Attention, Perception, & Psychophysics
2009, 71 (7), 1469-1477
doi:10.3758/APP.71.7.1469

Research Articles

Coherent plaids are preattentively more than the sum of their parts

Jong-Ho Nam
Catholic University of Korea, Puchon City, Korea

Joshua A. Solomon and Michael J. Morgan
City University, London, England

and

Charles E. Wright and Charles Chubb
University of California, Irvine, California

We investigated whether plaids activate preattentive mechanisms that are distinct from those activated by their component gratings. Observers searched for a target plaid, the sum of two perpendicular components in a circular window (radius = 0.65º). The target was present on half the trials. On all trials, half of the distractors had the same frequency and orientation as one component of the plaid, and the rest were the same as the other component. The target and the distractors were arrayed evenly on a circle (radius = 2.36º) around fixation. Target and distractor contrasts were randomly perturbed up to ±30%. The following results held for each of the 6 participants tested. (1) When $F1 = 2$ c/deg and $F2 = 5.25$ c/deg, response times (RTs) increased significantly when set size (number of distractors plus target, if present) was increased from four to eight. (2) When the spatial frequencies of both plaid components were the same (i.e., both 2 c/deg or both 5.25 c/deg), RTs increased very slightly, if at all, when set size was increased from four to eight. These results suggest the existence of a preattentive, plaid-sensitive mechanism with band-limited input that does not respond to individual grating components.

The research reported here was based on several assumptions about low-level visual processing, all of which are common to the standard model of visual texture perception (e.g., Bergen & Landy, 1991). First, we follow Adelson and Bergen (1991) in assuming that, before extracting such things as object boundaries, identities, and locations from the retinal input, the visual system first applies a battery of fast, spatially parallel image transformations whose response images reflect "the amounts of various kinds of visual 'substances' present in the image" (p. 3). It is useful to imagine these substance-sensing transformations as being implemented in retinotopically organized neural arrays. In a given array, all neurons are assumed to perform the same computation on the visual input impinging on a small region of the retina but at different locations in the visual field. Thus, any such substance-sensing array operates like a movie camera to make a neural image available to higher level vision (Robson, 1980), reflecting the rapidly changing distribution of a particular substance across the visual field. Because these hypothetical arrays are assumed to operate automatically and prior to any conscious effort on the part of the participant, we call them *preattentive mechanisms*. Each preattentive mechanism requires a dedicated array of neurons, all of which are continuously active in normal vision. The high expense in neural computational resources of a single such mechanism makes it likely that human vision has only a modest number of them, raising the prospect that we may be able to catalog them. Thus, two of the most compelling open questions in the field of low-level vision are as follows: How many preattentive mechanisms exist in human vision, and what properties of the visual input do these mechanisms sense? Or, more grandly, what are the elementary substances of human vision?

The picture sketched above echoes the original theory offered by Treisman and Gelade (1980) to account for instances of parallel versus serial search. They proposed that, when search is parallel (i.e., when a target pops out from a field of distractors irrespective of how many there are), it is because there exists in human vision one or more preattentive mechanisms (of the sort hypothesized above) activated by the target but not by the distractors. Wolfe and Horowitz (2004) have argued that this picture is too simple. They pointed out that there exist certain cases of

C. Chubb, c.chubb@uci.edu

parallel search that seem to require more complicated preprocessing than we imagine as being achieved by a preattentive mechanism. In particular, search is parallel for an oriented bar among bars perpendicular to the target bar, even when all bars in the scene are largely occluded by amorphous blobs (one for each bar) so that the small bits of bar that remain revealed offer no cues of the sort that might drive a standard, orientation-selective preattentive mechanism. The impression is compelling that the features driving the parallel search in these displays are the orientations of the *occluded* bars, cues that can be derived only through segmentation of the scene and analysis of the spatial relations between segmented regions. Preattentive mechanisms, as they were originally conceived by Treisman and colleagues and are posited by standard models of rapid texture segmentation, were assumed to operate in a bottom-up fashion prior to any sort of segmentation such as would be required to construct the occluded bars that support parallel search in these displays.

The parallel search afforded by the displays of Wolfe and Horowitz (2004) thus implies the existence of what might be called second-order preattentive mechanisms, which take as their input a signal stream that has already been transformed by first-order preattentive mechanisms to embody occlusory relations implicit in the visual input. However, it should be noted that all of the processing required to derive the orientations of the occluded bars in the examples demonstrated by Wolfe and Horowitz is generally assumed to take place preattentively. Thus, although these displays stretch our sense of what properties can be sensed by preattentive mechanisms, they do not overthrow the idea that texture discrimination, visual search, and any other tasks requiring rapid analysis of qualitative differences across visual space are mediated by a fixed battery of spatially parallel visual transformations that we continue to call preattentive mechanisms.

The present study focused on the kinds of textures commonly called *plaids*,[1] which have been studied mainly in the domain of motion perception (Adelson & Movshon, 1982). As has often been noted, a plaid composed of gratings with similar spatial frequencies yields a percept of coherent, rigid motion with velocity roughly equal to that with which it is translated; however, a plaid composed of gratings that differ in spatial frequency by more than an octave yields a strikingly different percept: The plaid fails to cohere, and one sees the two components moving separately in directions perpendicular to their bars, each component evoking the same percept as it would if it were presented alone (e.g., Hupé & Rubin, 2003; Krauskopf & Farell, 1990). Thus, when it moves, a plaid is perceived as being more than the sum of its moving parts if and only if its component gratings are at least roughly matched in spatial frequency.

We speculated that the motion percepts evoked by plaids might be rooted in the architecture of human preattentive vision. Perhaps when a plaid is seen as coherent, it is because it activates a preattentive mechanism that is not activated by either of its component gratings. We refer to such hypothetical preattentive mechanisms as *plaid grabbers*. The percepts evoked by moving plaids suggest that, if plaid grabbers exist, they should exist for only those plaids whose components are matched (at least roughly) in spatial frequency. The point of the present experiments was to test whether plaid grabbers exist and, if so, whether they conform to this prediction.

Most of the prominent models of preattentive visual processing offer no reason to think plaid grabbers exist in human vision. First, the literature in visual search does not suggest their existence. Plaids are not on the usual lists of basic features in visual search (for a recent list, see Wolfe & Horowitz, 2004). Moreover, standard "back-pocket" models of texture segregation (Chubb & Landy, 1991) typically hypothesize multiple mechanisms (corresponding to different classes of complex cells), each sensitive to energy in a particular orientation and spatial-frequency band (e.g., Bergen & Adelson, 1988; Bergen & Landy, 1991; Bovik, Clark, & Geisler, 1990; Caelli, 1985; Fogel & Sagi, 1989; Graham, 1991; Knutsson & Granlund, 1983; Landy & Bergen, 1991; Malik & Perona, 1990). However, any such mechanism that responds to a plaid also responds to at least one component of the plaid.

There are, however, reasons to suspect that human vision may incorporate plaid grabbers. Texton theory (Julesz, 1981; Julesz & Bergen, 1983) proposed that human vision is preattentively sensitive to (among other things) "crossings," points of intersection between line segments in the visual input. As those two studies showed, it is possible to produce pairs of textures with identical space–average energy spectra that nonetheless segregate preattentively due to intertexture differences in the density of line segment crossings. Whatever mechanism is selectively sensitive to crossings may well also operate as a plaid grabber. A plaid is, after all, a pattern in which the bars of two sinusoids cross over each other.

More recently, Barth, Zetzsche, and Rentschler (1998) have offered a reinterpretation of texton theory in which they emphasized that human vision may comprise preattentive mechanisms specifically sensitive to the local Gaussian curvature of the intensity map. Imagine a surface covering the visual field whose height at a given point is equal to the light intensity at a point. For any point p in the visual field, the Gaussian curvature of the surface at p is the product of the maximum and minimum curvatures at p taken across all curves through p that lie in the surface. Thus, for example, singular intensity peaks (and valleys) have positive Gaussian curvature. Saddle points have negative Gaussian curvature. Importantly, however, any point p that lies in a region that is locally one-dimensional (i.e., that comprises a pattern of bars all of a fixed orientation) has a Gaussian curvature of 0. We follow Barth et al. (1998) in calling any preattentive mechanism an i2D sensor (intrinsic two-dimensionality sensor) if the mechanism in question is sensitive to rectified (e.g., squared) Gaussian curvature in the visual intensity map.

By itself, a grating is one-dimensional. That is, there exists a curve through any point along which intensity is constant (the curve that runs parallel to the bars of the grating); this means that the Gaussian curvature of the grating is 0 at every point.

On the other hand, the Gaussian curvature is nonzero at all locations in a plaid other than those at which one of

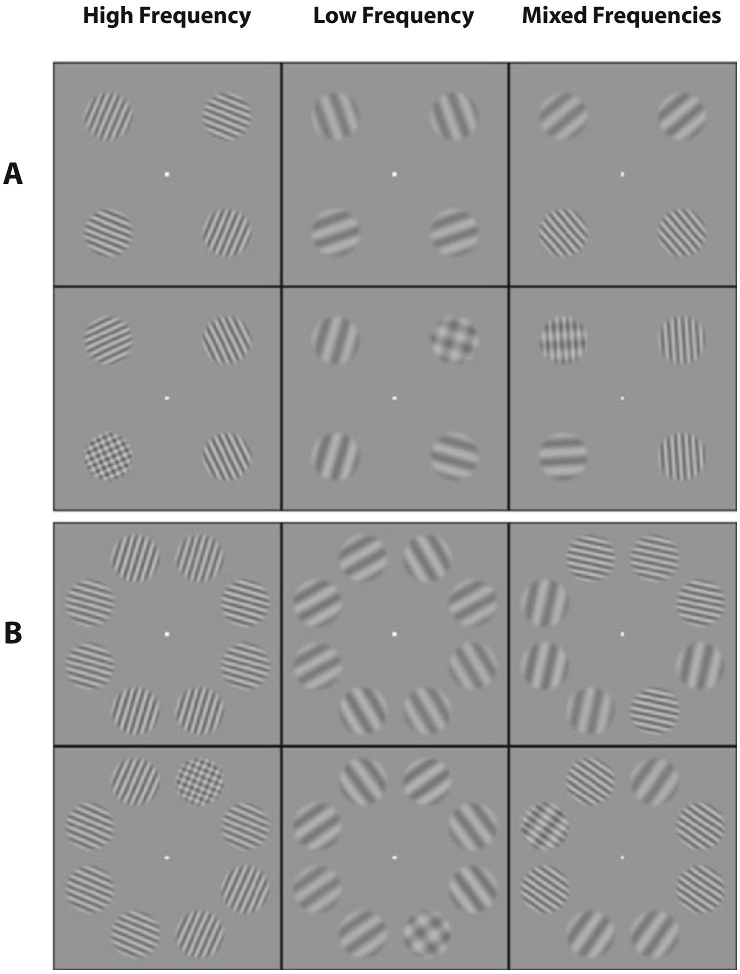

Figure 1. Stimuli from the search task. (A) Examples of stimuli from the high-4, low-4, and mixed-4 conditions. (B) Examples of stimuli from the high-8, low-8, and mixed-8 conditions. In both (A) and (B), the top row gives examples of stimuli on target-absent trials, and the bottom row gives examples of stimuli on target-present trials.

the components takes the value 0 (because curvature goes to 0 at the zero crossings of the components). The Gaussian curvature achieves maxima at the contrast maxima and minima of the plaid (where a peak of one component adds to a peak of the other or a trough of one component adds to a trough of the other) and minima at the saddle points of the plaid (where a peak of one component adds to a trough of the other).

Thus, if human vision comprises a preattentive mechanism specifically tuned to rectified intrinsic two-dimensionality, it must be highly activated by plaids but not by gratings; this means that any i2D sensor must be an excellent plaid grabber. It follows that, if plaid grabbers do not exist in human vision, neither do i2D sensors.

In the present study, we used a search task to test first whether human vision has plaid grabbers. Specifically, we required observers to search for a target plaid among an array of distractors, where each distractor was one of the two components of the plaid. As we show, search was highly efficient (slope of search time vs. number of distractors was near zero) if the two plaid components were equal in spatial frequency (either both high or both low). This implies that human vision does contain plaid grabbers, consistent with the hypothesis that human vision comprises i2D sensors. Moreover, if the two components of the plaid differ sufficiently in spatial frequency, search efficiency drops precipitously. Thus, as we shall show, plaid grabbers exist only for the narrowband plaids that tend to evoke percepts of coherent motion when translated. As we will argue in the Discussion section, this result implies that, if the plaid grabbers in human vision are i2D sensors, they extract Gaussian curvature, not from the broadband visual input, but from band-pass filtered versions of the input.

METHOD

Participants

Three of the authors (Observers 1, 2, and 6) and 3 naive observers participated in the experiment. All had normal or corrected-to-normal visual acuity.

Apparatus

Stimuli were generated by Dell OptiPlex GX1 and displayed on a Dell monitor (17-in., UltraScan 1000HS Series, D1025). Experimental procedure was controlled by MATLAB 7.1 and Psychtoolbox 2.54 (Brainard, 1997). A lookup table was used to ensure linearity. Mean luminance was 51 cd/m^2.

Task

On each trial, the participant searched for a circular plaid patch among distractor grating patches. Each target and distractor patch had a radius of 0.65º and was windowed at the edge by a raised cosine function. The target was present on half the trials. Target and distractors were arrayed evenly on a circle (radius 2.36º) around fixation. On each trial, there were two types of distractors, one randomly oriented and the other perpendicular to the first: On target-absent trials, half of the $n = 4$ or $n = 8$ distractors were one type, and half were the other; on target-present trials with $n = 3$ or $n = 7$ distractors, $(n + 1)/2$ were one of these distractor types and $(n - 1)/2$ were the other. If the target was present, it was the sum of the two distractor types, and the amplitudes of the two plaid components were adjusted (by dividing them by $\sqrt{2}$), so that the expectation of the total contrast energy in the target was equal to the expectation of the contrast energy in a given distractor.

In separate blocks, we tested six conditions (high-4, low-4, mixed-4, high-8, low-8, and mixed-8); stimuli from each are shown in Figure 1. The gratings used in the high-4 and high-8 conditions were 5.25 c/deg, and those used in the low-4 and low-8 conditions were 2 c/deg. Both sorts of gratings were used in the mixed-4 and mixed-8 conditions. Target and distractor Michelson contrasts were uniformly distributed between .175 and .325 for low-frequency gratings ($M = .25$) and uniformly distributed between .259 and .481 for high-frequency gratings ($M = .37$). (The rationale for this manipulation is explained in the Discussion section.) The mean Michelson contrast levels of the high- and low-frequency gratings were roughly matched for apparent contrast.

Before a trial, the participant faced a blank monitor screen of mean luminance and then pressed a button to start the trial. The buttonpress produced a cue spot slightly brighter than the background. This cue spot pulsed off and then on twice more; that is, after its initial 400-msec display, it went off for 400 msec, on again for 400 msec, off again for 400 msec, and then remained on for the duration of the trial. The onset of the stimulus patches that were presented (in a circular array around the cue spot) occurred a random amount of time after the final onset of the cue. This delay was distributed exponentially with a mean of 400 msec (truncated at 1 sec). The participant was instructed to then respond as quickly as possible, pressing a "1" on the keyboard if the stimulus was present and a "0" if it was not.

RESULTS

Results averaged across all 6 participants are shown in Figure 2. Individual results for all 6 observers are given in Tables 1A, 1B, and 1C, where each cell gives the 95% confidence interval for the indicated reaction time (RT). Note

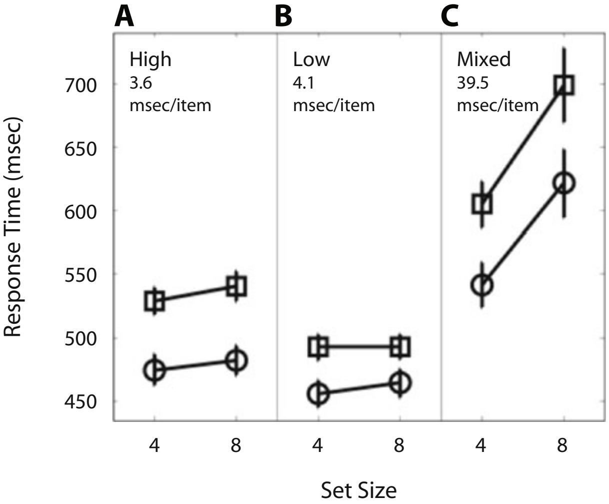

Figure 2. Results averaged across 6 observers: Response times as a function of set size when the target plaid comprised two high-spatial-frequency gratings (A), two low-spatial-frequency gratings (B), and one high- and one low-spatial-frequency grating (C). Note that increases in search time per item were not significantly different from zero when the spatial frequencies of plaid components were identical; however, increases in search time per item were elevated when the target plaid mixed different spatial-frequency gratings. Squares give the results for target-absent trials; circles, for target-present trials.

that, when the target was a plaid whose components were equal in spatial frequency, RT was nearly the same regardless of set size. By contrast, when the target was a plaid whose components differed in spatial frequency, RTs were significantly longer when set size was eight than when it was only four. Specifically, on average, search in the mixed-spatial-frequency condition required 39.5 msec per item (as compared with 3.6 msec/item for the high-frequency plaid search and 4.1 msec/item for the low-frequency plaid search, neither of the latter estimates being significantly different from zero (as shown by the confidence intervals in Figure 3). In this figure and elsewhere in the article, increases in search time per item are estimated by taking the average of $(MRT_8 - MRT_4)/2$, where MRT_k is the mean RT observed in the target-present condition with set size k. The reasoning behind this statistic is as follows: On average, when set size is eight, the participant must check four items to find the target, as compared with two items when set size is four. Thus, we expect the participant to have to check two more items (on average) when set size is eight than when it is four. Error rates for all observers in all conditions (including two control conditions described below) are given in Table 2. Error rates were less than 10% for all observers in all conditions.

Individual participants all showed the same general pattern of results. The plot in Figure 3 shows the increase in search time per item for each participant in each of the three conditions. The single number on the left side of each participant's plot gives the increase in search time per item for the mixed-spatial-frequency target. Error bars give 95% confidence intervals.

Table 1A
Search Times (in Milliseconds) for High-Spatial-Frequency Plaid Amid High-Spatial-Frequency Distractors

	Set Size 4 (95% Confidence Intervals)		Set Size 8 (95% Confidence Intervals)		Search Time per Item
Participant	Target Present (MRT_4)	Target Absent	Target Present (MRT_8)	Target Absent	$(MRT_8 - MRT_4)/2$
1	431 ±7.0	497 ±7.4	436 ±7.2	488 ±6.5	2.7 ±5.0
2	605 ±17.8	689 ±10.2	616 ±13.0	718 ±15.3	5.4 ±11.0
3	443 ±10.3	468 ±11.9	443 ±9.5	478 ±9.1	0.3 ±7.0
4	453 ±9.1	481 ±9.6	473 ±10.2	513 ±15.4	10.1 ±6.8
5	546 ±15.4	604 ±15.6	541 ±17.7	629 ±14.2	−2.5 ±11.7
6	371 ±11.5	435 ±10.0	382 ±8.4	417 ±6.9	5.6 ±7.1

Table 1B
Search Times (in Milliseconds) for Low-Spatial-Frequency Plaid Amid Low-Spatial-Frequency Distractors

	Set Size 4 (95% Confidence Intervals)		Set Size 8 (95% Confidence Intervals)		Search Time per Item
Participant	Target Present (MRT_4)	Target Absent	Target Present (MRT_8)	Target Absent	$(MRT_8 - MRT_4)/2$
1	428 ±7.3	480 ±6.9	432 ±6.1	485 ±7.0	2.0 ±4.7
2	557 ±14.0	610 ±9.1	581 ±13.2	606 ±7.9	11.9 ±9.6
3	439 ±9.5	447 ±12.0	436 ±11.2	461 ±9.3	−1.6 ±7.3
4	439 ±8.5	448 ±11.1	451 ±11.3	452 ±9.5	5.9 ±7.0
5	500 ±12.2	567 ±11.5	515 ±16.4	564 ±14.4	7.3 ± 10.2
6	374 ±12.7	404 ±8.9	372 ±8.3	389 ±7.2	−0.9 ±7.6

Table 1C
Search Times (in Milliseconds) for Mixed-Spatial-Frequency Plaid Amid Mixed-Spatial-Frequency Distractors

	Set Size 4 (95% Confidence Intervals)		Set Size 8 (95% Confidence Intervals)		Search Time per Item
Participant	Target Present (MRT_4)	Target Absent	Target Present (MRT_8)	Target Absent	$(MRT_8 - MRT_4)/2$
1	414 ±10.7	472 ±10.1	462 ±13.1	527 ±15.6	24.0 ±7.7
2	768 ±26.2	806 ±22.1	878 ±46.0	980 ±53.3	55.0 ±26.5
3	502 ±15.2	529 ±16.5	582 ±24.3	610 ±21.3	39.9 ±14.3
4	538 ±16.6	587 ±15.2	591 ±17.2	711 ±24.4	26.4 ±12.0
5	561 ±21.3	722 ±29.2	684 ±29.3	767 ±27.5	61.3 ±18.1
6	414 ±10.7	472 ±10.1	462 ±13.1	527 ±15.6	23.5 ±8.5

The main point to note is that, for all participants, search was highly efficient for plaids as long as their component gratings had the same spatial frequency; indeed, in nearly all such searches (excepting the high-frequency plaid search for Observer 4 and the low-spatial-frequency plaid search for Observer 2), the increase in search time per item did not differ significantly from 0. By contrast, for all participants, increase in search time per item for the mixed-spatial-frequency plaid was significantly greater than 0.

DISCUSSION

We would like to conclude that human vision is equipped with plaid grabbers for plaids whose component gratings have the same spatial frequency but not for plaids whose components differ sufficiently in spatial frequency. However, three objections might be raised.

First, note that in the conditions using plaids of only one spatial frequency the individual component gratings had lower amplitudes on average (by a factor of $1/\sqrt{2}$) than did the distractors. It could be, then, that the participants were using ordinary orientation-tuned channels to perform the search. They could simply have identified targets as those items activating a particular orientation-selective channel less than did the other items activating the same channel. There are two reasons to think that this was not the case. First, as was described in the Method section, the amplitude of each distractor (and also of the target) was subject to substantial random perturbation (±30%). A consequence of this manipulation was that on more than half of all target-present trials, the amplitude of each individual target component could have been drawn from the distribution of distractor components. However, it is nonetheless true that the amplitudes of individual target components tended to be lower than the amplitudes of distractors; thus, the proposed strategy might have provided some traction. The crucial reason to think that it was not used is that it was also available in the task in which plaids contained components of different spatial frequencies; obviously, as the results make clear, it did not help in this case. We conclude that participants did not solve the search problem for same-spatial-frequency plaids using a single, oriented channel.

Second, Olzak and Thomas (1991, 1992, 1999) and Thomas and Olzak (1996, 1997) have presented strong empirical evidence for the existence of preattentive mechanisms that they call *doughnuts* (because of the annular shape of their spatial-frequency sensitivity functions) that pool energy across different orientations within a given spatial-frequency band. How can we be sure that the search for the same-frequency plaids is not accomplished by using such mechanisms? Our original reason for dividing the amplitudes of plaid components by $\sqrt{2}$, relative to distractor amplitudes, was so that the total energy in the target would be equal to the total energy in a given distractor. In theory, this should ensure that the target in the single-spatial-frequency plaid conditions cannot be discriminated from the distractors by a doughnut mechanism; however, it is possible that doughnuts do not actually integrate energy across different orientations in precisely the way we have assumed. Indeed, available evidence suggests that pooling performed by doughnuts across different orientations may be linear in component amplitudes (Thomas & Olzak, 1997). If this is so, our same spatial-frequency plaid targets may activate doughnut mechanisms systematically more on average than do distractors.

However, this activation difference is not large enough to enable doughnuts to be used to detect same-spatial-frequency plaid targets in our displays. This is ensured by the $\pm 30\%$ random amplitude perturbations injected into targets and distractors. Even if same-spatial-frequency target plaids were producing a level of activation in doughnut mechanisms equal to the sum of the amplitudes of their individual components (which would, on average, elevate target doughnut activation by a factor of $\sqrt{2}$ above mean distractor doughnut activation), this boost would not be sufficient to support the levels of performance observed. In particular, under these conditions, the (optimal) strategy of saying a target was present only when an element produced doughnut activation greater than the maximum possible distractor activation would yield 63% correct.[2] As is shown in Table 2, error rates for all observers in all conditions are less than 10%.

But perhaps (one might suggest) rather than using a strictly parallel doughnut-activation strategy, observers used doughnut activation to guide their search for same-spatial-frequency plaids. In other words, perhaps they first checked the display element that produced maximal doughnut activation and, if it was not a plaid, checked the element producing the next highest level of activation, until all possible targets had been exhausted. In this case, simulations reveal that displays of set size 8 should require the observer to check 35% more items than should displays of set size 4, leading (contrary to our findings) to longer search times for displays of set size 8 than for those of set size 4. We conclude that performance in searches for same-spatial-frequency plaids is not mediated by "doughnuts."

A third possible objection springs from the observation that search times per item increase as the distractor set increases in heterogeneity (Rosenholtz, 2001a, 2001b). The two types of distractors in the mixed-spatial-frequency

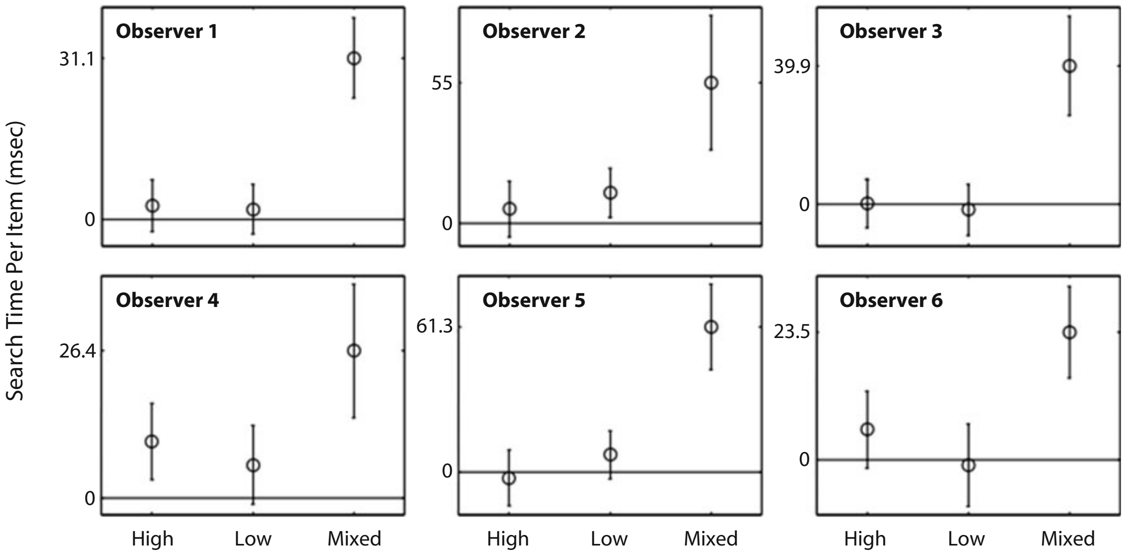

Figure 3. Increases in search times per item for 6 observers. Each panel plots mean increase in search time per item for each of three conditions: high (target plaid comprised two high-frequency gratings), low (target plaid comprised two low-frequency gratings), and mixed (target plaid comprised one low- and one high-frequency grating). In each condition, each distractor was one of the two component gratings used to make the target plaid. Note that for all except Observer 2 in the low condition and Observer 4 in the high condition, increases in search time per item were not significantly different from zero when the spatial frequencies of plaid components were identical, but they were elevated when the target plaid mixed different spatial-frequency gratings.

Table 2
Error Rates (in Percentages) for All Observers in All Conditions

	Observer					
Condition	1	2	3	4	5	6
High T–High D	3	3	2	3	3	4
Low T–Low D	1	4	3	2	3	4
Mixed T–Mixed D	1	8	5	3	4	7
High T–Mixed D		7				3
Low T–Mixed D		2				4

condition differ from each other both in orientation and in spatial frequency, whereas those in the same-spatial-frequency conditions differ only in orientation. Thus, the distractor set in the mixed-spatial-frequency condition is likely to be more heterogeneous than the distractor sets in either of the two same-spatial-frequency conditions. Perhaps this difference in distractor heterogeneity is the sole source of the high increases in search time per item in the mixed- vs. same-spatial-frequency conditions.

To test this possibility, we tested Observers 2 and 6 in two additional control conditions, using mixed high- and low-frequency gratings (perpendicular in orientation) as distractors and same-frequency plaids as targets (high-frequency plaids in one condition and low-frequency plaids in another). Specifically, one type of distractor grating was identical in orientation and spatial frequency to one component of the target plaid, and the other type of distractor grating was perpendicular to the first type and different in spatial frequency (if the target plaid comprised high-frequency gratings, the second type of distractor was low frequency, and vice versa).

The results are plotted in Figure 4 (along with the previous data for these observers from Figure 3). (95% confidence intervals for RTs in all individual conditions are given in Tables 3A and 3B, and error rates are given in the last two rows of Table 2.) For each observer, performance in the condition using a same-frequency (either high or low) plaid as a target and mixed high- and low-frequency gratings as distractors was similar to the performance observed when distractors were all same-spatial-frequency gratings. In all cases, regardless of the distractor set, search was significantly less efficient when the target was a plaid whose components had different spatial frequencies than when the target was a plaid whose components had equal spatial frequencies. This result strengthens the argument in favor of same-spatial-frequency plaid grabbers. Regardless of whether distractors vary in spatial frequency, none of them will activate a plaid grabber; thus, if search for a same-spatial-frequency plaid is mediated by a plaid grabber, the search should be immune to this type of distractor heterogeneity.

We conclude that human vision does indeed embody plaid grabbers—preattentive mechanisms activated by plaids, but not by their component gratings. However, our results suggest that we have plaid grabbers only for plaids whose component gratings have similar spatial frequencies.

This finding implies that if the plaid grabbers in human vision are i2D sensors, then they operate not on the spatially broadband visual input but, rather, on band-pass-filtered input streams. We introduced the notion of an i2D sensor in the introduction. To be clearer about this, we call a hypothetical mechanism an *i2D mechanism* if its response is an increasing function of locally space-averaged rectified Gaussian curvature of the surface defined by the contrast variations in the visual input. For this definition to make sense, the contrast-defined surface must be differentiable. We assume that this is ensured by isotropic spatial filtering of the visual input.

Suppose that this filtering is low-pass (for smoothing) but broadband (so as to filter out neither of the two com-

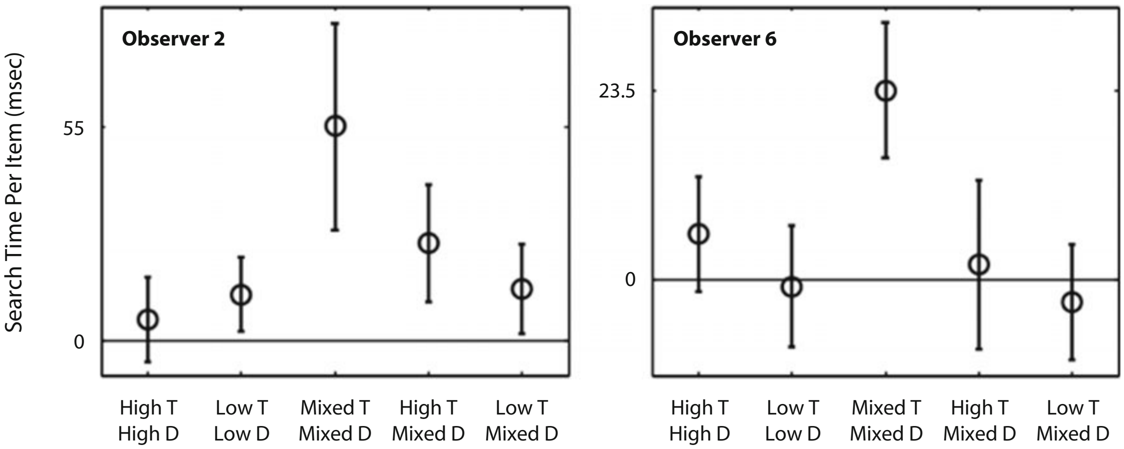

Figure 4. Control experiment results for 2 observers. The first three points in each panel replot results shown in Figure 3. The fourth and fifth points in each panel show increases in search time per item when the target was a plaid, for which both components had the same spatial frequency; whereas the distractors were mixed in spatial frequency. The fourth point gives the results when both target components were high frequency, and the fifth point gives the results when both target components were low frequency. Note that increases in search time per item were much lower than they were when the target contained components of different spatial frequencies. We conclude that the higher increases in search time per item in the mixed T–mixed D condition were not due merely to the greater distractor heterogeneity in this condition.

Table 3A
Search Times (in Milliseconds) for High-Spatial-Frequency Plaid Amid Mixed-Spatial-Frequency Distractors

	Set Size 4 (95% Confidence Intervals)		Set Size 8 (95% Confidence Intervals)		Search Time per Item
Participant	Target Present (MRT_4)	Target Absent	Target Present (MRT_8)	Target Absent	$(MRT_8 - MRT_4)/2$
2	566 ±21.6	661 ±18.4	616 ±20.7	639 ±18.9	25.0 ±15.0
6	425 ±14.2	461 ±9.6	429 ±15.3	461 ±10.6	1.8 ±10.4

Table 3B
Search Times (in Milliseconds) for Low-Spatial-Frequency Plaid Amid Mixed-Spatial-Frequency Distractors

	Set Size 4 (95% Confidence Intervals)		Set Size 8 (95% Confidence Intervals)		Search Time per Item
Participant	Target Present (MRT_4)	Target Absent	Target Present (MRT_8)	Target Absent	$(MRT_8 - MRT_4)/2$
2	526 ±17.2	574 ±13.1	552 ±15.3	625 ±19.9	13.3 ±11.5
6	420 ±10.9	448 ±8.9	414 ±9.5	448 ±8.9	−2.9 ±10.4

ponents in our heterogeneous plaids). The space-average absolute Gaussian curvature of a plaid with perpendicular components of equal amplitude is proportional to the product of the spatial frequencies of the component gratings. In itself, this would seem to suggest that an i2D sensor should respond more strongly to high- than to low-frequency plaids. However, space-average absolute Gaussian curvature is also an increasing function of each of the amplitudes of the gratings in a plaid. If the plaid grabbers in human vision are actually i2D sensors operating on the broadband visual input, then a search for the mixed-spatial-frequency plaid should be comparable in efficiency to a search for the same-spatial-frequency plaids. The reason for this is straightforward: The degree of activation produced in a broadband i2D mechanism by a mixed-frequency plaid must lie between the activations produced in it by the high- and the low-frequency plaid. Our results show clearly that this is not the case: The search for the mixed-frequency plaid is much less efficient than either of the same spatial-frequency plaid searches. We therefore conclude that, if the plaid grabbers in human vision are i2D sensors, they operate not on the spatially broadband visual input but, rather, on band-pass-filtered streams.

Plaids behave analogously in motion and search experiments. In our same-spatial-frequency plaid searches, the targets are plaids that would perceptually cohere if we translated them in space. That is, if one pushed either of these (high- or low-frequency) plaids in a direction diagonal with respect to either of their component orientations, the moving plaid would elicit a distinct and unambiguous percept of translatory motion in the translation direction. By contrast, the plaid used as a target in the mixed-spatial-frequency condition would not cohere. This leads us to conjecture that the mechanism mediating search in the conditions using same-spatial-frequency targets is one and the same as that which produces coherence in moving plaids.

Previous evidence also supports the existence of band-limited mechanisms sensitive to multiple orientations at the same location. Olzak and Thomas (1992) demonstrated that acuity for tilt was better when both components of a plaid rotated in the same direction than when they rotated in opposite directions, but only when both components had the same spatial frequency. This result suggests a band-limited mechanism that can encode the orientation of a plaid.

A simple model encompasses the findings of Olzak and Thomas (1992) and the present results. Our experiments say nothing about the orientation selectivity of plaid grabbers; in particular, we have no reason to think that they are not orientation selective. Thus, for example, a particular plaid grabber might be strongly activated by plaids with vertical and horizontal component gratings, but not by plaids with diagonal components. As long as any given plaid activates at least one such orientation-selective plaid grabber, the band-limited targets in our experiments should pop out from the nonplaid distractors. If individual plaid grabbers are orientation selective, then their ensemble response to a given plaid would be expected to peak at the orientation of the plaid. Under this model, the findings of Olzak and Thomas (1992) suggest that observers can use this peak response to discriminate plaid orientations.

Finally, the present experiments implicate a class of preattentive mechanisms that we have called *plaid grabbers*. This name might be taken to suggest that these mechanisms are tuned specifically to plaids; we emphasize that this need not be the case. All we know about plaid grabbers is that they are activated by plaids comprising perpendicular gratings of the same spatial frequency, but not by their component gratings; however, this bare fact tells us little about what these mechanisms are sensing. Indeed, any property that marks a plaid but neither of its component gratings (e.g., having nearly circular contrast extrema) might be the "visual substance" sensed by a plaid grabber—for example, the orientation-variance of the texture (Morgan, Chubb, & Solomon, 2008). Precise characterization of the visual property sensed by plaid grabbers and the sharpness of their tuning awaits future investigation.

AUTHOR NOTE

This research was supported by Engineering and Physical Sciences Research Council Grant EP/E064604. This article was greatly improved by comments from Sabina Wolfson and Jeremy Wolfe. Address correspondence to C. Chubb, Department of Cognitive Sciences, University of California, Irvine, CA 92697 (e-mail: c.chubb@uci.edu).

REFERENCES

Adelson, E. H., & Bergen, J. R. (1991). The plenoptic function and the elements of early vision. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 3-20). Cambridge, MA: MIT Press.

Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, **300**, 523-525. doi:10.1038/300523a0

Barth, E., Zetzsche, C., & Rentschler, I. (1998). Intrinsic two-dimensional features as textons. *Journal of the Optical Society of America A*, **15**, 1723-1732. doi:10.1364/JOSAA.15.001723

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

Bergen, J. R., & Landy, M. S. (1991). Computational modeling of visual texture segregation. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 253-271). Cambridge, MA: MIT Press.

Bovik, A. C., Clark, M., & Geisler, W. S. (1990). Multichannel texture analysis using localized spatial filters. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **12**, 55-73.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**, 433-436. doi:10.1163/156856897X00357

Caelli, T. (1985). Three processing characteristics of visual texture segmentation. *Spatial Vision*, **1**, 19-30.

Chubb, C., & Landy, M. S. (1991). Orthogonal distribution analysis: A new approach to the study of texture perception. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 291-301). Cambridge, MA: MIT Press.

Fogel, I., & Sagi, D. (1989). Gabor filters as texture discriminator. *Biological Cybernetics*, **61**, 103-113.

Graham, N. (1991). Complex channels, early local nonlinearities, and normalization in texture segregation. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 273-290). Cambridge, MA: MIT Press.

Hupé, J. M., & Rubin, N. (2003). The dynamics of bi-stable alternation in ambiguous motion displays: A fresh look at plaids. *Vision Research*, **43**, 531-548. doi:10.1016/S0042-6989(02)00593-X

Julesz, B. (1981). Textons, the elements of texture perception and their interactions. *Nature*, **290**, 91-97. doi:10.1038/290091a0

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

Knutsson, H., & Granlund, G. H. (1983). Texture analysis using two-dimensional quadrature filters. In *Proceedings of the IEEE Computer Society Workshop on Computer Architecture for Pattern Analysis and Image Database Management* (pp. 206-213). Silver Spring, MD: IEEE Computer Society.

Krauskopf, J., & Farell, B. (1990). Influence of colour on the perception of coherent motion. *Nature*, **348**, 328-331. doi:10.1038/348328a0

Landy, M. S., & Bergen, J. R. (1991). Texture segregation and orientation gradient. *Vision Research*, **31**, 679-691. doi:10.1016/0042-6989(91)90009-T

Malik, J., & Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America A*, **7**, 923-932.

Morgan, M. J., Chubb, C., & Solomon, J. A. (2008). A "dipper" function for texture discrimination based on orientation variance. *Journal of Vision*, **8**(11, Art. 9), 1-8. doi:10.1167/8.11.9

Olzak, L. A., & Thomas, J. P. (1991). When orthogonal orientations are not processed independently. *Vision Research*, **31**, 51-57. doi:10.1016/0042-6989(91)90073-E

Olzak, L. A., & Thomas, J. P. (1992). Configural effects constrain Fourier models of pattern discrimination. *Vision Research*, **32**, 1885-1892. doi:10.1016/0042-6989(92)90049-O

Olzak, L. A., & Thomas, J. P. (1999). Neural recoding in human pattern vision: Model and mechanisms. *Vision Research*, **39**, 231-256. doi:10.1016/S0042-6989(98)00122-9

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.

Rosenholtz, R. (2001a). Search asymmetries? What search asymmetries? *Perception & Psychophysics*, **63**, 476-489.

Rosenholtz, R. (2001b). Visual search for orientation among heterogeneous distractors: Experimental results and implications for signal-detection theory models of search. *Journal of Experimental Psychology: Human Perception & Performance*, **27**, 985-999. doi:10.1037/0096-1523.27.4.985

Thomas, J. P., & Olzak, L. A. (1996). Uncertainty experiments support the roles of second-order mechanisms in spatial frequency and orientation discriminations. *Journal of the Optical Society of America A*, **13**, 689-696.

Thomas, J. P., & Olzak, L. A. (1997). Contrast gain control and fine spatial discriminations. *Journal of the Optical Society of America A*, **14**, 2392-2405. doi:10.1364/JOSAA.14.002392

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, **12**, 97-136. doi:10.1016/0010-0285(80)90005-5

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how they do it. *Nature Reviews Neuroscience*, **5**, 1-7.

NOTES

1. Strictly speaking, these patterns should be called "tartans" not "plaids," since a plaid is a material, not a pattern.

2. The value of 63% arises as follows: Distractor amplitudes are uniformly distributed between $0.7k$ and $1.3k$ (where $k = 0.25$ for low-frequency plaids and 0.37 for high-frequency plaids). Target component amplitudes are uniformly distributed between $0.7k/\sqrt{2}$ and $1.3k/\sqrt{2}$. If component amplitudes add to produce doughnut activation, then targets produce doughnut activations that are uniformly distributed between $\sqrt{2} \times 0.7k$ and $\sqrt{2} \times 1.3k$. This means that the probability that a target will produce doughnut activation outside the range of activations possible for distractors is $(\sqrt{2} \times 1.3k - 1.3k)/[\sqrt{2} \times (1.3k - 0.7k)] = 0.6346$.

(Manuscript received February 24, 2008;
revision accepted for publication April 5, 2009.)