54)$ = 2.64, p < .05 for A target-present; $t(54)$ = 5.63, p < .01 for A target-absent; $t(54)$ = 2.02, p < .05 for B target-present; $t(54)$ = 5.05, p < .01 for B target-absent; $t(54)$ = 1.89, p < .05 (1-tailed) for C target-present; $t(54)$ = 4.27, p < .01 for C target-

absent]. The existence of this region of relatively stable search speed suggests that—at least at far as length estimation is concerned—the assemblies governing search in these conditions were the same, and were also similar to those that governed search in Experiment 1.

In the next condition (D), the test lines were removed completely. A rather striking slowdown now occurred: mean RT slopes were 19.5 and 29.5 ms/item for different-overall test, and 29.6 and 57.0 ms/item for different-segment test (all ps <.01). Observers could no longer rapidly detect the target, even though only two short line segments had been removed from the stimuli of Condition C. Evidently, the removal of the small segments caused the loss of a highly salient difference between target and distractor.

A dependence on overall length, however, remained: different-overall rates were significantly faster than the different-segment rates [$t(54)$ = 2.64, p < .01 for target-present; $t(54)$ = 7.19, p < .01 for target-absent]. Search was apparently still based on some property derived from the group of lines in each item.

In order to investigate this further, the final three conditions (E-G) had gaps placed at the ends of the test lines. To allow comparison with the results of Conditions A-C, these gaps were split equally among the two locations; sizes were therefore 0.17^O, 0.35^O, and 0.52^O respectively. As for conditions A-C, gaps were the same size in both different-overall and different-segment tests.

Results showed the same indifference to gap size as found in Conditions A-C. However, search here was much slower, being similar to that for Condition D. Indeed, the results for Conditions E-G all followed the same pattern. Mean different-overall rates were not significantly different in the three conditions [all ps > .05] and were also close to those of Condition D [all ps > .05]. For all three conditions, different-segment rates were significantly slower than the corresponding different-overall rates [range of $t(54)$ = 2.02, p < .05 to $t(54)$ = 6.86, p < .01], and were similar to the different-segment rates of Condition D [range of $t(54)$ = 0.52 to $t(54)$ = 2.82, p < .01].

In effect, the central test lines of Conditions E-G had very little influence on search speed, no matter what their length. Since the different-segment test lines are sufficiently distinct to support rapid search on their own (Experiment 2), it appears that they are still preempted. Thus, although a salient property has been lost, both signatures of grouping—the dependence on overall length and the preemption of segments—are still present.

Two different kinds of group will therefore be distinguished: *assemblies* and *bundles*. Search based on assemblies is reflected in an assembly-based pattern, with speeds that are relatively high (i.e., similar to those in Experiments 4A-C). Search based on bundles follows a separate *bundle-based* pattern that is similar but slower by a factor of roughly four for the different-overall stimuli and roughly two for the different-segment stimuli (i.e., rates are similar to those of Conditions 4D-G). The relatively slow speed of bundle-based search suggests that bundles are somehow less distinct; given that bundles are formed when no central test line is present, the loss of the salient assembly-based feature is likely to be related to a loss of interior structure.

Table 3
Summary of Data in Experiment 4

Condition	Test	Target	Mean reaction time (ms) Display Size			Mean % Correct Display Size		
			2	8	14	2	8	14
A	Overall	Present	510	517	547	92	94	94
		Absent	545	581	603	96	95	95
	Segment	Present	626	701	783	91	93	85
		Absent	637	905	955	94	96	97
B	Overall	Present	509	519	542	97	98	97
		Absent	518	575	626	97	99	99
	Segment	Present	606	666	736	97	97	95
		Absent	653	806	1001	98	98	98
C	Overall	Present	508	549	561	99	96	97
		Absent	536	591	632	96	99	99
	Segment	Present	599	671	736	96	96	96
		Absent	616	771	906	97	99	99
D	Overall	Present	544	646	776	98	91	89
		Absent	556	716	912	98	99	97
	Segment	Present	714	963	1070	94	85	73
		Absent	658	1089	1340	97	98	94
E	Overall	Present	602	710	796	98	96	93
		Absent	636	815	1015	98	98	99
	Segment	Present	658	830	963	97	95	88
		Absent	687	991	1324	98	98	98
F	Overall	Present	628	733	829	97	94	87
		Absent	651	810	999	98	98	97
	Segment	Present	633	783	925	95	93	85
		Absent	641	954	1301	98	98	97
G	Overall	Present	564	649	721	97	97	97
		Absent	583	747	942	99	96	97
	Segment	Present	634	781	938	95	89	84
		Absent	657	915	1213	97	99	95

The existence of these two search-speed patterns presents a problem for any account based on filtering. Filtering could account for part of these effects, in that search is always slower when the filtered stimuli have the same overall size. But if the effects are mostly due to small-scale filters, why should the preemption of the individual junctions (Conditions A-C) be indifferent to the width of the gap separating them? And if the effects are mostly due to large-scale filters, why should the small end gaps of Condition E not fill in and lead to search speeds like those for their contiguous counterparts? And why should the removal of small test line fragments (Conditions C-D) cause such a sudden transition in search speed?

Figure 8 shows the effect of filtering the center-gap stimuli of Condition C and the end-gap stimuli of Condition E with the circularly-symmetric Gaussian filters used in many models of texture perception. A range of filter sizes is shown, corresponding to various degrees of blurring. As seen from the figure, there is no point at which a blurred version of the center-gap stimuli has a greater overall extent or more cohesiveness than the end-gap stimuli. If anything, the filtered center-gap stimuli resemble the wings-alone stimuli of Condition D, while the filtered end-gap stimuli resemble the contiguous stimuli of Experiment 1. These similarities in shape are a complete reversal of the similarities in search rates. Thus, low-level grouping is

unlikely to be due to the simple spatial filtering that occurs at the earliest levels of visual processing. Rather, it would appear to be carried out by more active, sophisticated processes at somewhat higher levels.

Experiment 5

The preemption of individual segments by low-level groups does not imply that the segments are completely impossible to access—indeed, they are clearly visible in any casual viewing of the figures. At some point, therefore, low-level preemption can be overridden by higher-level control. But how might this be done?

A natural place to begin is with the influence of selective attention. Selective attention can have a considerable influence on early vision (Driver, McLeod, & Dienes, 1992; Treisman, 1988; Treisman & Sato, 1990; Wolfe, Franzel, & Cave, 1989). According to Treisman & Sato (1990), attention can guide search by inhibiting outputs from all locations that show activity in a particular feature map; search can then be restricted to the remaining uninhibited items. Wolfe et al. (1989) have proposed a similar model based on excitation rather than inhibition.

If low-level grouping involves structures (e.g., lines, junctions) bound to each other only weakly, observers should be able to extract these structures in order to facilitate
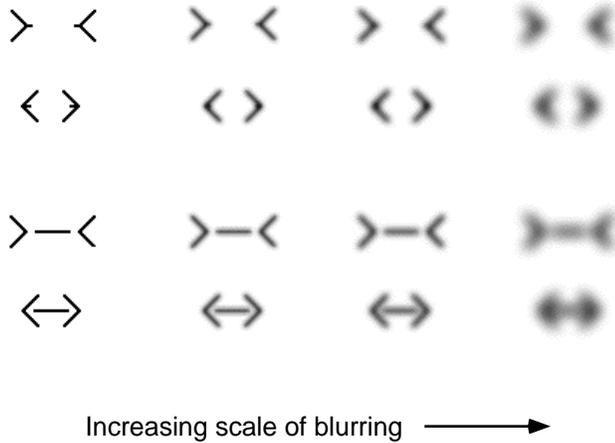
Increasing scale of blurring ⟶

Figure 8. Effect of filtering the center-gap stimuli of Condition C and the end-gap stimuli of Condition E (Experiment 4) with circularly-symmetric Gaussian filters of increasing spatial scale. Filtering the center-gap stimulus results in a shape similar to that of the wings alone; search rates, however, are similar to those for the complete Mueller-Lyer (ML) configuration. Conversely, filtering the end-gap stimulus results in a shape similar to the complete ML configuration; search rates are now similar to those for the wings alone.

Condition		Different Overall	RT Slope Present	RT Slope Absent	Different Segment	RT Slope Present	RT Slope Absent
A	T / D		5.2	8.9		32.1	44.5
B	T / D		23.1	43.3		12.6	24.1
C	T / D		4.0	6.4		13.8	30.6
D	T / D		6.2	11.9		4.9	7.3
E	T / D		14.8	18.6		6.6	5.7

Figure 9. The stimuli and mean RT slopes (in ms/item) in Experiment 5. T = target item; D= distractor item.

search. The preceding experiments have shown that this cannot be done via line orientation, for otherwise the wings would not have slowed search in the different-segment tests. But selection might be done via a nongeometrical property, such as the sign of contrast. This possibility is, of course, difficult to verify completely: if a structure can be rapidly accessed, this could be due to its extraction from its group, or it could be that the structures were never grouped together to begin with. Either way, however, it is of interest to determine how grouping is affected by contrast sign.

The first condition (A) examined whether search could be guided by selective attention to individual lines of the ML configurations. Stimuli (Figure 9A) had black test lines and white wings, and were placed on a gray background. If selective attention could be used, observers in the different-segment test could ignore the white wings or focus on the black test lines. Target and distractor would then be lines differing by a ratio of 3:2 in length, and search would be rapid (cf. Experiments 2 and 3). But if selective attention cannot break the groups, speeds should be much the same as for stimuli of uniform contrast (Experiment 1).

The mean RTs and percentage error data for Condition A are given in Table 4, with mean RT slopes shown in Figure 9. Search was clearly assembly-based: search for the different-overall test was fast (mean RT slopes of 5.2 and 8.9 ms/item), while search for the different-segment test was much slower [mean RT slopes of 32.1 and 44.5 ms/item, $t(36) = 8.15$, $p < .01$ for target-present; $t(36) = 10.79$, $p < .01$ for target-absent]. Although this latter speed was slower than that for the different-segment tests of Experiment 1, this may simply have been due to the noise introduced by the two different contrasts, which would act to slow down search in the absence of distinct differences in length (Duncan & Humphreys, 1989). Furthermore, the pre-

emption of the test lines again occurred, in spite of the different contrast signs of the wings. Thus, the grouping of line segments attached to each other was not weakened by selection based on sign of contrast.

The next condition (B) examined whether the unit of selection might be an entire configuration rather than an individual line. As shown in Figure 9B, stimuli were formed by adding to each ML configuration a second set of wings that have opposite contrast sign and point in the opposite direction (Coren & Porac, 1983; Tsal, 1984). If observers could separate the configuration from the added wings, the search task would be facilitated, at least for the different-overall tests. However, the results (Table 4 and Figure 9) showed no sign of this: rates for the different-overall test were slow (mean RT slopes of 23.1 and 43.3 ms/item) and rates for the different-segment test were significantly faster [mean RT slopes of 12.6 and 24.1 ms/item; $t(36) = 3.18$, $p < .01$, for target-present; $t(36) = 5.82$, $p < .01$ for target-absent]. Evidently, the black and white segments had been bound into unitary assemblies.

Although attentive control cannot select structures that are physically attached to each other at their endpoints, it might be able to select structures that—although grouped together—are physically separated. This was tested in the next condition (C), where the ML stimulus was split into two junctions by placing a small (0.40°) gap in the center of the test line. Stimuli were therefore similar to those of Experiment 4A, but with opposite contrasts for the two halves (Figure 9C). As the data in Table 4 and Figure 9 show, a strong assembly-based pattern was present, with different-overall rates of 4.0 and 6.4 ms/item, and different-segment rates of 13.8 and 30.6 ms/item [$t(36) = 2.97$, $p < .01$ for target-present; $t(36) = 7.33$, $p < .01$ for target-absent].

Table 4.
Summary of Data in Experiment 5

Condition	Test	Target	Mean reaction time (ms)			Mean % Correct		
			Display Size			Display Size		
			2	8	14	2	8	14
A	Overall	Present	501	526	555	97	94	87
		Absent	540	608	657	98	99	97
	Segment	Present	597	800	983	96	87	74
		Absent	630	950	1156	97	99	97
B	Overall	Present	652	831	929	97	89	82
		Absent	705	1012	1215	97	99	98
	Segment	Present	578	663	731	99	93	85
		Absent	663	869	959	99	98	98
C	Overall	Present	499	523	546	94	96	96
		Absent	541	601	619	95	97	98
	Segment	Present	573	652	738	90	91	92
		Absent	613	848	980	95	98	95
D	Overall	Present	499	524	572	88	90	91
		Absent	516	599	658	98	99	97
	Segment	Present	505	541	563	90	93	87
		Absent	502	564	590	98	97	97
E	Overall	Present	509	608	686	96	94	90
		Absent	527	666	748	98	99	99
	Segment	Present	498	545	576	96	92	89
		Absent	522	567	594	94	97	97

To determine whether the indifference to gap size found in Experiment 4 was maintained, the next condition (D) used stimuli similar to those of Condition C, but with a larger gap (0.80°). The results (Table 4 and Figure 9D) show a striking change—a junction-based pattern was now found, with different-overall rates of 6.2 and 11.9 ms/item, and different-segment rates of 4.9 and 7.3 ms/item [all ps < .05]. Thus, the grouping of the junction pairs has failed or else is sufficiently weak that it can be overridden by selective attention. Either way, the individual junctions could now be accessed in a way that was not possible for the same-contrast stimuli of Experiments 4A-C.

To determine the sensitivity of bundles to contrast sign, the next condition (E) used stimuli like those of Experiment 4E—a test line separated from the wings by end gaps of 0.20°—but with a black test line and white wings. The results (Table 4 and Figure 9E) show a clear segment-based search pattern. Search for the different-overall test remained relatively slow (14.8 and 18.6 ms/item); the target-present rates were much the same as for Experiment 4E (16.2 ms/item, p > .05), although the target-absent rates there were significantly slower (31.5 ms/item, $t(18) = 2.62$, $p < .05$). Search for the unequal segments, on the other hand was considerably faster [6.6 and 5.7 ms/item; $t(36) = 2.48$, p < .05 for target-present; $t(36) = 3.91$, $p < .01$ for target-absent].

Both of these effects are consistent with a failure to form groups—either assemblies or bundles—from the separated line segments. Changing the contrast sign does not greatly affect search for the different-overall stimuli, but the extraction of the target and distractor test lines would not help here, since they have the same physical lengths. However, the change of contrast in the different-segment

condition now allows rapid access to the distinctly shorter test line of the target, which in turn allows search to speed up. The isolated components were apparently not pre-empted in the way they had been in Experiments 4E-G.

In summary, then, the effects of contrast sign depend critically on the type of group involved, and on whether the line segments are attached to each other. In Condition A, for example, search for completely-connected stimuli slowed down considerably for the different-segment conditions, whereas in Condition E, similar—but disconnected—segments led to a speed-up of search. More generally, low-level groups are sensitive to contrast sign when the endpoints of the component segments are sufficiently separated, but not when these segments are attached to each other.

This sensitivity to endpoint attachment and to endpoint separation indicates that rapid access to the components (at least for assemblies) is unlikely to result from selective attention, for it is difficult to imagine why selection based on contrast sign should depend upon such geometric factors for its success or failure. Rather, it would appear that the rapid access to components found in Condition D simply result from a failure to form low-level assemblies.

Experiment 6

Given that low-level grouping can be broken by differences in contrast sign (at least when the components are separated), can it also be broken by differences in geometric structure? To find out, we used stimuli similar to the center-gap stimuli of Experiment 4, but with the test-line segments altered so that the interior contour could no longer

be completed by simple straight-line interpolation (Figure 10). Each condition not only had different-overall and different-segment tests, but also had two sub-conditions: one with gaps of width 0.40° (small gap) and one with gaps of width 0.80° (large gap).

In the first condition (A), the two halves of each item were displaced vertically with respect to each other by 0.17°; the direction of displacement (i.e., left side up or down) was assigned at random. The small-gap condition (A1) was examined first. As Figure 10 shows (details in Table 5), the vertical displacement did not affect search: it was still fast for the different-overall test (2.7 and 12.4 ms/item) and significantly slower for the different-segment test [12.0 and 30.2 ms/item; $t(18) = 4.62$, p < .01 for target-present, $t(18) = 8.86$, p < .01 for target-absent]. This clearly followed an assembly-based pattern, with speeds not significantly different from those for the "undisplaced" center-gap stimuli in Experiment 4A [both ps > .05].

In contrast, the corresponding large-gap condition (A2) yielded a junction-based pattern: search was fast in both the different-overall test (4.2 and 10.5 ms/item) and the different-segment test (4.9 and 7.3 ms/item); the difference between these two speeds was not significant [both ps > .05]. This sensitivity to gap size parallels the pattern found in Experiments 5C/D, showing that grouping has broken down in the wide-gap conditions, with search now able to rapidly access the individual junctions.

The next condition (B) examined gaps surrounded by test-line segments of different orientations. The horizontal widths of the gaps were again either 0.40° or 0.80°; segments were either +20° or -20° from horizontal. These orientations were randomly assigned, except that the lines on opposite sides of a gap had opposite orientation signs (see

Figure 10). Note that in contrast to Condition A, these segments could be extended to contact each other—a situation similar to the "relatability" condition of Kellman and Shipley (1991, 1992). However, since the restrictions on relatability (that the lines must correspond to one-sided surface edges, and that their extensions must meet at an oblique angle) need not apply here, we will refer to such lines in a more neutral way as "intersectable".

Search for the small-gap condition (B1) was again assembly-based, with search fast for the different-overall stimuli (2.0 and 8.5 ms/item) and slower for the different-segment stimuli [13.1 and 31.9 ms/item; $t(18) = 5.52$, p < .01 for target-present, $t(18) = 11.64$, p < .01 for target-absent]. Thus, line orientation appeared to have little effect when small gaps were used—search rates were not significantly different from those for the center-gap stimuli of Experiment 4A [both ps > .05].

Search for the large-gap condition (B2), however, differed considerably from that for its "non-intersectable" counterpart (A2): an assembly-based pattern was now found. Search was fast for the different-overall test (3.8 and 9.9 ms/item), and slow for the different-segment test [16.7 and 30.3 ms/item; $t(18) = 6.42$, p < .01 for target-present, $t(18) = 10.14$, p < .01 for target-absent]. This distance-indifference effect was similar to that found in Experiment 4: small- and large-gap rates did not differ significantly either for the different-overall tests [both ps > .05] or for the different-segment tests [both ps > .05]. Thus, the formation of assemblies appears to have taken place.

The next condition (C) used the same stimuli as Condition B, but with the segments on each side of a gap having the *same* orientation, so that they were no longer intersectable (see Figure 10). For the small-gap condition (C1), speeds followed the assembly-based pattern found in Conditions A1 and B1: search was fast for the different-overall test (3.8 and 9.6 ms/item) and slower for the different-segment test [14.3 and 26.0 ms/item; $t(18) = 5.22$, p < .01 for target-present, $t(18) = 8.16$, p < .01 for target-absent]. There were no significant differences from the rates for Condition B1 of this experiment or from those of Experiment 4A [all ps > .05], indicating that line orientation had no effect on speeds for small-gap stimuli.

In contrast, the results for the large-gap stimuli in Condition C2 showed a return to the junction-based pattern. Search was fast for both the different-overall test (5.6 and 9.4 ms/item) and different-segment test [7.1 and 9.4 ms/item; both ps > .05]. This sensitivity to distance indicated that grouping failed for this condition, with search based on individual junctions. Thus, the mere presence of oriented line segments did not cause the grouping found in Condition B. Rather, grouping depended on the relative orientation of the line segments, maintaining its indifference to distance only when these segments could be extended to contact each other.

In summary, then, assembly-based search patterns are found when the segments surrounding the gap can be extended to contact each other, or when the endpoints of the junction stems are close to each other; if the junctions cannot be related to each other in this way, a junction-based search pattern is found, indicating that they are no longer being grouped together. These results are difficult to explain via

Condition	Different Overall	RT Slope Present	Absent	Different Segment	RT Slope Present	Absent
A1		2.7	12.4		12.0	30.2
A2		4.2	10.5		4.9	7.3
B1		2.0	8.5		13.1	31.9
B2		3.8	9.9		16.7	30.3
C1		3.8	9.6		14.3	26.0
C2		5.6	9.4		7.1	9.4

Figure 10. The stimuli and mean RT slopes (in ms/item) in Experiment 6. T = target item; D = distractor item.

Table 5
Summary of Data in Experiment 6

			Mean reaction time (ms)			Mean % Correct		
			Display Size			Display Size		
Condition	Test	Target	2	8	14	2	8	14
A1	Overall	Present	548	568	576	98	98	96
		Absent	589	682	737	98	98	99
	Segment	Present	541	633	688	96	98	98
		Absent	591	749	952	97	99	99
A2	Overall	Present	561	603	612	97	97	96
		Absent	568	690	697	99	98	99
	Segment	Present	511	532	568	94	95	89
		Absent	526	607	615	96	98	95
B1	Overall	Present	543	548	549	98	97	99
		Absent	577	636	663	99	99	99
	Segment	Present	618	723	774	97	96	91
		Absent	680	914	1063	97	99	99
B2	Overall	Present	497	520	545	99	97	94
		Absent	529	594	648	99	99	98
	Segment	Present	493	565	691	96	95	93
		Absent	544	827	907	95	99	98
C1	Overall	Present	527	528	574	97	98	95
		Absent	530	601	645	99	99	99
	Segment	Present	541	613	714	96	97	95
		Absent	618	851	933	98	99	99
C2	Overall	Present	483	508	549	99	98	94
		Absent	505	582	618	98	99	99
	Segment	Present	454	483	537	94	95	87
		Absent	476	554	588	96	98	98

spatial filtering, for how could blurring induce a sensitivity to the orientation of widely-spaced segments? Rather, the pattern of results echoes those found in the earlier experiments—grouping occurs in a largely "all or nothing" way, indicating the involvement of mechanisms sharply sensitive to particular aspects of structure.

Experiments 7-8: Inputs to low-level grouping

The first set of experiments presented in this paper (Experiments 1-3) showed that visual search cannot access visual primitives, the individual components being pre-empted by the groups that they form. The second set (Experiments 4-6) showed that this effect is not due to simple filtering at the level of primitive measurements, but rather, is due to a more sophisticated form of grouping which has a sharp sensitivity to the contrast sign, relative orientation, and endpoint separation of adjacent line segments. Thus, bounds can be set to the level at which this grouping occurs: a level below those easily accessible to search, but above those involved with the measurement of image properties.

Can this level be located more precisely? In what follows, we attempt to do so by determining the kinds of inputs used by the grouping process.

Experiment 7

Structures that are somewhat more abstract than simple luminance-based lines are the "virtual lines" formed from arrays of dots or endpoints[2]. While equivalent in many ways to luminance-defined "physical" lines, these structures have no physical basis in the image (see, e.g., Petry & Meyer, 1987; Zucker, 1987a). They are generally believed to result from interactions among receptive fields at early stages of vision. Cells responsive to such lines have been found in cortical area V2 (von der Heydt & Peterhans, 1989; Peterhans & von der Heydt, 1989, 1991), and it has also been shown that rapid visual search can be based on their shape (Gurnsey, Humphrey, & Kapitan, 1992).

---

[2] The term *virtual line* is used here in preference to *subjective contour* (or *illusory contour*) for two reasons. First, there is no assumption that the result itself is consciously experienced as a structure—it is only detected here by its effects on grouping. Second, the result need only be equivalent to a physical *line* rather than a contour, since it is not necessarily treated as anything more complex at this stage.

Table 6.
Summary of Data in Experiment 7

			Mean reaction time (ms)			Mean % Correct		
			Display Size			Display Size		
Condition	Test	Target	2	8	14	2	8	14
A	Overall	Present	496	544	576	98	98	97
		Absent	535	604	629	99	98	98
	Segment	Present	681	930	1037	96	88	78
		Absent	724	1169	1425	99	96	95
B	Overall	Present	554	631	654	99	91	86
		Absent	599	705	735	99	98	98
	Segment	Present	615	871	1050	99	91	76
		Absent	725	1185	1414	99	98	97
C	Overall	Present	673	819	934	98	95	89
		Absent	712	1023	1257	99	98	96
	Segment	Present	689	932	1126	98	91	82
		Absent	767	1293	1632	99	98	96
D	Overall	Present	493	533	550	94	94	89
		Absent	528	566	593	97	98	97
	Segment	Present	594	692	778	92	90	81
		Absent	622	840	981	96	98	96
E	Overall	Present	510	536	555	97	96	95
		Absent	559	603	663	97	98	97
	Segment	Present	614	709	838	97	96	87
		Absent	651	891	1135	97	99	99

To determine whether low-level grouping can operate on such lines, the first condition (A) replaced the lines of the stimuli of Experiment 1 by dot arrays in which dots were separated by 0.35$^\mathrm{O}$. As the data in Table 6 and Figure 11 show, an assembly-based pattern was found: search was fast for the different-overall test (6.0 and 7.0 ms/item) and considerably slower for the different-segment test (29.7 and 60.0 ms/item). These differences were all significant, $t(36)$ = 5.35, $p < .01$ for target-present; $t(36) = 11.96$, $p < .01$ for target-absent. Thus, it appears that low-level grouping can be based on virtual lines.

To determine how these virtual lines depended on contrast sign, a second condition (B) used dots of alternating contrast on a gray background. Interestingly, speeds were largely unaffected (Figure 11)—the search pattern was similar to that of condition A, with different-overall rates of 8.2 and 11.3 ms/item, and different-segment rates of 36.2 and 57.4 ms/item [$t(36) = 6.32$, $p < .01$ for target-present; $t(36) = 10.40$, $p < .01$ for target-absent].

To determine whether the input had been split into channels of opposite contrast prior to grouping, a third condition (C) used only the subsets of black dots in the previous stimuli. If the effects resulted from groups formed separately in each channel, it should have remained strong when these subsets were used. But search (see Figure 11) slowed down considerably, with different-overall rates of 21.0 and 45.5 ms/item, and different-segment rates of 37.2 and 72.3 ms/item. The slowdown in different-overall rates was significant [$t(36) = 2.89$, $p < .01$ for target-present; $t(36)$ = 8.76, $p < .01$ for target-absent], indicating that the removal of dots caused a much more severe disruption than did the

alternation of dot contrast. Either grouping can be based on virtual lines formed of opposite-contrast dots, or else there is considerable sophistication in the way the black- and white-defined virtual lines are combined.

Note that search for the different-segment tests in Conditions A and B was slower than for the corresponding tests in Experiments 1 and 4, indicating that grouping might have failed, perhaps because of the virtual lines. But this slowdown only occurred in the absence of strong differences in the overall lengths of target and distractor (cf. the alternating-contrast stimuli of Experiment 5A). Thus, the effect was likely due to the noise introduced by the dots themselves, which—in the absence of strong differences between target and distractor—would act to slow down search (Duncan & Humphreys, 1989).

As a test of this hypothesis, the next condition (D) replaced the dots in the wings by lines and used smaller dots spaced apart by 0.28$^\mathrm{O}$. If dot-induced noise is the cause of slowdown, search speeds should approach those found for the assemblies in Experiment 4; if virtual lines are somehow the cause, search speeds should stay much the same. The results (Table 6 and Figure 11) support the hypothesis of dot-induced noise: Mean rates for the different-overall tests did not differ reliably from the center-gap tests in Experiment 4 [all ps > .05], nor did the mean rates for the different-segment tests differ from the corresponding center-gap tests [all ps > .05].

The assembly-based search found for Condition D contrasts oddly against the much slower bundle-based search found for Experiment 4G (see Figure 7), where the stimuli had an extremely short line (essentially a dot) between the

Condition		Different Overall	RT Slope Present	RT Slope Absent	Different Segment	RT Slope Present	RT Slope Absent
A	T D		6.0	7.0		29.7	60.0
B	T D		8.2	11.3		36.2	57.4
C	T D		21.0	45.5		37.2	72.3
D	T D		4.7	5.8			
E	T D		3.9	8.5		18.5	40.6
(4G)	T D		14.0	28.5		25.4	46.3

Figure 11. The stimuli and mean RT slopes (in ms/item) in Experiment 7. For purposes of comparison, a lower row has been added to show the RT slopes of Condition G of Experiment 4. T = target item; D = distractor item.

wings. Apparently, there is a critical dependence on the number of dots between the wings. To explore this further, the next condition (E) used stimuli in which the four dots between the wings were replaced by two dots, again separated by 0.28°. Results (Table 6 and Figure 11) show an assembly-based pattern, with fast search for the different-overall test (3.9 and 8.5 ms/item), and slower search for the different-segment test [18.5 and 40.6 ms/item; $t(36) = 3.29$, $p < .01$ for target-present, $t(36) = 7.24$, $p < .01$ for target-absent]. Thus, the replacement of four dots by two did not have any significant effect [all $ps > .05$, except for target-absent in the different-segment test, $t(36) = 2.28$, $p < .05$], whereas the replacement of two dots by one caused a dramatic slowdown [$t(18) = 6.35$, $p < .01$ for target-present, $t(18) = 4.41$, $p < .01$ for target-absent]. Evidently, the virtual lines used in low-level grouping can be formed from only two dots.

The large difference in search speeds between the one- and two-dot cases again reinforces the idea that low-level grouping is not due to spatial filtering, but is due to more sophisticated mechanisms operating on relatively complex representations. In particular, search for the dot-array stimuli followed assembly-based patterns similar to those found in Experiment 1. This indicates not only that virtual lines can be used as the basis of grouping, and that these lines are preempted as readily as their luminance-defined counterparts. It also indicates that virtual lines act much like physical lines without explicit terminations. However, given the reluctance to group lines of opposite contrast sign when separated (Experiment 5), grouping does not appear to

operate on the more abstract "media-indifferent" structures that can also support rapid search (Cavanagh, Arguin, & Treisman, 1990).

Experiment 8

Input structures can be characterized not only by their *abstractness*, but also by their *complexity*. It is therefore of interest to ask whether low-level grouping depends exclusively on the properties of uniform line segments (virtual or real), or whether it is also sensitive to sub-structures in the input. Experiment 8 examined this issue from two different points of view.

The first two conditions (A and B) examined how grouping is affected when a geometrically more complex structure is placed between the wings. Here, the intervening structure was a square composed of four line segments (each of length 0.35°) attached at their endpoints (see Figure 12A). This structure differed from a simple test line in that it was a complex of lines, and contained no free line endings. In the first condition (A), the square had no connections to the wings; it effectively replaced the test line in the bundle-forming stimuli. If the square is treated as an isolated line, Condition A should lead to a bundle-based pattern similar to that of Experiment 4D. If the square has a different status, a different pattern might be expected. The data (see Figure 12 and Table 7) show that replacing the line segment by the square had very little effect: mean different-overall rates were 10.2 and 20.0 ms/item, and mean different-segment rates were 18.8 and 40.1 ms/item, showing a bundle-based pattern [$t(27) = 3.84$, p < .01 for target-present; $t(27) = 8.98$, $p < .01$ for target-absent]. Apparently, grouping is not affected by the placement of a more complex structure between the wings.

In the second condition (B), the square was connected to the wings by lines, with T-junctions formed at their point of contact (see Figure 12B). These T-junctions could potentially separate the square from the rest of the figure, as occurs in several other kinds of early visual processing (e.g., Enns & Rensink 1991a, 1992; He & Nakayama, 1992). If

Condition		Different Overall	RT Slope Present	RT Slope Absent	Different Segment	RT Slope Present	RT Slope Absent
A	T D		10.2	20.0		18.8	40.1
B	T D		2.5	8.5		11.1	25.2
C	T D		12.8	25.9		18.4	42.2
D	T D		5.9	11.7		11.3	28.3

Figure 12. The stimuli and mean RT slopes (in ms/item) in Experiment 8. T = target item; D = distractor item.

Table 7
Summary of Data in Experiment 8

			Mean reaction time (ms)			Mean % Correct		
			Display Size			Display Size		
Condition	Test	Target	2	8	14	2	8	14
A	Overall	Present	516	576	639	98	98	95
		Absent	515	612	756	98	98	99
	Segment	Present	538	649	774	99	95	88
		Absent	585	796	1052	98	99	98
B	Overall	Present	472	497	500	98	97	97
		Absent	491	550	606	97	98	98
	Segment	Present	536	608	675	98	97	95
		Absent	570	718	874	98	99	99
C	Overall	Present	535	600	688	98	96	89
		Absent	549	671	863	98	99	99
	Segment	Present	576	668	796	98	94	85
		Absent	591	796	1094	97	98	97
D	Overall	Present	507	559	577	99	98	98
		Absent	514	593	658	99	99	99
	Segment	Present	533	582	669	97	97	94
		Absent	546	692	891	97	99	98

separated out in this way, the interposed square could potentially interfere with the formation of the assemblies, and thus slow down search. However, the results (see Figure 12 and Table 7) show that this does not happen: rates for the different-overall test were 2.5 and 8.5 ms/item, and for the different-segment test were 11.1 and 25.2 ms/item, an assembly-based pattern not significantly different from that of Experiment 1 (Conditions A and B, $ps > .05$) or Experiment 4 (Conditions A-C, $ps > .05$). Conditions A and B therefore show that the square had no special status in regard to grouping, at least in its ability to *interfere* with the grouping process.

The next two conditions (C and D) examined the case where complexity stemmed from luminance pattern rather than geometrical shape. In these conditions, stimuli were presented against a uniform gray background. Condition C served as a control, using stimuli (those of Experiment 4E) known to yield a bundle-based pattern. As shown in Table 7 and Figure 12C, mean rates for the different-overall test were 12.8 and 25.9 ms/item, while rates for the different-segment test were 18.4 and 42.2 ms/item [$t(27) = 2.50$, p < .05 for target-present; $t(27) = 7.28$, $p < .01$ for target-absent]. Comparisons against Experiment 4E showed that the gray background did not significantly affect search speed [both $ps > .05$].

In Condition D, the solid black test line was replaced by a tightly-packed array of alternating black and white dots (see Figure 12D). If the dot array is treated as a single complex structure (namely, a striped line), search speeds in Conditions C and D should be virtually identical, or might even slow down. But if the dots retain their separate identities, they should generate virtual lines and so create an assembly (as in Experiment 4A) that will yield faster search. As seen from Figure 12 and Table 7, the latter result was found: rates for the different-overall test were 5.9 ms and

11.7 ms/item, and for the different-segment test were 11.3 ms and 28.3 ms/item [$t(27) = 2.41$, p < .05 for target-present; $t(27) = 7.41$, $p < .01$ for target-absent], an assembly-based pattern not significantly different from that of Experiment 7D (all $ps > .05$). Given the inability to select image components of opposite contrast sign that are connected together (Experiment 5), it would appear that these tightly-packed dot arrays have maintained their identities, at least as far as generating virtual lines is concerned. Taken together, the results of Conditions A-D indicate that if complex structures are involved in low-level grouping, they at least have no special status or property that adversely affects its operation.

General Discussion

The experiments presented above show that visual search for ML stimuli was consistently influenced by the complete configurations, and not by their component lines. Indeed, this effect was strong enough to yield high-speed "pop-out" search even when no differences existed in the lengths of the component lines (Experiments 1-3). When the lines of the ML stimuli were attached to each other at their endpoints, no amount of effort or selective attention could weaken the effect (Experiment 5).

In addition, search was highly sensitive to the particular arrangement of the line segments in each item. Although it is difficult to make direct comparisons of search rates for all the different stimuli, many of them differed only slightly from the "standard" ML configurations used in Experiment 1. A comparison of the rates for these stimuli is presented in Figure 13. Here, the two axes of the graph represent the search rates for the two kinds of stimulus tests: different-segment and different-overall[3]. Values along each axis rep-

---

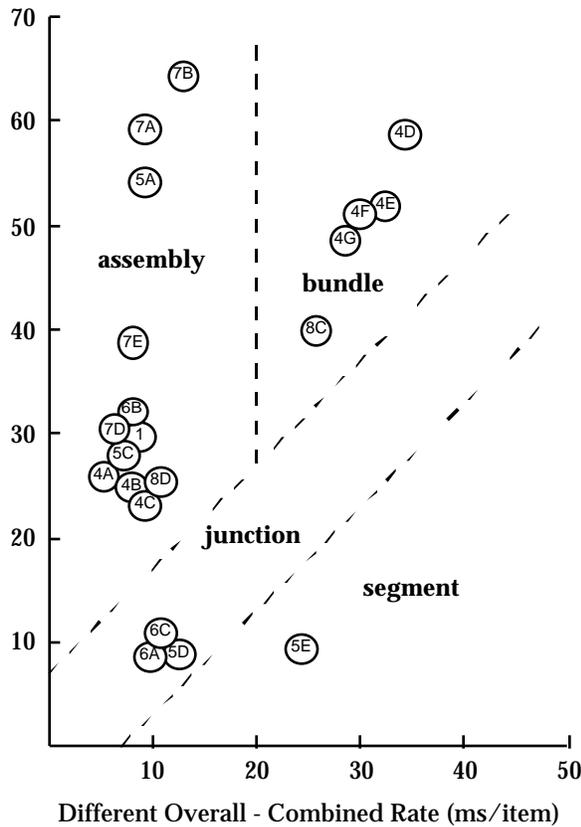[3] We wish to thank Steve Palmer for suggesting this format.

Figure 13. Graph of the combined search rates for stimuli similar to the "standard" ML configurations of Experiment 1. Plotted search speeds are the mean of the target-present rate (multiplied by 2) and the target-absent rates (see text for detailed explanation). To avoid cluttering the graph, the narrow-gap conditions of Experiment 6 have been omitted; the corresponding points fall within the assembly-based cluster.

resent a "combined rate" that is the average of the target-present and target-absent rates. In this measure, the target-present rates are multiplied by a factor of two to compensate for the fact that a self-terminating search will on average take only half as long to detect target presence as it will target absence. Note that the combined rate is not meant to have any theoretical significance—it merely serves to help summarize the data.

An inspection of Figure 13 reveals that search rates clustered into at least four different classes. One of these is *junction-based* search, represented by points that fall on or slightly below the main diagonal of the graph. For this class, search was fast for different-segment stimuli and even slightly faster for different-overall stimuli, indicating that search was based on junctions alone (Experiment 2). Another class is *segment-based* search, represented by points that fall well below the diagonal. Here, different-segment stimuli were detected much faster than different-overall stimuli, indicating that search was based only on test-line segments.

The other two classes are located above the diagonal: a relatively fast *assembly-based* pattern, and a somewhat slower *bundle-based* pattern. In both cases, search was governed entirely by the complete stimulus configuration, with subjects being unable to rapidly access individual

segments or junctions. Both patterns were also affected by gaps in a similar way: if disconnected components had the same contrast sign, search was unaffected (Experiment 4); if they had opposite contrast signs, search became much faster, as if the components could now be rapidly accessed (Experiment 5). The transition between these two patterns is sharp—even the addition of a few small line segments between isolated wings was sufficient to change bundle-based into assembly-based search (Experiment 4).

Taken together, these results indicate that groups are formed at levels below those easily accessible to visual search, and that these groups preempt their component segments. If one believes that visual search has rapid access to image-based primitives at the earliest levels of vision, two routes are available to explain these effects. The first appeals to mechanisms at the lowest levels of visual processing, such as one- or two-stage spatial filtering. But filtering is unable to explain either the large changes in search speed with small changes in stimulus shape, or the uniformity of speeds within each class. Alternatively, it could be argued that these effects are due to higher-level mechanisms acting on the outputs of preattentive maps. But this is unable to account for either the indifference of the effects to selective attention, or the preemption of individual components.

We therefore argue that visual search is based on groups formed by processes more sophisticated than simple spatial filtering. We further argue that two different kinds of groups are formed: *assemblies* and *bundles*. Note that in this view, junctions are also assemblies, but with a different shape than the assemblies corresponding to ML configurations. Search for the relatively distinct (ML) assemblies is based on the length estimates assigned to them via their (interior) contours; search for the relatively indistinct bundles—which have no such contours—is based on overall extent[4]. Since the pixel density of the items was controlled, as was the shape of the resulting groups, the particular pattern of search rates simply reflected the success or failure of these intra-item grouping processes.

Low-Level Grouping

The experiments presented here point clearly toward the existence of some form of low-level intra-item grouping. Because only ML configurations (and a few variants) were used, it is almost certain that these experiments did not examine all aspects of its operation. And given that relatively few types of ML stimuli were tested, it is likely that they did not even examine all aspects that apply to ML configurations. Nevertheless, the formation of the assemblies and bundles found here can be accounted for by three simple grouping processes, which would presumably be involved in the formation of other, more complex structures as well.

---

[4] The measurement of the overall extent (or "size") of a bundle *can* be done via simple spatial filters; indeed, such a mechanism has been proposed to explain the sensitivity of texture perception to the spatial extent of texture elements (Bergen & Adelson, 1988; Gurnsey & Browse, 1989). Measurements of overall extent can be assigned to assemblies as well, but are likely to be less salient than the estimates derived from the interior contours.

Grouping Processes

From a purely descriptive point of view, three grouping processes suffice to account for the results obtained here (Figure 14). It should be emphasized that the description of these processes—and their relation to each other—is not meant to be definitive in all its details. Rather, it is only intended as an illustration of how low-level grouping might be carried out.

alpha (attachment): Groups formed of lines (including virtual lines of dots) attached to each other at their endpoints. Note that the endpoints of *both* segments must contact each other—T-junctions formed by the attachment of only one endpoint are not grouped together. Contrast sign is unimportant. The result is a "strong" assembly, i.e., one that cannot be broken by subsequent processing. This type of grouping is the only one used when no gaps are present in the test lines (Experiments 1-3), or when dot arrays generate virtual lines that have no well-defined endings (Experiment 7).

beta (bundle): Groups formed of structures (alpha-groups, isolated lines and dots) that are relatively near to each other (e.g., within some distance limit). Contrast sign is unimportant. The result is a "bundle" with no interior structure; only its overall extent is available for rapid search. Such grouping is reflected in the bundle-based search patterns of Experiments 4 and 8. The resultant bundle is "fragile", i.e., can be disbanded by subsequent processing.

gamma (contour): Groups formed from structures within a bundle that have "intersectable" interior segments (i.e., segments that contact each other when extended—

see Figure 14), or have endpoints that are between the interior segments are large. The result is a "strong" assembly. This type of grouping is found, for nearby. Contrast signs must match when distances example, in Experiments 4A-C. Note that segments must be in the interior (or obey some related "trigger" condition) for gamma grouping to occur—otherwise, *any* set of intersectable segments (e.g., the "wings-in" brackets of Experiments 1D-G) would be grouped in this way. If the interior segments are non-intersectable and not sufficiently near to each other (e.g., Conditions A and C of Experiment 6), the bundle is disbanded, and the component structures become available for rapid search.

Explanation of Empirical Results

To show that the grouping processes proposed here can explain the results of the experiments, we make the generally-held assumption that search rates reflect the signal-to-noise ratio of the target amid the distractors, with fast search when this ratio is high (Duncan & Humphreys, 1989; Treisman & Gormican, 1988). We also assume that search is based on the overall length of the assemblies and the bundles, and that the assemblies support an additional property (likely based on interior contour) that makes targets and distractors more distinct from each other.

The results of Experiments 1-3 follow straightforwardly. Since all segments are attached to each other, they are alpha grouped into assemblies, with search governed by overall length. Although the results of Experiment 2C appear to be anomalous in that speed increased as the test lines were shortened, this could have been simply due to the relatively large difference in pixel density (or equivalently, in overall size), a quantity known to support rapid search (Gurnsey & Browse, 1987).

The results of Experiment 4 are also easy to explain. In Conditions A-C, the wings and attached test line segments were alpha grouped into assemblies, and then beta grouped. Since the interior segments were intersectable, gamma-grouping formed assemblies similar to those of Experiment 1 (see Figure 14). The stimuli of Conditions D-G did not have interior segments that could be completed (see Figure 14), and so remained as relatively indistinct bundles.

The results of Experiment 5 follow from the sensitivity of the various grouping processes to contrast sign. Grouping was maintained for the alpha groups of Conditions A and B, leading to assembly-based search. The contrast-indifference of short-range gamma grouping also caused assemblies to be formed in Condition C; the sensitivity of long-range grouping to contrast sign caused the bundles to be disbanded in Condition D, allowing search to be based on individual junctions. Note however that the dependence in search speed with gap size (or equivalently, segment length) found in Conditions C/D is also found for isolated junctions (Experiment 2) indicating that search in Condition C might have been based on isolated junctions. The experiments here are not sufficient to settle this issue unequivocally; if individual junctions were being accessed in Condition C, it would imply that the short-range, contrast-indifferent aspect of gamma grouping does not exist.
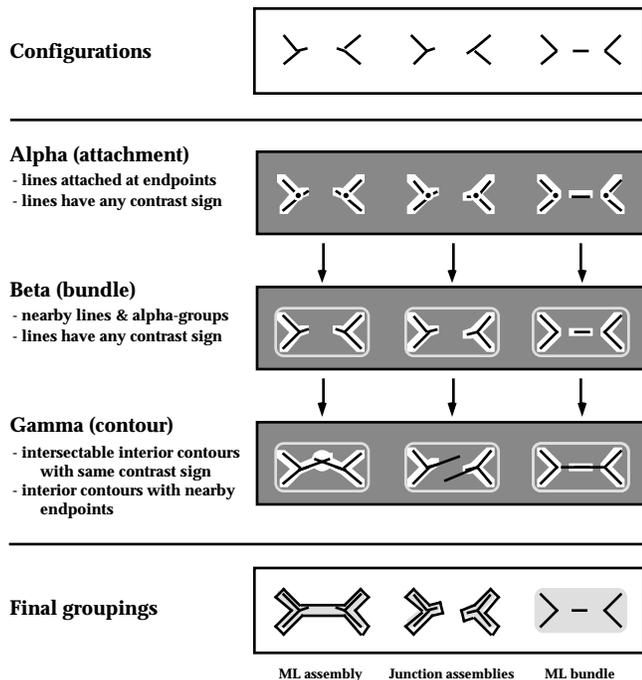

| | ML assembly | Junction assemblies | ML bundle |

Figure 14. Illustrations of the three types of low-level grouping proposed to account for the visual search data in this study (see text for details). ML = Mueller-Lyer.

The search pattern for Condition E clearly indicates that the central test line can be rapidly accessed. This is somewhat anomalous—since alpha- and beta-grouping are contrast indifferent, a search pattern similar to that for the uniform-contrast stimuli of Experiment 4 should be expected. However, it may be that the bundles formed by beta grouping are sufficiently fragile that line segments can be extracted by selective attention.

The results of Experiment 6 can be explained in a fashion similar to that for Experiments 5C/D. The line segments were first alpha grouped into junctions, and nearby junctions then beta grouped into a bundle. When the interior segments were intersectable (Condition B), they were gamma grouped into an assembly. This also occurred when the endpoints of the interior segments were sufficiently close to each other (Conditions A1 and C1). When these segments were non-intersectable and their endpoints too distant (Conditions A1 and C1), the beta group was disbanded, allowing the individual junctions to become available for rapid search. As for the case of Experiment 5C/D, the sensitivity to gap size found in Conditions A and C could indicate that the assemblies accessed in A1 and C1 were junctions rather than ML configurations; if so, it would imply that there is no short-range component to gamma grouping. Further experiments are required for a definitive conclusion.

Experiment 7 used the same stimuli as Experiment 1, except that dot arrays replaced the line segments. Since the dot arrays generated virtual lines largely equivalent to luminance-defined lines, similar grouping operations took place, resulting in assemblies of similar form. The assembly-based pattern follows straightforwardly.

Finally, the indifference to sub-structure found in Experiment 8 was a direct consequence of the simplicity of the grouping process. Since there were no free endpoints in the stimuli of Condition A, the isolated wing segments were only beta grouped. The line segments occluded by the square in Condition B, however, could be gamma grouped since their endpoints were effectively still free. The results of Condition C followed directly from beta grouping, while those of Condition D showed that the central dots retained their ability to link up with the wings into a gamma group. This is consistent with the alpha grouping of the adjacent dots, provided that the virtual line generated by the dots prior to alpha grouping is not discarded.

Computational Analysis

From a computational point of view, it is useful to have a linking operation of some sort in early vision, not only to determine extensive properties like length, but more generally, to determine what parts of the two-dimensional image correspond to the same object in the three-dimensional scene (see, e.g., Horn, 1986). Due to factors such as noise and occlusion, an object in the scene can give rise to several disconnected elements in the image, and it is important to identify these elements and link them together as soon as possible. Owing to the time and space limitations on early vision (see Zucker, 1987b), the linkages formed at this stage are not likely to be optimal, since this would require collecting and interpreting information from the entire image. But a time-limited process can be effective when limited to areas that are contiguous and compact (Enns

& Rensink, 1992; Rensink, 1992), so that rapid linking would have a reasonable chance of success when gaps are relatively small. This restriction to small gaps permits a powerful "local coherence" constraint that can simplify its operation: *since objects in the world are usually coherent on a local scale, changes across small gaps are highly unlikely*.

The rules of low-level grouping have only been partially explored here, and a complete computational theory of the processes involved must await future experiments. But a preliminary analysis of the processes proposed here may be carried out, based on considerations of time limitations and local coherence.

i) Alpha grouping

If objects are to be formed via the linking of contour fragments, one of the most reliable cues is *cotermination*: lines attached at their endpoints in the image are likely to correspond to edges attached in the scene. Although this association requires some constraints on the scene if it is to be valid, these are usually obeyed in the real world (Richards and Jepson, 1992).

Thus, a plausible candidate for the first stage of rapid grouping is the alpha grouping process proposed above. This indifference of this process to the medium used (black, white, or virtual lines) provides a simple way to combine information from several sources, allowing use of the best contour estimates. The advantage is that the effects of noise and uncertainty are minimized, increasing the robustness of the operation. The disadvantage is that contours from different objects (e.g., those of different contrast sign) might be erroneously bound together. It would appear that attachment is regarded as sufficiently restrictive that the gain in robustness is worth the higher risk of error.

This type of grouping can be readily carried out via the binding[5] of lines attached at their endpoints. Since only the information in the neighborhood about the attachment point is required, the line-binding process is extremely local and can be carried out in parallel with a minimum of processing time. Particular details need not be discussed here; algorithms for such processes have been developed, including some that can operate on noisy gray-scale images (e.g., Walters, 1987). Note that cells have been found in visual area V1 that respond selectively to virtual lines (Grosof et al., 1993), supporting the idea that this grouping may take place at very early levels.

ii) Beta grouping

The next stage in the formation of assemblies is *beta grouping,* defined here to be the linking of all structures (alpha groups, isolated lines and dots) sufficiently near each other in the image (e.g., nearest neighbors, within some spatial range). This type of linking is taken to be relatively fragile, the result being a bundle that can be disbanded by subsequent processing. Such "fragile" forms allow a sense of interior to be defined without commitment to a more definite assembly. This is particularly important given the

---

4 The terms *binding* and *grouping* are not used as synonyms here. *Grouping* is used in this paper to refer to the effect of image elements being linked into wholes. *Binding* refers to the causal mechanism invoked to explain how the grouping is achieved.

need for interior contours by gamma grouping. If grouping requires an interior, and the interior is unavailable without grouping, a "chicken-and-egg" situation arises. The use of fragile forms would appear to be a way out of such a situation.

There are several ways of accomplishing beta grouping. For example, the nearest-neighbor structures could be labelled as belonging to the same assembly, with the outer boundary of the resulting assembly (perhaps its convex hull) being the only explicit structure formed. Given a limit on the distance between structures in a bundle, these operations will only require information from local neighborhoods, and so can be carried out in parallel within a fixed amount of time.

iii) Gamma grouping

Once a tentative interior has been formed, it is advantageous to link together those contours that have been interrupted by noise, occlusion, etc. This is the purpose of *gamma grouping*, which binds together interior structures that are spatially separated. In particular, gamma grouping binds together interior segments with nearby endpoints, or that are *intersectable*, i.e., that contact each other when extended.

The linkage of nearby segments is fairly easy to understand; the linkage of more distant segments is justified by more complex factors. As shown in Figure 15, three types of gap can exist between a pair of lines: (A) the

Type of Gap	Possible Occlusion

Figure 15. Possible types of relation between line segments. (A) The extension of each contacts the extension (or endpoint) of the other; interpretation requires at most one change in direction of the edge behind the occluder. (B) Neither line extends to meet at a common point; interpretation requires at least two changes in direction behind the occluder. (C) One extension contacts the other line at its body; interpretation again requires two changes in direction behind the occluder.

extension of each contacts the extension (or endpoint) of the other, (B) neither extension contacts the other, and (C) the 1987). Since smooth curves can change their orientation over small distances, a change in orientation within the occluded region is not exceptional. Although the odds are not high that an arbitrary pair of intersectable lines belongs extension of one contacts the body of the other. Gaps of type A (i.e., those between intersectable lines) can be interpreted as due to an occlusion of at most one corner (Figure 15); a special case of this is collinearity, an important cue in many models of grouping (e.g., Lowe, together in the scene, they improve considerably if the segments have the same contrast sign and have been tentatively identified as part of the same object. Although this does not guarantee that the structures actually *are* part of the same object, the results of Experiments 4-6 show that the likelihood of this is deemed sufficiently high that such linking is worth the potential error that might arise.

In contrast, if gaps of types B or C are due to occlusion, at least *two* corners of the occluded edge must be hidden (or, for curved edges, a change in the sign of the curvature). From the results of Experiment 6, it is evident that such gaps are not bridged; indeed, if interior segments form gaps of this type, it is taken as evidence that the relevant structures are *not* part of the same object. This response is a natural consequence of the principle of local coherence: the likelihood of a narrow occluder hiding two changes in orientation (or a change in the sign of the curvature) is extremely small, and so the linkage is not made. This reluctance to posit such interpretations reflects a highly conservative strategy, which enables rapid processing to avoid forming structures that might later need to be undone (Enns & Rensink, 1992; Rensink, 1992).

Several different kinds of mechanism could carry out gamma grouping rapidly and in parallel. For example, a binding operation could operate directly on the structures formed by beta grouping. Since an outer boundary has been defined, interior segments could be identified as those segments that do not contact it. Binding itself could then be done by having extensions "shoot out" of the free endings of the interior segments, with segments linked if these extensions contact each other.

Relation to Other Studies of Grouping

The low-level processes postulated here are similar in several ways to those found in other studies of visual grouping. However, it must be kept in mind that a wide variety of processes underlie grouping (see, e.g., Pomerantz & Kubovy, 1986; Zucker, 1987a), and so it is unlikely that all these similarities reflect common mechanisms.

Gestalt Theory

One of the earliest and most influential approaches to the study of grouping was Gestalt theory, which proposed that the visual system operates on a principle of figural "goodness" (see, e.g., Pomerantz & Kubovy, 1986; Rock & Palmer, 1990). This principle gave rise to several laws of grouping, such as grouping by proximity, similarity, and good continuation. Several of these factors also appear to govern low-level grouping—proximity via gap size, similarity via contrast sign, and good continuation via
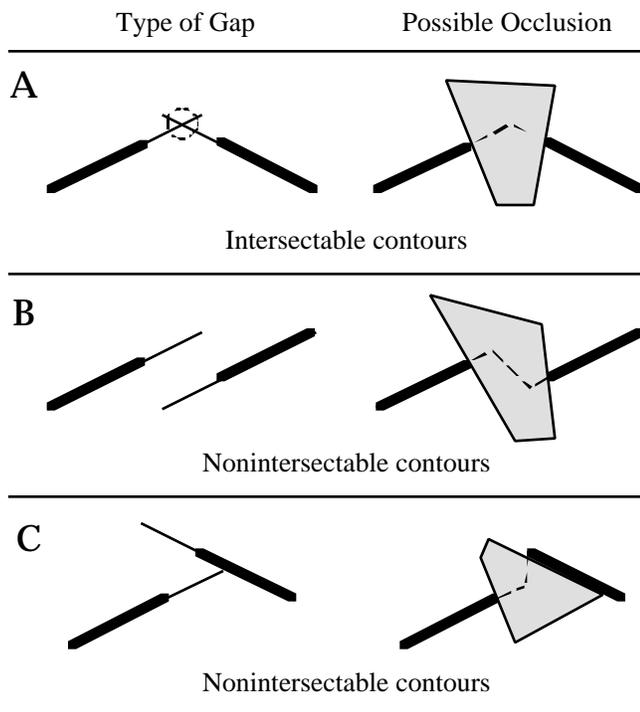
intersectability. Some laws are echoed even more directly. For instance, the law of element connectedness (Palmer & Rock, 1994) is much like alpha grouping, and the law of common region (see Rock & Palmer, 1990) has a correlate in gamma grouping.

These similarities are striking, and suggest that at some level both Gestalt and low-level grouping may draw upon common principles of object (or "proto-object") formation. But important differences also exist. First, Gestalt grouping emphasizes "goodness" and global structure, while low-level grouping emphasizes speed and local processing. Global properties such as symmetry (at least at large scales) therefore have no influence on low-level grouping, which is based instead on local considerations such as gap type. In keeping with its emphasis on more integrated structures, Gestalt grouping also appears to operate at a relatively high level of visual processing (see e.g., Palmer & Rock, 1994). For example, Gestalt grouping acts on representations of three-dimensional surfaces rather than two-dimensional images (Rock & Brosgole, 1964) whereas low-level grouping operates on much simpler structures (although the involvement of three-dimensionality cannot be completely ruled out at this point). Gestalt grouping also appears to require attention (e.g., Ben-Av, Sagi, & Braun, 1992; Mack, Tang, Tuma, Kahn, & Rock, 1992), at least for the grouping of items into rows and columns; presumably this would also apply to the grouping of segments into items. In contrast, low-level grouping acts "preattentively", i.e., at a level where attention is not required (see, e.g., Treisman & Gormican, 1988).

Palmer & Rock (1994) suggest that Gestalt grouping may act on "entry-level" structures that correspond to regions of relatively uniform image-based (or even scene-based) properties. If so, low-level grouping could play a role in the formation of these regions, e.g., via the alpha grouping of same-contrast segments into uniform assemblies. However, there are many differences from the process suggested by Palmer & Rock. For example, alpha grouping also links together segments of opposite contrast sign, something not possible when forming regions of uniform color. And although assemblies might be identified with uniform regions (or at least their boundaries), there is nothing that can be identified with bundles. There is also no provision for handling virtual lines of widely-spaced dots. More generally, there is no low-level grouping of widely-separated elements in Palmer & Rock's model, since they regard this to be handled by higher-level Gestalt operations. However, the uniform regions they propose might serve as inputs to low-level grouping.

Visual Interpolation

An approach to grouping based primarily on considerations of occlusion is the visual interpolation theory of Kellman and Shipley (1991). Here, unit formation is taken to occur when two surface edges are *relatable*, i.e., when their extensions contact each other at an angle of at least 90°; the contrast signs of the edges are irrelevant. If a pair edges are relatable, a contour is then interpolated between them. This process is triggered by local spatiotemporal discontinuities, with no explicit search through all possible (global) configurations. The result, however, is an object with a minimal number of

discontinuities in the tangent of its boundary (Kellman and Shipley, 1991).

The emphasis on object contours and their spatial relations have strong parallels in low-level grouping. In particular, there is a strong connection between "relatability" and the intersectability condition of gamma grouping. In a sense, this aspect of low-level grouping could be seen as a simple form of the process Kellman and Shipley propose, based on upon many of the same considerations.

Some differences do exist, however. Relatability applies to *surface edges* rather than lines. Since edges must "belong" to one of their flanking regions (Nakayama, Shimojo, & Silverman, 1989), border ownership must be consistent across the completed edge. No such concern arises in low-level grouping, where linkage is determined entirely by the contrast and spatial relations of the lines. Also, relatability has no sensitivity to contrast sign, whereas (long-range) gamma grouping does. Finally, relatability applies to three-dimensional constructs formed at relatively high levels (Kellman & Shipley, 1992) rather than to low-level structures such as dots and lines.

Configural Dimensions

An approach more in the spirit of the present inquiry is the study of grouping via its effects on selective attention (Pomerantz, 1978, 1981; Pomerantz & Pristach, 1989). If a set of line segments has been grouped into a unitary (or "configural") dimension, it will be easy to spread attention among them and difficult to attend to only one; otherwise, it will be difficult to spread attention among them and easy to attend to only one. Unitary dimensions are therefore those that allow the fastest classification of individually-presented line drawings. Although the range of figures tested this way is not extensive, effects have been found that are similar to those found here. For example, the grouping of brackets sufficiently close to each other (Pomerantz, 1978) is similar to beta grouping, as is the indifference to small displacements of line endpoints (Pomerantz & Pristach, 1989).

It would be interesting to see if the grouping rules found here also govern the formation of configural dimensions. However, a complete match of these rules is unlikely. To begin with, configural dimensions are believed to be "emergent features" that do not preempt the properties of the component line segments (Pomerantz & Pristach, 1989). If the overall lengths of the assemblies or bundles are instances of these emergent properties, they would be rather anomalous, since the experiments here indicate that the components *are* preempted by the low-level groups. In addition, some configural dimensions found in individually-presented figures are not detected when multiple copies are presented simultaneously (Pomerantz & Pristach, 1989). This suggests that configural dimensions may originate at a somewhat different (and presumably higher) level concerned with the formation of individual objects.

Primal-Sketch Grouping

The first computational account of grouping at early levels was given by Marr (1976, 1982) in his theory of the primal sketch. The primal sketch serves as an explicit representation of the interesting structure in the image, and is assumed to be formed "nonattentively" (Marr, 1976). The

basic elements to be grouped are localized image discontinuities that have been detected by spatial filters. Grouping takes place at two levels: an intra-item grouping of the elements into meaningful "tokens" (e.g., edge segments and endpoints) and an inter-item grouping of the tokens into larger-scale structures such as texture regions.

The intra-item grouping in the primal sketch was developed via quite general considerations, and there are several interesting similarities with the process proposed here, at least at a general level. To begin with, tokens include not only edges and edge endpoints, but also virtual lines formed of dots (Marr, 1976, 1982); as such, they parallel the structures that enter into the basic alpha groups. Primal-sketch grouping is almost completely recursive, with higher-order tokens built out of lower-order ones (Marr, 1982). The limited time and processing power available to early vision make it unlikely that recursion would be unlimited, but the three sequential grouping stages proposed here could reflect a practical embodiment of this idea.

There are also some similarities in the particular grouping rules used. One type of primal-sketch grouping is *curvilinear aggregation*, which joins together tokens that are aligned so as to form curves. This is somewhat like gamma grouping, except that tokens must be very close together, and need not be in the interior of an object.

However, other types of grouping proposed by Marr—while based on the similarity and the spatial arrangement of the tokens—differ substantially from those proposed here. One these is *theta-aggregation*, which groups together structures that are arrayed in a line, regardless of the orientations of the structures themselves; in essence, this is a higher-order form of curvilinear aggregation. There is no known correlate to this in low-level grouping, although further tests may reveal one. Another type of grouping is *clustering*, which joins nearby tokens that are similar to each other. This has no direct analog in the processes proposed here; at best it corresponds to a weakened form of beta grouping, but with a sensitivity to contrast sign. Also, the output of this process is a cluster that cannot be disbanded by subsequent processing, which contrasts with the fragile bundles formed by beta grouping. Indeed, the concept of a structure that can be broken down by subsequent processes is at odds with Marr's principle of least commitment, which insists that no computation be undone (Marr, 1976).

Spatial Filtering

Several kinds of grouping can result from the blurring induced by spatial filters (see, e.g., Ginsburg, 1986; Watt, 1988). Interestingly, many aspects of texture perception can also be accounted for via such filters (e.g., Beck, Sutter, & Ivry, 1987; Bergen & Adelson, 1988; Gurnsey & Browse, 1989; Malik & Perona, 1990; Rensink, 1987; Voorhees & Poggio, 1988), and it is tempting to explain low-level group- ing in the same way. Indeed, several of the experiments presented in this paper were explicit tests of this hypothesis.

The virtual lines used in alpha grouping—at least those based on same-contrast dots—could certainly be obtained via spatial filtering, since dot arrays can be blurred into lines nearly indistinguishable from luminance-based ones (Ginsburg, 1986). And the sensitivity to the overall extent of beta groups may be based on their compactness, a quantity that is also easily determined via such filters (Beck et al., 1987; Gurnsey & Browse, 1987; Watt, 1988).

But as shown in Experiments 4-8, simple filtering—even if carried out over several different spatial scales—cannot easily account for everything. It cannot easily explain the sensitivity to gap type found in Experiment 4: blurring would always cause narrow gaps to be filled, regardless of the structures that surround them. It cannot easily explain the sudden transition in search speed when the small line fragments are dropped from the wings (Experiment 4), or when a second dot is added between the wings (Experiments 4 and 7). And it cannot easily explain the "emergence" of individual junctions when the interior line segments no longer extend to contact each other (Experiment 6).

Any change in an image will always induce a change in activity among a sufficiently rich set of filters, and so explanations in terms of spatial filtering can never be ruled out completely. However, the appeal here is to the particular *pattern* of results—in particular, to the apparent involvement of likelihoods based on the separation, orientation, and contrast sign of the line segments. This pattern could be interpreted as due to "side effects" or "accidental sensitivities" of a simple filtering mechanism concerned primarily with tasks such as edge detection. But the functional advantages of the rules uncovered here would seem to favor the involvement of rapid, sophisticated grouping operations directly concerned with surface and object formation.

Receptive-Field Interactions

Grouping has also been studied via the subjective (or "illusory") contours that are induced between some grouped structures. Although these contours do not correspond directly to any structure in the image, they are equivalent in many respects to the "physical" contours derived from luminance differences (see, e.g., Petry & Meyer, 1987). It is generally believed that subjective contours result from boundary-completion processes carried out via nonlinear interactions among receptive fields at early stages of visual processing (e.g., Grossberg & Mingolla, 1985a; von der Heydt & Peterhans, 1989; Zucker, 1987a). Cells in cortical area V2 have been found that are responsive to such contours (von der Heydt & Peterhans, 1989; Peterhans & von der Heydt, 1989, 1991), and psychophysical evidence has also been found for their existence at early levels (Gurnsey et al., 1992).

Although the various models differ in detail, they are usually based on simple filters assumed to correspond to the receptive fields of cells in visual area V1. In many of these models the filters are combined in some nonlinear way to become responsive to line endings; relative activity among neighboring filters then determines whether adjacent endings should be joined via a subjective contour (Grossberg & Mingolla, 1985a, 1985b; Peterhans & von der Heydt, 1991).

Such nonlinear mechanisms can lead to grouping processes that are considerably sophisticated. But are they sophisticated enough? The basic structures used in low-level grouping could certainly be based on such mechanisms, which readily form virtual lines from dot arrays (e.g., Peterhans & von der Heydt, 1989, 1991; Zucker, 1987a). The indifference of alpha- and beta grouping to contrast sign could likewise be explained by the discarding

of this quantity at the earliest stages of operation (Grossberg & Mingolla, 1985a).

However, these mechanisms cannot account for all the phenomena encountered here. Even if they could carry out alpha- and beta grouping, they could not explain the contrast-dependence of gamma grouping (although the contrast-sensitive feature-contour system of Grossberg and Mingolla might be of some relevance here). In addition, these models do not distinguish between interior and exterior contours, which makes the triggering of gamma grouping difficult to explain. Thus, although nonlinear interactions among local filters may well be the means by which low-level grouping is implemented, the interactions required would need to be more complex than those used in models to explain subjective contours.

Relation to the Mueller-Lyer Illusion

In the experiments here, most stimuli were designed to yield groups of potentially the same form. This allowed length assignment to be factored out, so that search rates could directly signal the existence or nonexistence of grouping. But what of the length-assignment process itself? Given that strong configural effects exist even for attentively-viewed ML configurations (i.e., the ML illusion), it is worth examining whether analogous (or even the same) mechanisms are involved.

As a first step, it is important to note that the connection between the two sets of configural effects is not necessarily a direct one. To begin with, it was only assumed that length was the basis of search in our experiments. But any assembly-based property (such as area) could have been involved, provided only that it led to an illusion of some kind. However, the transition between the assembly-and bundle-based patterns indicates that a property directly obtained from the interior contour was almost certainly responsible. Length is the most likely candidate in this regard: It is one of the simplest and most fundamental properties, entering into the determination of many quantities and structures (see, e.g., Marr 1976, 1982) and it is difficult to see why such a basic property would be irrelevant.

Given that length estimates were the basis of search for both assemblies and bundles, it is unlikely that these estimates were based only on their maximum extent, for otherwise there would be no difference between the two patterns. One possibility is that the estimates obtained from the assemblies also include an estimate of the test-line length. Given the inaccessibility of individual segments, this estimate would not be a physical length obtained directly from the segment, but an "apparent" length recovered from the entire assembly. If such an estimate were involved, the difference between the assembly- and bundle-based patterns could be attributed to a low-level length illusion directly analogous to the attentively-viewed ML illusion.

However, the particular length estimates assigned to assemblies and bundles cannot be determined from the experiments here. The focus of this paper has been on the existence of configural effects, with only a rough comparison of test line lengths being carried out. To determine the length estimates assigned to the groups would require establishing the point of subjective equality, i.e., the length for which search is maximally slow (taking into account the effects of other quantities, such as overall and segment length). Given the lack of parametric variation in the stimuli used here, the length-assignment mechanisms can only be compared to those proposed for the ML illusion in the most general terms.

The ML illusion is believed to arise from mechanisms at several different levels of processing (Coren & Girgus, 1978). The simplest of these are "structural" mechanisms, including blurring via optical scattering in the eye, lateral inhibitory connections in the retinal ganglion cells, and neural pooling in the cortex (Coren, 1970). This blurring causes the ML figure to distort, with the wings-out configuration becoming longer than the wings-in configuration (Coren, 1970; Ginsburg, 1984, 1986). Since these mechanisms operate at the earliest stages of vision, they act prior to the formation of low-level groups, and so are unlikely to be involved here. This is consistent with the indifference of search to the replacement of lines by dot arrays—if blurring were involved, it would have to act on virtual as well as on luminance-defined lines.

The ML illusion may also be influenced by higher-level cognitive factors. One of these is size constancy, which would rescale the length of the test line based on the depth suggested by the wing arrangement (Gregory 1963; Tausch, 1954). However, the difference between physical and apparent length for any section of test line depends on the distance of that section from the wings (Morgan, Hole, & Glennerster, 1990), something not consistent with uniform rescaling. In any event, inferred depth is almost certainly not part of low-level length estimation—its assignments are maintained when gaps are placed in the center of lines, whereas this manipulation destroys the rapid recovery of three-dimensional structure at early levels (Enns & Rensink, 1991a). In addition, low-level estimation is almost certain to be free of several other factors believed to contribute to the ML illusion, such as focused attention (Coren & Girgus, 1978; Coren & Porac, 1983) and eye movement patterns (Coren, 1986).

Thus, the only length-determining mechanisms that the ML illusion could share with its low-level counterpart are those that operate on the line assemblies themselves. Several such mechanisms are possible. For example, length might be taken as the distance between the "average" positions of the endpoints, where each position is the centroid of the structures in the nearby neighborhood (Morgan et al., 1990). Note that there are also several ways in which the overall length can be defined, such as the maximum extent, or the diameter of the *minimal enclosing circle*, i.e., the smallest circle containing the assembly (Gurnsey & Browse, 1987). Determination of the particular mechanisms used must await more detailed parametric studies.

Relation to Current Theories of Visual Search

The existence of grouping at relatively early levels of processing implies that visual search can rapidly access structures with a considerable degree of complexity. One way to better understand the nature of these groups—and the preemption they induce—is to carefully examine the relation of low-level grouping to the other processes believed to exist at this level of vision.

Grouping and Visual Search

Several theories have explicitly postulated a role for perceptual grouping in visual search and texture perception (e.g., Beck, 1982; Humphreys et al., 1989; Julesz, 1984; Treisman, 1988). Among the more influential of these is *feature integration theory* (Treisman & Gelade, 1980; Treisman & Gormican, 1988) where the basic structures are simple *features* formed rapidly and in parallel across the visual field. Although these features can be arrayed homogeneously across the visual field, they can also be grouped into clusters, each of which can be inspected via a "spotlight of attention". Thus, when observers are shown small clusters of items and asked to report on the presence of a single target, they search serially from cluster to cluster, with parallel search within each cluster (Treisman, 1982).

Such inter-item grouping has also been postulated as the reason for the dependence of search speed on target-distractor similarity (Duncan & Humphreys, 1989; Treisman & Gormican, 1988). It has even been postulated as the basis of object formation (Donnelly, Humphreys, & Riddoch, 1991; Pomerantz, 1981). For instance, Donnelly et al. (1991) showed that a conjunction-defined target (e.g., an oriented L) embedded among a globally-defined shape (e.g., square or hexagon) could be quickly detected if it violated the properties of closure and/or good continuation; otherwise, search was slow. Note that this form of grouping differs from the low-level grouping found here not only because it acts *between* rather than *within* items, but also because it differs in its rules of operation: interior contours are not distinguished, and collinearity—rather than proximity or intersectability—is the key geometric constraint.

The existence of inter-item grouping does not necessarily conflict with the grouping process suggested here, for inter-item grouping occurs *after* the formation of the rapidly-accessible features, whereas low-level grouping occurs *before* their formation. Thus, our proposal is consistent with—and in fact is complimentary to—the various models of inter-item grouping. The existence of these two levels simply reflects the fact that different kinds of grouping process are suited for different kinds of tasks (see Zucker, 1987a); indeed, computational considerations suggest that intra- and inter-item grouping are both needed for the formation of early visual representations (Marr 1976, 1982).

Although relatively little attention has been paid to intra-item grouping, its existence has been suggested by several studies. For example, Prinzmetal (1981) studied the role of Gestalt grouping factors on illusory conjunctions. Although that paper focused on the effects of spatial proximity and color similarity on the integration of features into more complex conjuncts, the experiments also showed significant grouping influences on the detection of the features themselves.

The effect of spatial proximity on feature formation also became part of the texton theory of texture perception (Julesz, 1986, 1991). In the original account (Julesz, 1981), texture perception was based on primitive elements—called *textons*—that were registered independently and in parallel at early levels. These included oriented lines, line endings, and line intersections. Textons were somewhat similar to features, and were also considered to be the basic elements of visual search (Julesz & Bergen, 1983). The finding that

their formation was influenced by nearby textons (Sagi & Julesz, 1985) and line elements (Enns, 1985, Julesz, 1986) led to the addition of a spatial-proximity threshold that determined whether adjacent elements were part of the same texton. In other words, a rough grouping of elements was required *prior to* texton formation. But this grouping only determined *which* elements belonged to a texton—it never entered into the formation of the properties themselves, or caused preemption of their components. In addition, these spatial-proximity factors can be largely explained in terms of spatial frequency filtering (e.g., Bergen & Adelson, 1988; Gurnsey & Browse, 1987), something unlikely to account for the results found here.

It is only recently that studies have begun to examine complex intra-item processes in visual search and texture perception. This was primarily done in terms of the rapid recovery of scene-based properties, which depends critically on intra-item structure (e.g., Enns & Rensink, 1990a, 1991a; Ramachandran, 1988). An early attempt to show intra-item grouping was that of Bravo & Blake (1990), who suggested that regions of similarly-oriented line segments could be rapidly grouped and used as the basis of rapid search. However, since texture differences can lead to contours that support rapid search (Cavanagh et al., 1990), it is unlikely that true intra-item grouping was involved.

Besides the low-level grouping described in this paper, there also appears to be another type of intra-item grouping, involved in the determination of contour closure (Elder & Zucker, 1991, 1992); it is currently unknown whether it causes preemption. The rapid determination of contour closure is similar in spirit to the rapid recovery of scene-based properties, in that relatively complex outputs are obtained from simple inputs. The relation of this grouping process to those found here is currently unknown, although it is likely to be a relatively close one. In any event, our results extend the domain of such intra-item processing, showing that it is used even to determine length, a property considered to be "primitive" in both psychophysical and computational theories of vision (Marr, 1982; Treisman et al., 1990). They also show that the resultant groups pre-empt their components, indicating that it is *only* these more complex structures that are salient.

Preemption of Primitive Measurements

Much of the motivation for this work has been to determine whether the rapidly-constructed primitives formed at the earliest stages of visual processing can always be rapidly accessed by visual search. As the experiments here have shown, this is not the case—not only are simple line segments grouped together before being accessed, but they are *preempted* by the resultant groups (cf. Liberman & Mattingly, 1989), effectively becoming unavailable for rapid search.

In its weakest sense, "preemption" only refers to the failure to obtain *information* about a part of the stimulus when it is surrounded by a particular context. The center lines of the different-segment ML stimuli, for example, are preempted in this sense because the information they contain is not accessible to the search mechanisms when the wings are present. This effect could be due to a variety of factors, e.g., distortions induced by the surrounding parts when the image is blurred. In such a case, preemption arises from the

loss of information incurred at the very earliest stages of unit formation.

However, there is also a stronger sense of "preemption" which refers to the failure to obtain access to the *representation* of a stimulus part. This requires establishing the existence of some kind of perceptual element and then showing that this element in no longer accessible under certain conditions. We believe that our experiments also show evidence for preemption in this stronger sense. For example, the diagonal wings of the ML stimuli should have little effect on the outputs of the horizontally-oriented filters present at early levels. The large effect of the wings on search rates therefore indicates that if these filters do provide primitive measurements, their outputs are no longer readily accessible.

Indeed, there also appears to be preemption of higher-level structures. For example, the misaligned junctions in the wide-gap stimuli of Experiment 6A2 support rapid search, indicating that they have some kind of existence as a perceptual element. However, when aligned with each other (Experiment 4C), a slowdown of search occurs in the different-segment condition, showing that relevant information has somehow been lost. It is difficult to trace this effect to a loss of information in the junctions themselves; rather, it would appear that they have been pre-empted by their incorporation into assemblies.

Evidence for preemption has been found in earlier studies (Enns & Rensink 1991b, 1992; He & Nakayama 1992; He, Nakayama, & Tumosa, 1991). However, these involved relatively few conditions and so were vulnerable to other interpretations because of the existence of confounding factors. In the Enns & Rensink (1992) study, for example, the two-dimensional (2D) orientations of lines were apparently inaccessible to search when they were embedded in a drawing of a cube. This could have been due to the pre-emption of the lines by their three-dimensional (3D) interpretations as cubes, but it could also have resulted from an accidental sensitivity to particular image frequencies. Likewise, He & Nakayama (1992) used stereopsis to show that differences in 2D shape were apparently inaccessible when the shape was interpreted as the visible part of an occluded planar surface. This could have been due to pre-emption, but it could also have been due to an accidental sensitivity to particular conjunctions of depth planes and line junctions. Furthermore, neither study completely ruled out the possibility of access to lower-level structures: slowdown might have been due to preemption, but it might also have been due to a greater salience for 3D structure, with the 3D properties of the targets and distractors simply overwhelming the differences in their 2D properties.

However, the results here involve only 2D measures, thus eliminating any effects of 3D properties. And unlike those of He & Nakayama (1992), they involve reaction times typical of rapid search. Most importantly, they show a high degree of systematicity across a wide variety of conditions. This pattern not only rules out the possibility of access to primitive measurements, but the particular form of this pattern makes it unlikely to have resulted from accidental sensitivities of any kind.

Thus, visual search is not based on visual primitives. But what then *is* it based on? From a computational perspective, it is unlikely that visual search can rapidly access many different properties on many different levels of visual processing—this would require an enormous number of pathways, resulting in an immensely complex system (Tsotsos, 1990). Rather, it is likely to be based on a fairly restricted "rapid-access domain", involving relatively few properties and structures on relatively few levels. Our experiments indicate that this domain includes only the higher-level, more ecologically-relevant properties of surfaces and (proto-)objects, leaving behind the more primitive measurements of the image. Of course, primitive structures could still be effectively accessed if there is no context that causes them to be changed or incorporated into a more complex structure (as is the case, for example, of isolated lines).

Multiple Levels of Rapid Vision

The existence of low-level grouping implies the existence of at least two levels in early vision: one concerned with the formation of the basic structures to be grouped, and the other with their grouping into more complex assemblies. Given this, the question arises as to how many other levels there might be.

To begin with, there appears to be a considerable hetero-geneity in the processes that occur rapidly at early levels. The types of grouping found here differ in their sensitivity to contrast sign, distance, and type of gap. These in turn differ from the grouping found by Elder & Zucker (1991, 1992), which is always sensitive to the sign of contrast (even when structures are attached to each other) and insensitive to the type of gaps in the lines. A sensitivity to contrast sign also exists for the rapid determination of orientation of virtual lines defined by pairs of dots (O'Connell & Treisman, in preparation). Rapid amodal completion of occluded objects is likewise sensitive to contrast sign, but not to gaps (Enns & Rensink, in press). Rapid line interpretation, on the other hand, is destroyed by the presence of gaps (Enns & Rensink, 1991a).

The simplest view to take is that all these processes operate concurrently and independently on the same low-level representations, and simply differ in the rules that they follow. But might be that rapid processing is carried out over several different levels (and possibly even in parallel streams), each with its own set of operations. If so, interesting questions arise about the interactions and dependencies among these processes. Is the output of low-level grouping used for the interpretation of line drawings? Can the amodal completion of Enns & Rensink (in press) be based on the virtual lines of O'Connell and Treisman (in preparation)? Is there a maximum number of processes that can be used concurrently? An interesting area for future work is a systematic examination of the way in which these various processes interact with each other.

It is also of interest to determine how these *intra*-item processes are related to *inter*-item operations such as the clustering of items into homogeneous groups (Duncan & Humphreys, 1989; Grossberg et al., 1993; Treisman, 1982) or the formation of spatially-extended objects (Donnelly et al., 1991). Are these just the same processes operating over different spatial ranges, or are they fundamentally different? Do the inter-item processes also involve different levels and different streams? Texture boundaries appear to be formed

via a representation different from that accessed by visual search (Enns & Rensink, 1993; Wolfe, 1992b), and the success of spatial-filter models in accounting for texture segmentation (in contrast to their failure for the visual search data here) may reflect the fact that different levels are involved.

More generally, the heterogeneity and interconnection of these processes suggest that they might be part of a complex system that—although involving "early" or "preattentive" levels—is best characterized as a *rapid* visual system. More precisely, rapid vision would involve those aspects of vision carried out within the first few hundred milliseconds of stimulus presentation, i.e., the average time between saccades. This system would be in some sense "orthogonal" to the others, cutting across different levels of visual processing. The lower levels might operate rapidly and in parallel to create "islands" of locally-consistent inter-pretations (Enns & Rensink, 1992) and spatially-extended linkages of simple properties (Rensink, 1992). The higher levels might include those aspects of object and event perception that are carried out within a few hundred milliseconds (Biederman, 1981; Johansson, 1975). Such a system could provide a "quick and dirty" initial analysis of the image, with this "first pass" then serving as the basis for more elaborate analysis later on. Although no more than a tentative suggestion at the moment, this characterization of visual processing may provide a useful perspective that will allow us to better understand its operation.

Conclusions

The results presented here show that visual search is subject to a preemption effect induced by grouping at early visual levels. In particular, search was found to be governed by length estimates derived from the entire stimulus configuration, and not from its individual segments. Two distinct grouping patterns were found that differed in their dependence on the contrast sign, separation, and orientation of adjacent segments. It was shown that these effects cannot be accounted for by low-level spatial filtering or by high-level attentive control. It was then proposed that they are due instead to *low-level grouping*, a rapid intra-item linking of contour segments into spatially-extended groups at a level below those readily available to visual search. A model of this process was sketched and shown to account for the outcomes of the experiments described here.

Because the results of this study are based on a relatively restricted domain of stimuli (namely, ML configurations), many details are likely to be incomplete or inapplicable to larger domains. But the results within this domain are robust and have a considerable degree of systematicity; they also have several general implications for both theoretical and methodological approaches to vision. We briefly discuss four of the more important ones here.

1. Rapid (and possibly parallel) processing can create structures that cannot be accessed by rapid visual search. The results here indicate that line segments are grouped together at a level below those rapidly accessible to search and are then preempted, losing their identity as individual structures. Thus, the set of rapidly-accessible properties and structures does not include everything that is rapidly constructed. Only some rapidly-constructed properties are rapidly accessible, these apparently being ecologically-relevant quantities such as properties of surfaces and objects in the scene.

2. Rapid processing is pervasive throughout early vision, being used to compute a wide range of properties and structures. Such processes have been shown to carry out several aspects of scene recovery at early levels of vision (e.g., Enns & Rensink, 1991a, 1992). The findings here show that this kind of "quick and dirty" intra-item processing even enters into the determination of length, one of the most primitive of the image-based properties.

3. Rapid processing is distributed over several different levels and possibly several different streams. Low-level grouping involves at least two different levels: one concerned with the formation of the basic structures to be grouped (e.g., luminance-defined and virtual lines), and the other with their grouping into more complex assemblies. Grouping itself appears to involve sub-processes that differ in their sensitivity to the contrast sign, endpoint separation, and orientation of adjacent line segments. Indeed, some of these processes have characteristics not found in any other rapid-recovery process, such as indifference to contrast sign and sensitivity to gap type. A systematic comparison of the various kinds of rapid processes might illuminate the relations and interactions that exist between them, thereby providing a new source of information on early visual processing.

4. Visual search can be used to explore the operation of grouping. The techniques developed here allow grouping to be measured by its effect on performance (namely, the relative speed of search) rather than by subjective impression. Furthermore, this performance effect is sufficiently large that it provides a reliable tool for exploring the grouping processes involved. The stimuli here were restricted to a small subset of configurations, which did not even exhaust the possible variations of the ML configuration. The use of other variants may provide additional information on low-level grouping, as may the use of illusions and other configural effects.

References

Beck, J. (1982). Textural segmentation. In J. Beck (Ed.), *Organization and representation in perception* (pp. 285-317). Hillsdale, NJ: Erlbaum.

Beck, J., Sutter, A., & Ivry, R. (1987). Spatial frequency channels and perceptual grouping in texture segregation. *Computer Vision, Graphics, and Image Processing*, *37*, 299-325.

Ben-Av, M. B., Sagi, D., & Braun, J. (1992). Visual attention and perceptual grouping. *Perception & Psychophysics, 52,* 277-294.

Bergen, J. R., and Adelson, E. H. (1988). Early vision and texture perception. *Nature, 333*, 363-364.

Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy and J.R. Pomerantz (Eds.) *Perceptual organization* (pp. 213-253). Hillsdale, NJ: Erlbaum.

Bishop, P. O. (1984). Processing of visual information within the retinostriate system. In J. M. Brookhart and V. B. Mountcastle (Eds.) *Handbook of physiology, section I, volume III: Sensory processes, Part I* (pp. 341-424). Bethesda, MD: American Physiological Society.

Bravo, M., & Blake, R. (1990). Preattentive vision and perceptual groups. *Perception, 19*, 515-522.

Bruner, J. S., Goodnow, J., & Austin, G. A. (1956). *A study of thinking.* New York: Wiley.

Carnap, R. (1966). In M. Gardner (Ed.), *The philosophy of science* (pp. 70-77). New York: Basic Books.

Cavanagh, P., Arguin, M., & Treisman, A. (1990). Effect of stimulus domain on visual search for orientation and size features. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 479-491.

Coren, S. (1970). Lateral inhibition and geometric illusions. *Quarterly Journal of Experimental Psychology, 22*, 274-278.

Coren, S. (1986). An efferent component in the visual perception of direction and extent. *Psychological Review, 93*, 391-410.

Coren, S., & Girgus, J. S. (1978). *Seeing is deceiving.* Hillsdale, NJ: Erlbaum.

Coren, S. & Porac, C. (1983). The creation and reversal of the Mueller-Lyer illusion through attentional manipulation. *Perception, 12*, 49-54.

Donnelly, N., & Humphreys, G. W., & Riddoch, M. J. (1991). Parallel computation of primitive shape descriptions. *Journal of Experimental Psychology: Human Perception and Performance, 17*, 561-570.

Driver, J., McLeod, P., & Dienes, Z. (1992). Motion coherence and conjunction search: Implications for guided search theory. *Perception & Psychophysics, 51*, 79-85.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*, 433-458.

Elder, J.H., & Zucker, S.W. (1991). The importance of closure in visual search. *Investigative Ophthalmology & Visual Science, 32,* 715.

Elder, J.H. & Zucker, S.W. (1992). *Contour closure and the perception of shape.* Technical Report TR-CIM-91-08, McGill University, Montreal, Canada.

Enns, J. T. (1985). Seeing textons in context. *Perception & Psychophysics, 39,* 143-147.

Enns, J. T., Ochs, E. P., & Rensink, R. A. (1990). VSearch: Macintosh software for experiments in visual search. *Behavior Research Methods, Instruments, & Computers, 22,* 118-122.

Enns, J. T., & Rensink, R. A. (1990a). Influence of scene-based properties on visual search. *Science, 247,* 721-723.

Enns, J. T., & Rensink, R. A. (1990b). Sensitivity to three-dimensional orientation in visual search. *Psychological Science, 1,* 323-326.

Enns, J. T., & Rensink, R. A. (1991a). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review, 98,* 335-351.

Enns, J. T., & Rensink, R. A. (1991b). Apparent length in preattentive vision: Evidence for low level grouping. *Investigative Ophthalmology & Visual Science, 32,* 1039.

Enns, J. T., & Rensink, R. A. (1992). A model for the rapid interpretation of line drawings in early vision. In D. Brogan (Ed.), *Visual search II* (pp. 73-89). London: Taylor & Francis.

Enns, J. T., & Rensink, R. A. (1993). Conditional vs. unconditional pop-out. *Investigative Ophthalmology & Visual Science, 34,* 1236.

Enns, J. T., & Rensink, R. A. (in press). An object completion process in early vision. In A. Gale (Ed.), *Visual search III.* London: Taylor & Francis.

Ginsburg, A. P. (1984). Visual form perception based on biological filtering. In L. Spillman & B. R. Wooten (Eds.), *Sensory experience, adaptation, and perception* (pp. 53-72). Hillsdale, NJ: Erlbaum.

Ginsburg, A. P. (1986). Spatial filtering and visual form perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance, vol. II* (ch. 34). New York: John Wiley & Sons.

Graham, N. V. S. (1989). *Visual pattern analyzers.* New York: Oxford University Press.

Gregory, R. L. (1963). Distortions of visual space as inappropriate constancy scaling. *Nature, 199*, 678-680.

Grosof, D. H., Shapley, R. M., & Hawken, M. J. (1993). Macaque V1 neurons can signal 'illusory' contours. *Nature, 365*, 550-552.

Grossberg, S., & Mingolla, E. (1985a). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review, 92*, 173-211.

Grossberg, S., & Mingolla, E. (1985b). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception & Psychophysics, 38,* 141-171.

Grossberg, S., Mingolla, E., & Ross, W. D. (1993). *A neural theory of attentive visual search: Interactions of boundary, surface, spatial, and object representations.* Technical Report CAS/CNS-TR-93-038, Boston University, Boston, MA.

Gurnsey, R., & Browse, R. A. (1987). Micropattern properties and presentation conditions influencing visual texture discrimination. *Perception & Psychophysics, 41*, 239-252.

Gurnsey, R., & Browse, R. A. (1989). Asymmetries in visual texture discrimination. *Spatial Vision, 4*, 31-44.

Gurnsey, R., Humphrey, G.K., & Kapitan, P. (1992). Parallel discrimination of subjective contours defined by offset gratings. *Perception & Psychophysics, 52*, 263-276.

Haenny, P. E., & Schiller, P. H. (1988). State dependent activity in monkey visual cortex. *Experimental Brain Research, 69*, 225-259.

He, Z.J., Nakayama, K., & Tumosa, N. (1991). Surface representation vs. features in visual search. *Society for Neuroscience Abstracts, 17,* 9.

He, Z.J., and Nakayama, K. (1992). Surfaces versus features in visual search. *Nature, 359*, 231-233.

Helmholtz, H. von. (1867/1967). *Treatise on physiological optics* (Vol. 3). In J. P. C. Southall (Ed. and Trans.). NY: Dover.

von der Heydt, R., & Peterhans, E. (1989). Cortical contour mechanisms and geometrical illusions. In D.M. Lam & C.D. Gilbert (Eds.), *Neural mechanisms of visual perception.* Woodlands, TX: Portfolio Publishing.

Horn, B. K. P. (1986). *Robot vision.* Cambridge: MIT Press.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology (London), 160*, 106-154.

Hubel, D. H., & Wiesel, T. N. (1977). Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London (B), 198,* 159.

Humphreys, G. W., & Quinlan, P. T., & Riddoch, M. J. (1989). Grouping processes in visual search: Effects with single- and combined-feature targets. *Journal of Experimental Psychology: General, 118,* 258-279.

Johansson, G. (1975). Visual motion perception. *Scientific American, 232* (June): 76-88.

Julesz, B. (1981). Textons, the elements of texture perception and their interactions. *Nature, 290*, 91-97.

Julesz, B. (1984). A brief outline of the texton theory of human vision. *Trends in Neuroscience, 7*, 41-45.

Julesz, B. (1986). Texton gradients: The texton theory revisited. *Biological Cybernetics, 54,* 245-261.

Julesz, B. (1991). Early vision and focal attention. *Reviews of Modern Physics, 63,* 735-772.

Julesz, B. & Bergen, J. R. (1983). Textons, the fundamental elements in preattentive vision and perception of textures. *Bell System Technical Journal, 62,* 1619-1645.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology, 23,* 141-221.

Kellman, P. J., & Shipley, T. F. (1992). Perceiving objects across gaps in space and time. *Current Directions in Psychological Science, 1*, 193-199.

Kelly, D. H., & Burbeck, C.A. (1984). Critical problems in spatial vision. *CRC Critical Reviews in Biomedical Engineering, 10*, 125-177.

Kleffner, D.A., & Ramachandran, V.S. (1992). On the perception of shape from shading. *Perception & Psychophysics, 52*, 18-36.

Klein, R., & Farrell, M. (1989). Search performance without eye movements. *Perception & Psychophysics, 46*, 476-482.

Liberman, A. M., & Mattingly, I. G. (1989). A specialization for speech perception. *Science, 243*, 489-494.

Lowe, D. G. (1987). Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence, 31*, 355-395.

Mach, E. (1886/1959). *The analysis of sensations and the relation of the physical to the psychical.* New York: Dover.

Mack, A., Tang, B., Tuma, R., Kahn, S., & Rock, I. (1992). Perceptual organization and attention. *Cognitive Psychology, 24*, 475-501.

Malik, J., & Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America A, 2*, 923-932.

Marr, D. (1976). Early processing of visual information. *Philosophical Transactions of the Royal Society of London, B, 275*, 483-524.

Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229*, 782-784.

Morgan, M. J., Hole, G. J., & Glennerster, A. (1990). Biases and sensitivities in geometrical illusions. *Vision Research, 30*, 1793-1810

Nakayama, K., Shimojo, S., & Silverman, G.H. (1989). Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception, 18*, 55-68.

Nakayama, K., & Silverman, G.H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature, 320,* 264-265.

Neisser, U. (1967). *Cognitive psychology*. Englewood Cliffs, NJ: Prentice Hall.

O'Connell, K. M., & Treisman, A. (in preparation). *Shared orientation coding for lines, dot pairs, and edges.* (University of California, Berkeley).

Palmer, S., & Rock, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin & Review, 1,* 29-55.

Peterhans, E., & von der Heydt, R. (1989). The whole and the pieces — cortical neuron responses to bars and rows of moving dots. In J.J. Kulikowski, C.M. Dickinson, & I.J. Murray (Eds.), *Seeing contour and colour* (pp. 125-130) Oxford: Pergamon Press.

Peterhans, E., & von der Heydt, R. (1991). Elements of form perception in monkey prestriate cortex. In A. Gorea (Ed.), *Representations of vision* (pp. 111-124) Cambridge: Cambridge University Press.

Petry, S., & Meyer, G.L. (1987). *The perception of illusory contours.* New York: Springer.

Pomerantz, J.R. (1978). Are complex visual features derived from simple ones? In E. Leeuwenberg and H. Buffart, (Eds.), *Formal theories of visual perception* (pp. 217-229). New York: Wiley.

Pomerantz, J.R. (1981). Perceptual organization in information processing. In M. Kubovy and J.R. Pomerantz, (Eds.), *Perceptual organization* (pp. 141-180). Hillsdale, NJ: Erlbaum.

Pomerantz, J.R., & Kubovy, M. (1986). Theoretical approaches to perceptual organization. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance, vol. II* (ch. 36). New York: John Wiley & Sons.

Pomerantz, J.R. & Pristach, E.A. (1989). Emergent features, attention, and perceptual glue in visual form perception. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 635-649.

Prinzmetal, W. (1981). Principles of features integration in visual perception. *Perception & Psychophysics, 30*, 330-340.

Ramachandran, V. S. (1988). Perceiving shape from shading. *Scientific American, 259* (August), 76-83.

Rensink, R.A. (1987). On the discrimination of self-similar random textures. *Proceedings of the IEEE Society Workshop on Computer Vision,* Miami Beach FL, Nov. 30 - Dec. 2 1987. pp. 240-242.

Rensink, R. A. (1992). *The rapid recovery of three-dimensional orientation from line drawings.* Ph.D. Thesis (also Technical Report 92-25), Department of Computer Science, University of British Columbia, Vancouver, BC, Canada.

Richards, W. & Jepson, A. (1992). What makes a good feature? (Memo No. 1356). Cambridge MA: Artificial Intelligence Laboratory. Massachusetts Institute of Technology.

Robson, J.G. (1980). Neural images: The physiological basis of spatial vision. In C.S. Harris (Ed.), *Visual coding and adaptability* (pp. 177-214). Hillsdale, NJ: Erlbaum.

Robson, J.G., & Graham, N. (1981). Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research, 21,* 409-418.

Rock, I., & Brosgole, L. (1964). Grouping based on phenomenal proximity. *Journal of Experimental Psychology, 67,* 531-538.

Rock, I., & Palmer, S. (1990). The legacy of Gestalt psychology. *Scientific American, 263* (December), 84-90.

Sagi, D., & Julesz, B. (1987). Short-range limitation on detection of feature differences. *Spatial Vision, 2*, 39-49.

Schiller, P. (1986). The central visual system. *Vision Research, 26,* 1351-1386.

Selfridge, O. G., & Neisser, U. (1960). Pattern recognition by machine. *Scientific American, 203* (August), 60-68.

Tausch, R. (1954). Optische Täuschungen als artifizielle Effekte der Gestaltungsprozesse von Grössen- und Formenkonstanz in der natürlichen Raumwahrnehmung. *Psychologische Forschung, 24,* 299-348.

Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception & Performance, 8*, 194-214.

Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology, 40A*, 201-237.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12,* 97-136.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review, 95,* 15-48.

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception & Performance, 16,* 459-478.

Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General, 114*, 285-310.

Treisman, A., Cavanagh, P., Fischer, B., Ramachandran, V. S., & von der Heydt, R. (1990). Form perception and attention: Striate cortex and beyond. In L. Spillman & J. S. Werner (Eds.), *Visual perception* (pp. 273-316). New York: Academic.

Tsal, Y. (1984). A Mueller-Lyer illusion induced by selective attention. *Quarterly Journal of Experimental Psychology: A. Human Experimental Psychology, 36A,* 319-333.

Tsotsos, J. (1990). Analyzing vision at the complexity level. *Behavioral and Brain Sciences, 13*, 423-469.

Voorhees, H., & Poggio, T. (1988). Computing texture boundaries from images. *Nature, 333*, 364-367.

Walters, D. (1987). Selection of image primitives for general-purpose visual processing. *Computer Vision, Graphics, and Image Processing, 37,* 261-298.

Watson, A.B. (1987). *Estimation of local spatial scale. Journal of the Optical Society of America A, 4,* 1579-1582.

Watt, R. J. (1988). *Visual processing*. Hillsdale, NJ: Erlbaum.

Wilson, H. R., & Gelb, D. J. (1984). Modified line element theory for spatial frequency and width discrimination. *Journal of the Optical Society of America A, 1*, 124-131.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1988). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception & Performance, 15*, 419-433.

Wolfe, J. M. (1992a). The parallel guidance of visual attention. *Current Directions in Psychological Science, 1,* 124-128.

Wolfe, J. M. (1992b). "Effortless" texture segmentation and "parallel" visual search are not the same thing. *Vision Research, 32,* 757-763.

Yager, D., Kramer, P., Shaw, M., & Graham, N. (1984). Detection and identification of spatial frequency: Models and data. *Vision Research, 24,* 1021-1035.

Zeki, S. (1992). The visual image in mind and brain. *Scientific American, 267* (September), 68-76.

Zucker, S. W. (1987a). The diversity of perceptual grouping. In M. A. Arbib & A. R. Hanson (Eds.), *Vision, brain, and co-operative computation* (pp. 231-261). Cambridge, MA: MIT Press.

Zucker, S. W. (1987b). Early vision. In S. C. Shapiro (Ed.), *The encyclopedia of artificial intelligence* (pp. 1131-1152). New York: John Wiley