Change Blindness Reveals Distinctive Featural Processing in Size, Color, and Orientation

Stella X. Yu[†] and Dimitri Lisin*

TR-13-001

February 2013

Abstract

Visual search experiments in static displays have long established that size, color, and orientation are elementary features whose attributes are processed in parallel and available to guide the deployment of attention. Using a gaze-tracked flicker paradigm for change blindness and stimuli rendered identically in space and separately in the 3 feature dimensions, we investigate whether and how these features distinguish themselves in the active deployment of attention during prolonged visual search. We find out that visual search does not show any attentional modulation in orientation, whereas it engages spatial attention in color with shorter saccades between the same color, and it engages featural attention in size with shorter fixation from previewing the same size as well as tuning into a particular size. Thus, in terms of dynamic attribute processing over time, size, color, and orientation are highly distinctive: Between successive fixations, only orientation is truly pre-attentive without any form of priming, whereas size and color deploy attention in the featural and spatial domains respectively.

[†]International Computer Science Institute, 1947 Center Street, Suite 600, Berkeley, California, 94704
*Mathworks, Inc.

INTRODUCTION

Size, color, and orientation have long been considered elementary features (Treisman and Gelade, 1980) that are available to guide the deployment of attention and visual search (Wolfe and Horowitz, 2004). Their special status in early visual processing is supported by physiological evidence on how these features could be extracted with separate sets of dedicated detectors working in parallel across the entire space (Maunsell and Newsome, 1987). It is also supported by a large volume of psychophysical evidence on how they can mediate "effortless" texture segregation, recombine in illusory conjunctions, and pop out in feature search more with the presence than the absence of an attribute and despite some distractor heterogeneity (Treisman and Souther, 1985; Wolfe et al., 1989; Treisman, 1998; Sagi, 1988; Moraglia, 1989; Wolfe, 2001; Nagy and Sanchez, 1990; D'Zmura, 1991; Bauer et al., 1996; Bergen and Julesz, 1983; Foster and Ward, 1991; Wolfe et al., 1992; Cavanagh et al., 1990; Wolfe et al., 1999).

While search performance has been instrumental to the study of visual attention, its interpretation is often met with considerable reservations. For example, search is easier for a target defined by a single feature ("vertical") than by a conjunction of features ("red+vertical"), but efficiency is not guaranteed even when a unique attribute defines the target of search. As can be understood from a signal detection point of view, search efficiency increases with target-distractor difference (signal) and decreases with distractor-distractor difference (noise) (Duncan and Humphreys, 1989; Verghese, 2001; Eckstein, 1998; Palmer et al., 2000). Search for a target defined by a unique basic feature can be made arbitrarily difficult if the target-distractor difference can be made arbitrarily small, whereas search for targets not defined by a unique basic feature cannot be made arbitrarily easy by increasing that difference (Nagy and Sanchez, 1990; D'Zmura, 1991; Bauer et al., 1996).

We consider *search behaviours* instead of *search performance* to be more revealing about the active deployment of visual attention, and about fundamental differences in the dynamics of featural processing over time.

We employ change detection in a flicker paradigm as a difficult visual search task to engage active search behaviours from the subject. This paradigm induces the so-called *change blindness* phenomenon, where a large difference between two otherwise similar images becomes strikingly difficult to detect with a blank inbetween, even

with repeated presentations (Rensink et al., 1995). Without the blank, the change elicits a singular motion signal which automatically draws the viewer's attention to the location of change; with the blank, the motion signal associated with the change is disrupted and overwhelmed by those motion transients between either image and the blank, which effectively mask the location of change.

Change blindness has been observed on real scene images under flickers (Rensink et al., 1995), saccades (Grimes, 1996), blinks(O'Regan et al., 2000), movie cuts (Levin and Simons, 1997), real-world interactions (Simons and Levin, 1998), mudsplashes(O'Regan et al., 1999). The wide variety of blindness conditions suggests that our seemingly complete visual impression under normal viewing is achieved not by a spatiotopic buildup of visual details with multiple eye fixations (Feldman, 1985; Trehub, 1991), but by a constant validation of *visual aspects* through a continuous access to the visual scene (Irwin, 1991; Kahneman et al., 1992; Pollatsek and Rayner, 1992; Irwin and Andrews, 1996; Henderson, 1997; O'Regan et al., 2000; Rensink, 2002).

While the majority of change blindness works focus on how high-level knowledge such as tasks and scene context engages visual attention (Rensink et al., 1997), we show that change blindness also provides an excellent testbed for understanding how low-level visual processing engages visual attention (Wolfe et al., 2006).

First, change blindness allows us to decouple the target of feature search from visual saliency in the space. Our target is defined not by one of the feature attributes, whose effect on attention is the focus of our investigation, but by the fact that the attribute is changing over time. At any instance in time, the attributes of our stimuli are uniformly random everywhere, so that the target cannot draw attention to itself, but has to be discovered with active search in the space. Consequently, the search behaviour, such as the influence of what is being looked at on what is to be looked at, becomes more interesting than the performance of final detection, such as detection accuracy or detection time.

Secondly, change blindness allows us to study the effect of attentional cues in a more natural setting. A common experimental setup for studying attention is to have the subject detect a target after seeing a precue which may or may not be helpful for priming the target location or attribute (Eriksen and Yeh, 1985; Hoffman and Subramaniam, 1995; Yeshurun and Carrasco, 2008; Greenwood and Parasuraman, 1999, 2004; Muller et al., 2003; Geyer et al., 2006; Becker and Horstmann, 2009; Anderson and Folk, 2010; Kristjansson and Campana, 2010). The cue effectiveness is then reflected in the task perfor-

mance, e.g. shortened reaction time with an informative cue. Our experiments have no artificial precues. We record the subject's eye movements during the search, and examine whether the previous fixation has any effect on the current fixation. In essence, we regard the item of each fixation as a precue to the next one, with the precue itself carrying information not so much about the target's attribute or location as about the search behaviour itself.

We investigate how size, color, and orientation differ in the dynamic attribute processing over time.

We conduct a gaze-tracked change blindness experiment using stimuli rendered identically in the space and separately in size, color, and orientation dimensions. We first establish in Experiment 1 that the target defined by the flickering change between two attributes is equally difficult to detect among the 3 features when spatial search is not needed. We find out in Experiment 2 that, when spatial search is needed, size, color, and orientation are highly distinctive in the search behaviour between successive fixations: There is no priming of any kind for orientation, but there is spatial priming for color and featural priming for size.

In particular, the attribute itself has an impact on search only for size, but not for color or orientation (Farell and Pelli, 1993; Verghese and Pelli, 1994; Zelinsky, 2001; Hollingworth et al., 2001; Anderson and Folk, 2010). Furthermore, there is a preview benefit of shortened fixation which does not diminish over distance, opposite to what the zoom-lens metaphor of visual attention would predict (Eriksen and Yeh, 1985; Eriksen and James, 1986; Castiello and Umilta, 1990; LaBerge et al., 1997; Greenwood and Parasuraman, 2004).

EXPERIMENT 1

The purpose of this experiment is to establish that our target of search in Experiment 2 is comparable among the 3 features: Change detection performance is comparable between small and large, or between black and white, or between horizontal and vertical. The two attributes themselves also have comparable no-change detection performance within each dimension.

Method

Stimuli. There are 6 kinds of disks, with 2 attributes for each of the 3 features (Fig. 1). Size has 2 radii, $0.45°$ for small and $1.35°$ for large. Color has 2 values, 0.3 for black and 0.7 for white on $0 - 1$ value scale, i.e. 77 and 179 for the $0 - 255$ grayscale range. Orientation has 2 angles, $0°$ for horizontal and $90°$ for vertical, with disk radii $0.45° \times 1.35°$ along two directions.

Both size and orientation stimuli are of black value 0.3. Color stimuli are of medium disk radius $0.9°$. The background is of neutral gray value 0.5. A trial for each feature dimension could involve any of the 4 attribute pairs: $(1, 2), (2, 1), (1, 1), (2, 2)$, with a change present in the first two pairs and absent in the second two pairs.

attribute #	size	color	orientation
1			
2			

Figure 1: **Disk stimuli are rendered on a neutral gray background with two attributes for each dimension. Size has two radii,** $0.45°$ **for small and** $1.35°$ **for large. Color has two values,** 0.3 **for black and** 0.7 **for white on** $0 - 1$ **value scale. Orientation has two angles,** $0°$ **for horizontal and** $90°$ **for vertical, with disk radii** $0.45° \times 1.35°$ **along two directions. A trial could involve any of the 4 attribute pairs:** $(1, 2), (2, 1), (1, 1), (2, 2)$**.**

Apparatus. Stimuli are displayed using a Panasonic PT-LB50NTU Projector with 2000 ANSI lumens in image brightness onto a screen mounted at the back of a 4.3m × 8.8m room with no other lighting. The display extends $25.6° \times 34.1°$ at a viewing distance of 5 meters. A 3.2GHz Dell Precision computer controls an eye tracker as well as the stimulus presentation.

Procedure. Each trial begins with a 1-second display of a neutral gray background. A fixation cross of radius $0.5°$ rendered in a random-dot texture pattern is subsequently shown at the center of the display for 500 ms, prompting the subject to gaze at the center. The first disk is shown 1 second after the disappearance of the cross, followed by the blank background and then the second disk, both the disks and the blank presented for 120ms each. A choice screen is then on for 1.5 seconds, during which the subject is required to respond as soon as possible whether the two disks are the same or different by pressing a left or right key in correspondence to the two displayed choices. The next blank background image indicates the start of next trial (Fig. 2).

There are 16 trials for each of the 4 attribute pairs in each of the 3 feature dimensions, resulting in a total of $16 \times 4 \times 3 = 192$ trials. These trials are completely randomized for each subject. The subject is given as many as needed practice trials in the beginning of the experiment to be familiarized with the procedure.

Participants. 15 Boston College students, with normal or corrected-to-normal vision, naive with respect to the experimental hypotheses, participated after provid-
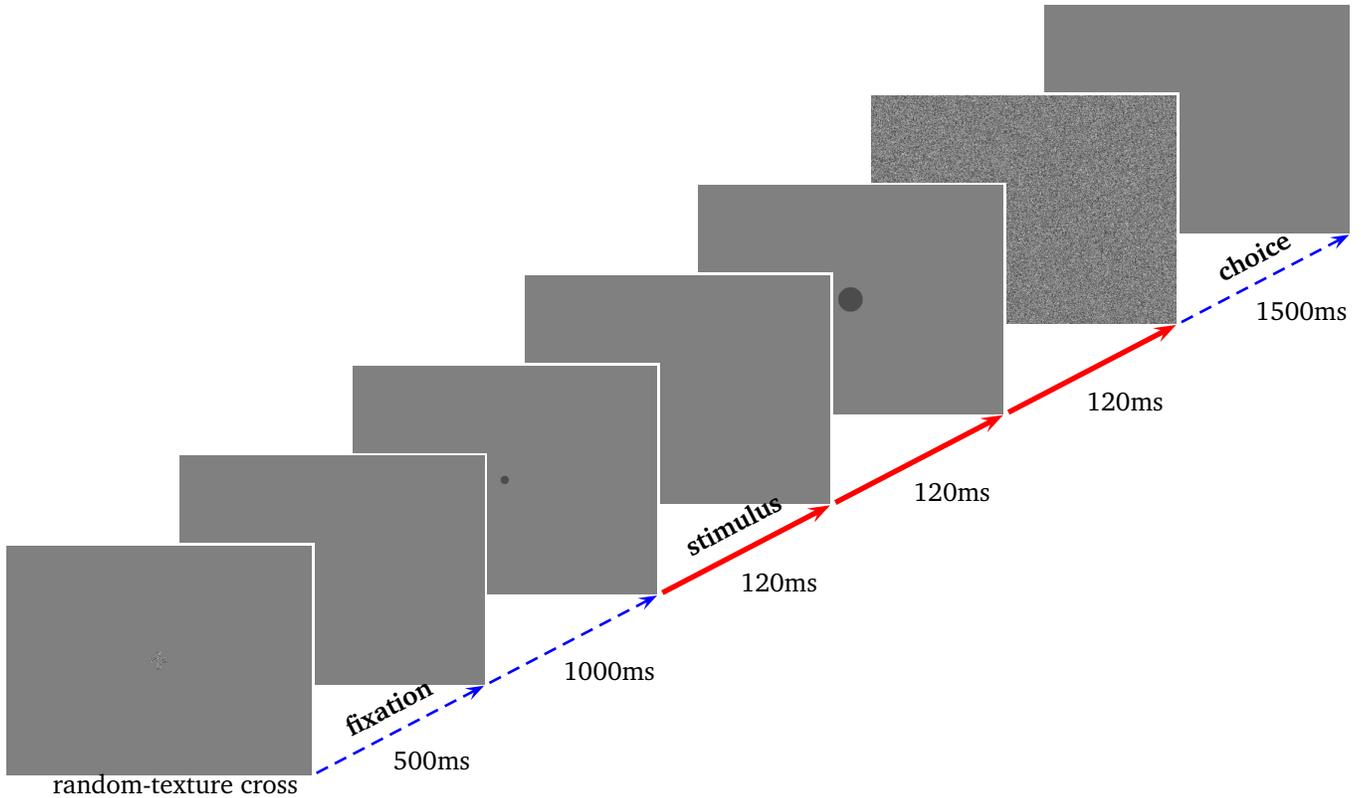
Figure 2: Each trial goes through fixation, stimulus, and choice stages. The fixation cross of radius $0.5°$, rendered in a random-dot texture to avoid interference with disk stimuli, is displayed for 500ms before disappearing. After 1 second of a blank screen, disk 1 (e.g. small), blank, and disk 2 (e.g. large) are shown in succession for 120 ms each. A choice screen is presented, and the subject is required to respond as soon as possible with a keypress before the blank screen is on in 1.5 seconds, indicating the start of next trial.

ing informed consent and were compensated with cash.

Analysis & Visualization. We treat the data from all the subjects as samples from a single subject population, since we are interested not in individual subjects' detection performance, but in the featural processing overall. We perform ANOVA tests on measurements to evaluate whether their differences are statistically significant between attributes or between feature dimensions.

Significant differences ($p < 0.05$) are visualized in black p values and **dashed lines** connecting two conditions in comparison. Insignificant differences ($p \geq 0.05$) are indicated with grayed out p values and solid lines.

We use error plots of a measurement, where the mean is marked by □ for size, ● for color, and △ for orientation. When comparing the 3 features, the p value from testing differences among 3 features (n-way ANOVA) is marked in the legend, and the p value from testing differences between any 2 features (two-sample t-test) is marked in the plot. When comparing 2 attribute conditions, we mark the p value in the legend, with significance visualized in

the line connecting them in the plot.

In short, dashed lines indicate significant differences.

Results

We evaluate detection performance with both accuracy and time. The detection time is measured from the choice screen onset to the subject's keypress response.

Fig. 3 shows that detecting a change or no-change is equally accurate and fast between the 2 attributes in each feature dimension. That is, there is no distinction between 2 attributes in the detection performance.

Fig. 4 shows that detecting a change is equally accurate and fast among the 3 feature dimensions, whereas detecting a no-change is more accurate and faster in color than in size or orientation.

Fig. 3 and Fig. 4 taken together, we have 3 conclusions. **1)** The two attributes are equally salient in its own feature dimension, with comparable no-change detection performance when spatial search is not needed.

change

change detection accuracy (%)

p=...50 □ size
p=...34 ● color
p=...25 △ orient.

no change

no–change detection accuracy (%)

p=0.89 □ size
p=...38 ● color
p=0.63 △ orient.

change detection time (ms)

p=...82 □ size
p=...90 ● color
p=...22 △ orient.

no–change detection time (ms)

p=... □ size
p=... ● color
p=0.22 △ orient.
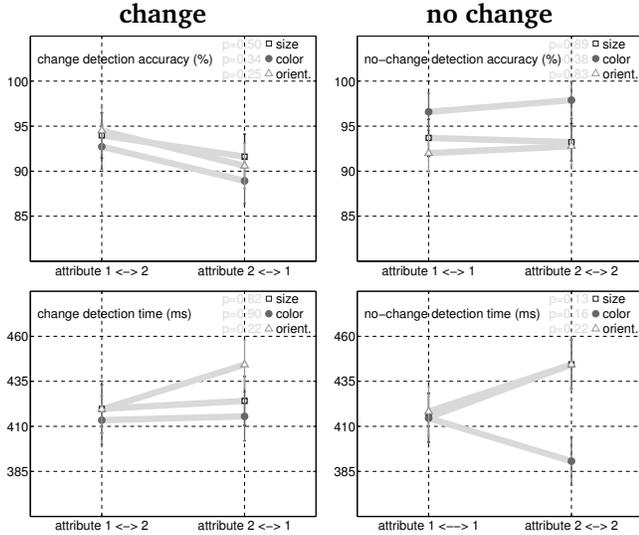
Figure 3: The 2 attributes in each feature are not different in the detection performance: The accuracy (top) and time (bottom) are comparable for detecting either a change (left) or a no-change (right). "Attribute $i \leftrightarrow j$" denotes attribute pairs of (i,j) and (j,i), e.g. "Attribute $1 \leftrightarrow 2$" for size refers to all the trials involving small to large and large to small changes.

2) The no-change detection performance is equivalent

change

change detection accuracy (%): p=0.69

p=0.45   p=0.92   p=0.50

change detection time (ms): p=0.45

p=0.50
p=0.57   p=0.20

no change

no–change detection accuracy (%): p=0.05

p=0.04   p=0.01
p=0.65

no–change detection time (ms): p=0.06

p=0.91
p=0.03   p=0.03
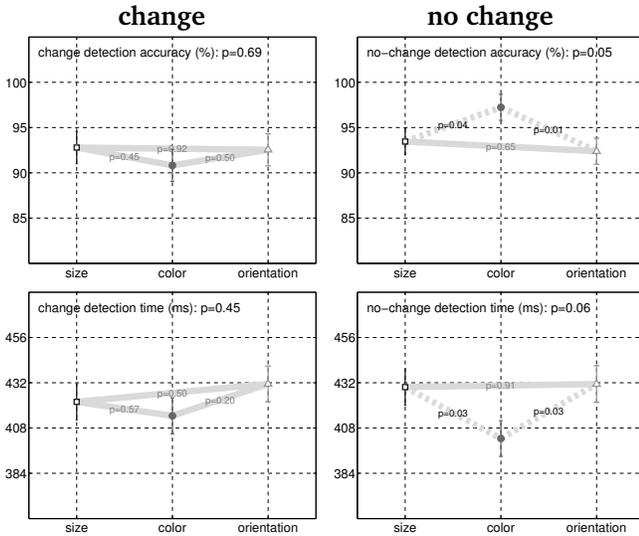
Figure 4: Detecting a change in the attribute is equally accurate (left top) and fast (left bottom) between the 3 features, whereas detecting a no-change is more accurate (right top) and faster (right bottom) in color than in size or orientation, the two of which are comparable. The significance of a difference is measured by the p value: the one in the top legend for comparison among 3 features, those in the plots for comparison between 2 connected features. Dashed lines indicate significant differences ($p < 0.05$).

between size and orientation, both worse than color: It is faster and more accurate in black or white than in small, large, horizontal or vertical. **3)** The change detection performance on a target flickering between our 2 attributes is equivalent among size, color, and orientation.

EXPERIMENT 2

The purpose of this change blindness experiment is to investigate how an attribute influences search behaviours for detecting and locating an attribute change which is comparable among the three feature dimensions. We show that size, color, and orientation engage attention in different ways: reduced fixation time for size, reduced saccade distance for color, and no impact for orientation.

Method

Stimuli. The flicker stimuli for the 3 feature dimensions are rendered in the identical spatial layout with the two-attribute sets studied in Experiment 1 (Fig. 1). Each stimulus involves a pair of 24-disk images which are identical except for one disk. These 24 disks are located centrally on a regular 4×6 grid, with an inter-disk distance of $5.4°$, which is 4 times the maximal radius a disk could assume. The 1st image of the stimulus consists of uniformly randomly distributed 12 attribute-1 disks and 12 attribute-2 disks. The 2nd image changes any one of the 24 disks from one attribute to the other (Fig. 5).

Apparatus. Stimuli are displayed using a Panasonic PT-LB50NTU Projector with 2000 ANSI lumens in image brightness onto a screen mounted at the back of a 4.3m $\times$ 8.8m room with no other lighting. The display extends $25.6° \times 34.1°$ at a viewing distance of 5 meters. Gaze data are recorded with a Tobii x50 eye tracker at 50Hz sampling rate and $0.5°$-$0.7°$ accuracy. Two clock-synced 3.2GHz Dell Precision computers control the eye tracker and the stimulus presentation respectively. The eye tracker is calibrated with the subject's gaze directions at the beginning of each data recording session.

Procedure. Each trial begins with 2-second display of a neutral gray background. A fixation dot of radius $0.5°$ is shown at the center of the display for 0.5 second, prompting the subject to gaze at the center. The flicker stimulus, in the sequence of disk image 1, blank, disk image 2, and blank, is then repeatedly presented for 120 ms each. Once the subject issues a mouse click to indicate his detection of the change, the last seen disk image is presented again till the subject indicates the location of change with another mouse click (Fig. 6).

There are 3 sessions, each using a different set of ran-

flicker stimuli	size		color		orientation	
identical spatial layout	1	2	1	2	1	2

1st image			
1 2 1 1 2 1			
2 1 2 2 2 1			
1 1 1 2 2 2			
2 2 2 1 1 ①			
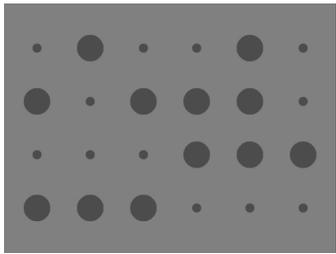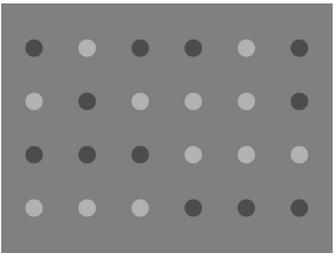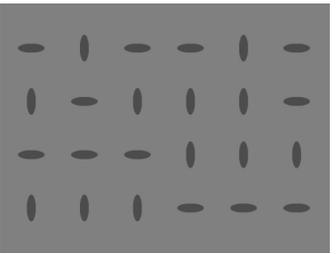
2nd image			
1 2 1 1 2 1			
2 1 2 2 2 1			
1 1 1 2 2 2			
2 2 2 1 1 ②			
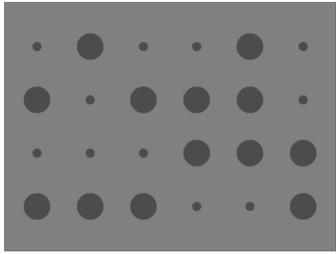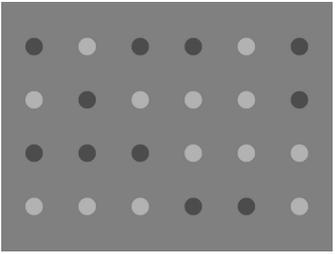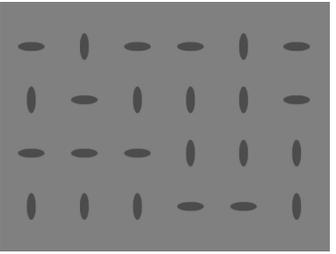
Figure 5: Flicker stimuli are rendered in the same spatial layout with the two- attribute sets studied in Fig. 1 for size, color, and orientation. The 1st image contains 12 attribute-1 disks and 12 attribute-2 disks in a uniformly random spatial distribution. The 2nd image is identical to the 1st image except that 1 disk changes its attribute. It could be any of the 24 disks. The disk of change here is circled in both layout matrices.
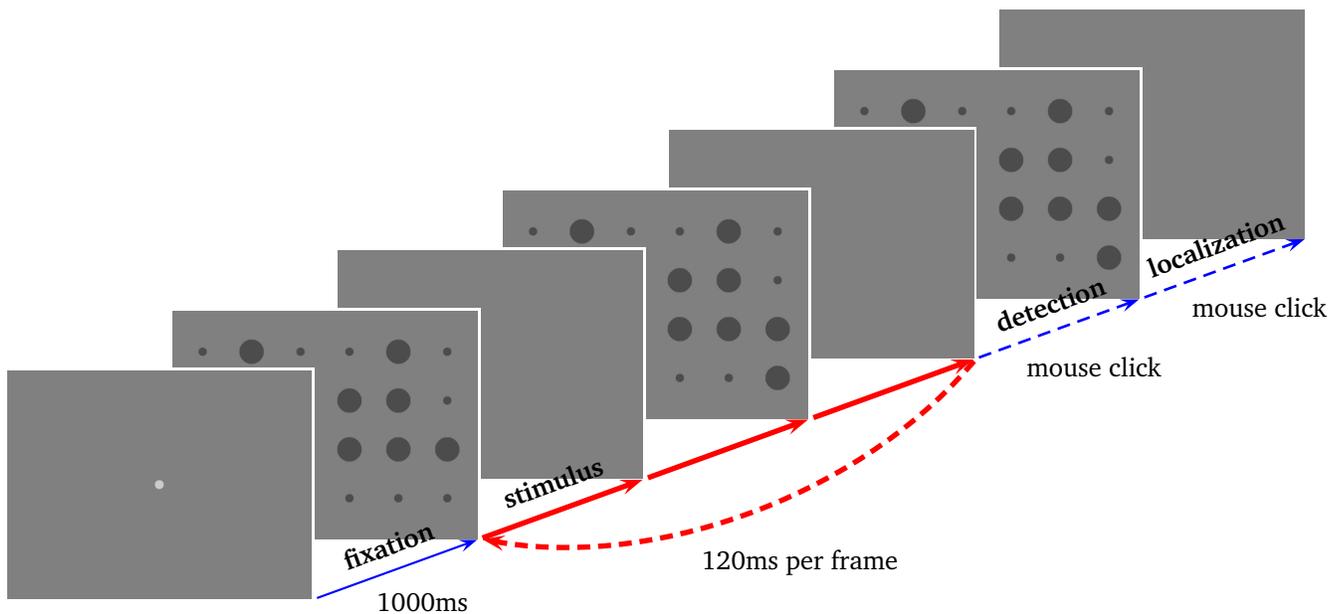
Figure 6: Each trial goes through fixation, stimulus, detection, and localization stages. A fixation dot is displayed for 1 second before the onset of the flicker stimulus, with disk image 1, blank, disk image 2, blank repeatedly presented at 120ms each. The subject issues a mouse click as soon as he detects the change, and the display is frozen at the last seen disk image, prompting him to localize the change by clicking the disk of change. A blank screen is then displayed for 2 seconds before the start of next trial.

dom stimuli. Each session has 3 blocks of 24 trials each, one trial for one change location and one block for one feature dimension. The trials are completely randomized in a block, and the order of blocks are also randomized and balanced among the subjects.

The subject is told that two images differing in only one disk are presented repeatedly. His task is to detect the changing disk. With his hand on a mouse, he should issue a click as soon as he detects the change. The flickering then stops at the last seen disk image, and he should click the disk which he believes has changed.

Participants. 25 Boston College students, with normal or corrected-to-normal vision, naive with respect to the experimental hypotheses, participated after providing informed consent and were compensated with cash.

Analysis & Visualization. As in Experiment 1, we treat the data from all the subjects as samples from a single subject population, since we are interested not in individual subjects' detection performance, but in the featural processing overall. We perform ANOVA tests on measurements to evaluate whether their differences are statistically significant between 2 attribute conditions in each feature dimension. We use error plots of a measurement, where the mean is marked by □ for size, ● for color, and △ for orientation.

Significant differences ($p < 0.05$) are visualized in black p values and **dashed lines** connecting two conditions in comparison. Insignificant differences ($p \geq 0.05$) are indicated with grayed out p values and solid lines.

Results

Fig. 7 shows how detection performance is different for the 3 feature dimensions when searching an equivalent change target among items of identical spatial layout. Fig. 9, Fig. 10, and Fig. 11 show respectively how search behaviours differ in the featural, spatial, and temporal domains to give rise to the distinctive performance.

Detection Performance. We evaluate detection performance with both accuracy and time. The detection time is measured from the flicker stimulus onset to the subject's first mouse click for indicating a detection.

Fig. 7 shows that when spatial search is needed, even among items of identical spatial layout, detecting an equivalent attribute change in the 3 feature dimensions yields different performances: it is most accurate and fastest in size, less in orientation, and least in color.

This result is the opposite of the detection performance without the need for spatial search, which is better in color than in size or orientation (Fig. 4). Specifically, detecting a change is equally accurate and fast,
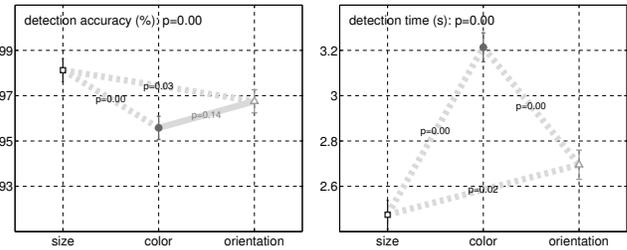
Figure 7: Detecting a flickering attribute change with spatial search in an array of items is best (most accurate and fastest) in size, worse in orientation, and worst in color. The accuracy (left) and time (right) are different ($p < 0.01$) between the 3 features (dashed lines), except that the accuracy in color is about the same as that in orientation (solid line).

and detecting a no-change is most accurate and fastest in color, and equally accurate and fast in size and orientation.

Clearly, the detection performance with spatial search cannot be understood singularly from the isolated performance on individual items. It must be the search behaviour from item to item that has altered the outcome. In particular, we examine how the feature attribute in a previous fixation influences what to look at (featural domain), where to look next (spatial domain), and how to look (temporal domain).

Featural Transitions. Fig. 9 shows that visual search is more likely to explore an attribute different from the current one in the 3 feature dimensions except size.

We first make a simplifying assumption to associate each fixation with the attribute of the nearest disk in the 4×6 disk array. This is supported by Fig. 8, where fixations cluster around the disk centers in the stimulus.

We can then count how many times an attribute is looked at and how many times visual search stays at the same disk or jumps to another disk with the same or different attribute. These are captured in statistics π_i, S_i, and P_{ij}, with i and j denoting attributes:

- The stationary probability π_i is the proportion of fixations at attribute i among all the fixations;
- The staying probability S_i is the proportion of successive fixations at the same disk of attribute i;
- The transition probability P_{ij} is the proportion of saccades towards j given the current fixation at i.

We disregard distance and contextual factors since: 1) All the attributes are uniformly randomly distributed in the space; 2) Stimuli have the identical spatial layout among the 3 features (Fig. 5). Differences in these statistics must be attributed to features themselves.
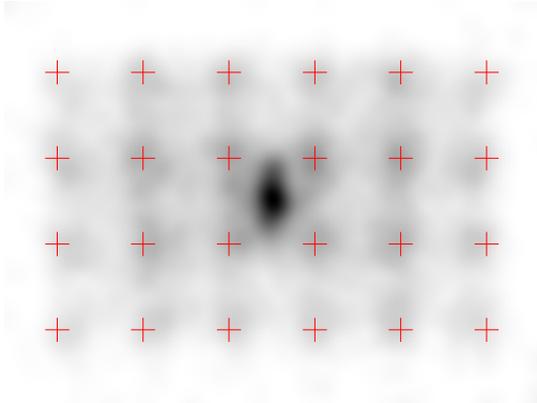
Figure 8: Fixations cluster around the disk centers (+) in the stimulus. Shown is the total fixation count distributed over the entire image, smoothed with a Gaussian of standard deviation $0.5°$, the fixation radius. The image center gets most fixations, since every trial starts with the subject fixating the center.

feature	stationary & transition probability				
	π_i	P_{ij}	•	●	S_i
size	.45	•	.27	.40	.33
	.55	●	.31	.39	.30
	π_i	P_{ij}	●	○	S_i
color	.53	●	.34	.36	.30
	.47	○	.40	.29	.31
	π_i	P_{ij}	⬭	⎮	S_i
orientation	.50	⬭	.31	.37	.32
	.50	⎮	.36	.32	.32

Figure 9: Feature transition with gaze shows that all the saccades are more likely to explore a different attribute except for size. Let i and j denote attributes, or row and column indices into the transition table. π_i is the overall probability of looking at i. P_{ij} is the probability of saccading to j at i. S_i is the probability of staying at the same disk of attribute i. For example, for size, π shows that 45% of all the fixations look at small and 55% at large. The 2nd row of P and S shows that upon fixating at a large disk, there is 30% chance of staying at it, 31% chance of saccading to a small disk and 39% chance to another large disk. These 3 numbers add to 1, since there are only 3 choices at each fixation: stay, or saccade away to the same or different type of disks. The most likely action is highlighted in red. While the chance of staying is about $\frac{1}{3}$, all the saccades tend to go to a different attribute except for size, where looking at another large disk is more likely. Search in color tends to stay away from white, whereas search in orientation switches between the two attributes indiscriminately.

There are essentially 3 action choices at each fixation: keep looking at the same disk, or jump away to a different disk with the same or different attribute.

According to S_i in Fig. 9, the chance of staying is about $\frac{1}{3}$ for every attribute. According to π_i, the chances of visiting the two attributes are overall equal for orientation, but different for size and color: more visits to large than to small, and more visits to black than to white.

The most interesting finding comes from P_{ij}: While the attributes are uniformly random, our eyes do not act like a blind space wanderer. For both color and orientation, it is more likely to explore a different attribute in the next fixation, whereas for size, it is more likely to explore the same large disk. In other words, search in size tends to stay within the group of large disks.

What distinguishes size from color and orientation is that visual search selects a particular size not a particular color and orientation. It is clear by all 3 accounts of π_i, S_i, and P_{ij}, search in orientation does not discriminate between horizontal and vertical. By π_i, search in color seems to stay away from white, yet by P_{ij}, it has no tendency to stay in the group of black (or white) disks. Only for size, by π_i there is a preference for large and by P_{ij} there is the same preference to stay with large.

In fact, the featural transition matrices in Fig. 9 represents 3 different types of Markov chains: Size has large as an attractive state, color has white as a repellent state, and orientation has horizontal and vertical as reversible states. Only search in size navigates by a particular attribute and is selective in the attribute itself.

Next we investigate whether the featural bias to the same attribute for size and to the different attribute for color and orientation manifests itself in the spatial and temporal behaviours of visual search. We analyze saccade distance and fixation time according to gaze transitions from the same and different attribute.

Saccade Distance. Fig. 10 shows that the saccade distance does not depend on what is looked at previously except for color.

For color, the saccade distance is $0.5°$ shorter if the previous attribute is the same as the current one, whereas for size and orientation, it is comparable whether the previous attribute is the same or different. Even in the same-attribute condition, the distance is shorter ($p < 0.001$) for color than for size and orientation, which are comparable ($p = 0.65$).
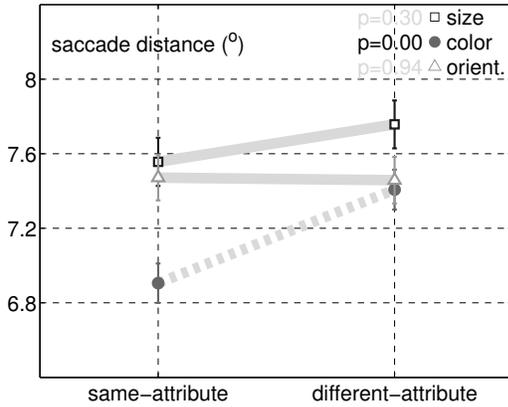
Figure 10: The saccade distance depends on what is looked at previously only for color, not for size or orientation. A shorter range of search ensues only when looking at the same color (dashed line).
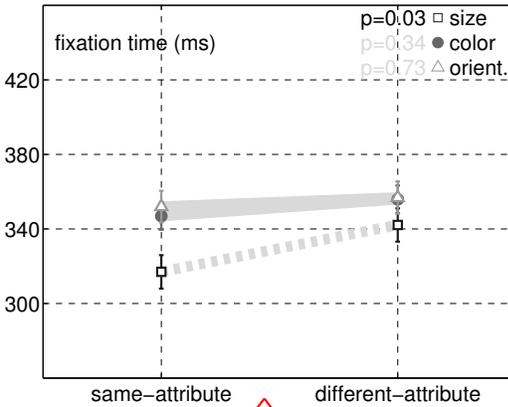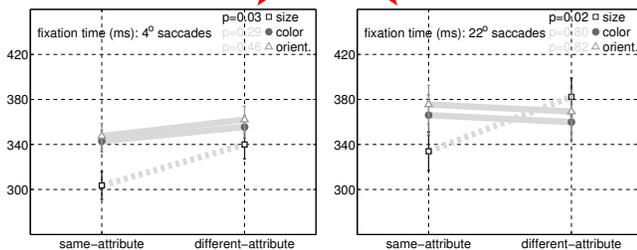
Figure 11: The fixation time depends on what is looked at previously only for size (dashed lines), not for color or orientation, and the effect of shortened fixation from previewing the same attribute does not diminish over distance. Top) Shortened fixation ensues only when looking at the disks of the same size. Bottom) This preview benefit applies to both short saccades of $[0.5°, 8.5°]$ (i.e. about 1 disk away) and long saccades of $[12°, 32°]$ (i.e. at least 2 disks away).

Several sources show that color search could be influenced by the previously looked color (Folk et al., 1992; Maljkovic and Nakayama, 1994; Found and Muller, 1996; Muller et al., 2003). We show that search in color engages spatial attention, and shorter-range search ensues when and only when moving between the same color.

Fixation Time. Fig. 11(Top) shows that the fixation time does not depend on what is looked at previously except for size.

In fact, if the previous attribute is different, the fixation time is the same ($p = 0.46$) for the 3 features; if the previous attribute is the same, the fixation time is 33 ms less ($p < 0.001$) for size than for color and orientation, with no difference between small and large ($p = 0.96$). These results suggest that only size has a preview benefit: While the fixation time on a disk does not change whether the color or orientation of disk is the same as the previous one, it is shortened if the size of the disk is the same as rather than different from the previous one.

Fig. 11(Bottom) shows that the preview benefit for size does not diminish over saccade distance. While the fixation time increases with the distance, the preview benefit nonetheless retains (if not increases) its magnitude even after saccades of more than $12°$.

Summary. According to Experiment 1, without the need for spatial search, detecting the presence of our 2-attribute change is equally accurate and fast among the 3 feature dimensions, whereas detecting the absence of attribute change is most accurate and fastest in color, and equally accurate and fast in size and orientation.

Given an identical spatial distribution of such attributes with a single change target, from the isolated change/no-change detection performance, one would expect change detection with spatial search to be most

attention	size	color	orientation
featural selection	✔	✘	✘
spatial selection	✘	✔	✘
temporal selection	✔	✘	✘

Figure 12: Size, color, and orientation engages attention in different forms of selection between successive fixations. There is a same-feature bias and shortened fixation on the same attribute for size, making detection more time efficient at the same spatial search capacity. There is a spatial focusing bias for color when successively looking at the same colors, reducing and jeopardizing the spatial coverage of search. There is not any form of priming in featural, spatial, or temporal domain for orientation.

accurate and fastest in color, and equally accurate and fast in size and orientation, yet according to Experiment 2, we find out that the performance is in fact the best in size, worse in orientation, and worst in color.

The key factor contributing to this discrepancy is that size, color, and orientation employ different forms of selection during prolonged active search (Fig. 12): The attribute in a previous fixation influences what to look at (featural domain), where to look next (spatial domain), and how to look (temporal domain).

For size, visual search is more likely to visit the same (and a particular) feature, which combined with shortened fixations at previewed attributes results in the best performance among the 3 features. For color, visual search reduces the range of spatial exploration when moving between the same colors, resulting in the worst performance among the 3 features. For orientation, there is no priming of any form, which results in a neutral performance that is worse than size but better than color.

DISCUSSIONS

Visual search behaviors in *static* displays have demonstrated that size, luminance, color, and orientation are elementary features extracted with separate sets of detectors responding in parallel across the visual scene. In schematic diagrams as well as computational models on visual saliency (Treisman, 1998; Itti and Koch, 2001; Torralba, 2004), image segmentation (Malik et al., 2001), or recognition (Riesenhuber and Poggio, 1999; Lowe, 2003; Zhang et al., 2006; Serre et al., 2007), it is invariably assumed that elementary features at all scales are processed and available simultaneously.

We use an *active* visual search task (Cavanagh and Mather, 1989; Cavanagh, 1991, 1992) in the change blindness flicker paradigm (Rensink et al., 1995; Rensink, 2000) to reveal their distinctive personalities in the dynamic deployment of visual attention, i.e., selective processing in the featural, spatial, and temporal domains between successive fixations.

When searching a target among an array of items, the performance is dominated not by the detection performance on individual items (Experiment 1) but by the search behaviour from item to item (Experiment 2). The former would predict better performance in color than in size and orientation, while the the actual performance is best in size, worse in orientation, and worst in color.

The performance ranking is reversed because, while visual search does not enhance processing in any domain for orientation between successive fixations, it deploys spatial focusing for color, and provides featural priming and temporal shortening for size.

Orientation turns out to be the only feature dimension that is truly pre-attentive (Treisman and Gelade, 1980; Maunsell and Newsome, 1987; Bergen and Julesz, 1983; Cavanagh et al., 1990; Foster and Ward, 1991; Wolfe et al., 1992, 1999; Wolfe and Horowitz, 2004), with no priming effects of any kind between fixations.

Color engages spatial attention when visual search navigates between items of the same attribute. The range of spatial search is reduced when looking at the items of the same color in successive fixations, thereby hurting the change detection performance.

This property is unique to color, not to size or orientation. It cannot be explained by any theory of a retinal origin: Our color stimuli are black and white disks, not disks of different hues which can only been seen in the fovea, and there is also no reduction in the distance when saccading between black and white disks.

Attribute-specific cuing effects, i.e. expedited detection with a top-down or inter-trial cuing, have been observed for color, but not for orientation (Folk et al., 1992; Maljkovic and Nakayama, 1994; Found and Muller, 1996; Muller et al., 2003). We show that visual search within the same trial narrows the spatial focus only when saccading between items of the same color. This result could thus reflect the engagement of spatial attention by color-specific cuing from one fixation to the next.

Size engages featural attention when visual search navigates between items of the same attribute. The duration of a subsequent fixation at the same attribute is shortened over any saccade distance, thereby improving the change detection performance.

A saccade towards a peripheral location in the visual field often requires orienting attention to that location (Hoffman and Subramaniam, 1995). Spatial attention enhances visual processing, e.g. faster and more accurate processing of luminance and form (Downing, 1988; Hawkins et al., 1990), and shortening the subsequent fixation on the same object surrounding that location (Henderson et al., 1989).

However, our preview benefit result of shortening the subsequent fixation cannot be explained by orienting attention spatially to the next item with every saccade in visual search, since it only exists for size, not for color or orientation, and it only exists between items of the same size. What it suggests instead is that attention can be allocated to a particular location as well as a particular size, but not a particular color or orientation.

This featural attention for size is different from the

spatial attention for color: The former enhances the processing of the size attribute itself through reduced fixation on the item of the same size irrespective of the saccade distance, whereas the latter only reduces the saccade distance to reach the next item of the same color without any benefit on processing that item.

Visual attention is often likened to the zoom-lens of a camera, where a beam of focus can be directed to a selected region in the visual field. Compared to the spotlight metaphor, the zoom-lens metaphor emphasizes the trade-off between the size of the beam and the resolution of the beam: Attention can either be sharply focused on a small region or coarsely distributed over a wide area (Eriksen and Yeh, 1985; Eriksen and James, 1986; Castiello and Umilta, 1990; Theeuwes, 1993; LaBerge et al., 1997; Greenwood and Parasuraman, 1999, 2004).

The attentional scaling model of visual search postulates that visuospatial attention possesses two independently deployed properties: shifting in space and varying in scale (Greenwood and Parasuraman, 1999). It has been shown that the size of the precue directly affects the size of the attentional beam, and it facilitates search only when the precue correctly indicates the location of a target fixed in size (Greenwood and Parasuraman, 2004). It has also been shown that implicit memory for the size of the attentional focus can guide visual search even in the absence of feature or position priming, or distractor contextual effects (Fuggetta et al., 2009). Our preview benefit result, which exists only for successive fixations on the same size, is consistent with these findings.

The zoom-lens metaphor implies that the scale of attention is narrowband: The visual system processes one stimulus size at a time, and it must pass attentionally through intermediate sizes in order to select a different size. In an attempt to measure the scale bandwidth of attention (Verghese and Pelli, 1994) and the relation between attending to stimulus size and attending to spatial area, Farell and Pelli have found out that the scale tuning is task-dependent: Subjects are poorer at localizing a target in mixed-scale displays than in single-scale displays, but their ability to identifying the target is unimpaired (Farell and Pelli, 1993). That is, one can only attend to one size at a time for localizing, but can attend to large and small at the same time for identifying. Since our change blindness task requires spatial search and localization of the change target, our observation that only size not color or orientation can tune into a particular attribute is in accordance with the previous work as well.

Our result that the preview benefit for size does not diminish over distance breaks the zoom-lens metaphor on its view that attention has a trade-off between resolution
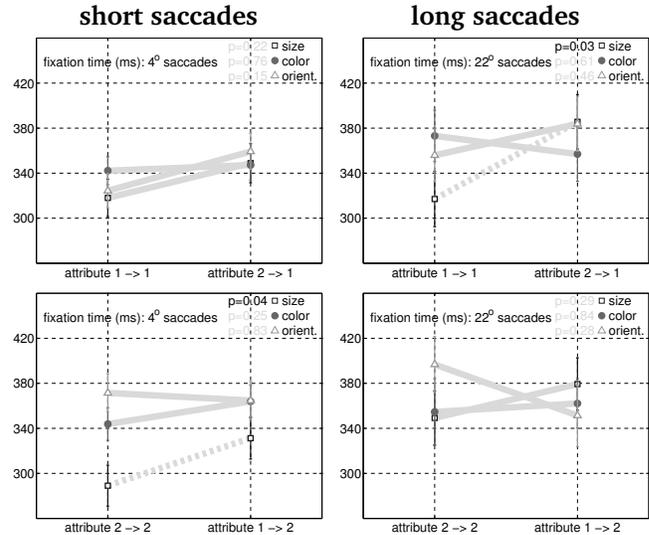
Figure 13: The fixation is shortened from previewing the same size (not color or orientation), and the effect is most significant (dashed lines) for small afar (top right) and for large nearby (bottom left). "Attribute $i \to j$" denotes previous fixation at i and current fixation at j, e.g. the top row compares the fixation time on small from previewing small and large, over short (left) and long (right) distances. Short saccades are about 1 disk away ($[0.5°, 8.5°]$) and long saccades are at least 2 disks away ($[12°, 32°]$).

and coverage. According to the zoom-lens metaphor, if the resolution of the attentional beam becomes finer, the size of the beam becomes smaller. In other words, if the visual system attends to small, the preview benefit for a subsequent fixation at small should reduce over the saccade distance. Fig. 11 shows that the preview benefit holds for a long saccade. Fig. 13 provides an even stronger argument: Previewing a small disk in fact most significantly shortens the fixation at another small disk far away, and previewing a large disk most significantly shortens the fixation at another large disk nearby.

It seems counter-intuitive that the preview benefit is most significant for small over long saccades and for large over short saccades. It may be explained by the retina's non-uniform spatial resolution. The benefit is not much for small over short saccades, probably because the high-resolution fovea is already tuned to small details, and it is the low-resolution periphery that could use the help. Conversely, the preview benefit is not much for large over short saccades, probably because the low-resolution periphery is already tuned to coarse details, and it is the fovea that could use the help.

The interaction between attention and the non-uniform resolution of the human eye has been investigated by Yeshurun and Carrasco (Yeshurun and Car-

rasco, 1998, 2008). Using a texture segmentation task which requires fine-scale processing, they show that performance is improved at a peripheral location (and impaired at a central location) only for a small cue attracting attention to that location. That is, attention enhances spatial resolution at the attended location when it is attracted to that location by a small cue but does not lower resolution when it is attracted by a large cue.

Our understanding of the preview benefit in size questions their interpretation of the results that attention cannot adapt its operation on spatial resolution based on the size of the attentional cue (Yeshurun and Carrasco, 2008). While they systematically manipulated the cue size, they did not vary the texture resolution. Consequently, they only observed small cues' benefits on fine resolution, not large cues' benefits on coarse resolution. Our prediction is that a texture segmentation or another visual task which requires coarse-scale processing would show improved performance with a large cue attracting attention to a central location not a peripheral location.

To conclude, change detection of a target flickering between two attributes among randomly distributed attributes rendered identically in space and separately in size, color, and orientation shows that the 3 elementary features have distinctive processing characteristics: Orientation is truly pre-attentive with no priming in featural, spatial, or temporal domain, color engages spatial attention with shorter saccades between the same colors, and size engages featural attention with attribute tuning and a preview benefit. In other words, with a preview benefit, attention can select a particular location and a particular size, but not a particular color or orientation.

AUTHOR NOTE

REFERENCES

Anderson, B. A. and Folk, C. L. (2010). Variations in the magnitude of attentional capture: Testing a two-process model. *Atten Percept Psychophys*, 72(2):342–52.

Bauer, B., Jolicoeur, P., and Cowan, W. B. (1996). Visual search for colour targets that are or are not linearly separable from distractors. *Vision Research*, 36(10):1439–66.

Becker, S. I. and Horstmann, G. (2009). A feature-weighting account of priming in conjunction search. *Atten Percept Psychophys*, 71(2):258–72.

Bergen, J. R. and Julesz, B. (1983). Rapid discrimination of visual patterns. *IEEE Trans Syst. Man Cybern.*, SMC-13:857–63.

Castiello, U. and Umilta, C. (1990). Size of the attentional focus and efficiency of processing. *Acta Psychologica*, 73.

Cavanagh, P. (1991). Short-range vs long-range motion: not a valid distinction. *Spatial Vision*, 5:303–9.

Cavanagh, P. (1992). Attention-based motion perception. *Science*, 257:1563–5.

Cavanagh, P., Arguin, M., and Treisman, A. (1990). Effect of surface medium on visual search for orientation and size features. *J. Exp. Psychol. Hum. Percept. Perform.*, 16:479–92.

Cavanagh, P. and Mather, G. (1989). Motion: The long and short of it. *Spatial Vision*, 4(2/3):103–29.

Downing, C. J. (1988). Expectancy and visual-spatial attention: effects on perceptual quality. *J. Exp. Psychol. Hum. Percept. Perform.*, 14:188–202.

Duncan, J. and Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, 96(3):433–58.

D'Zmura, M. (1991). Color in visual search. *Vision Research*, 31(6):951–66.

Eckstein, M. P. (1998). The lower visual search efficiency for conjunctions is due to noise and not serial attentional processing. *Psychol. Sci.*, 9:111–8.

Eriksen, C. W. and James, J. D. S. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception and Psychophysics*, 40:225–40.

Eriksen, C. W. and Yeh, Y.-y. (1985). Allocation of attention in the visual field. *J. Exp. Psychol. Hum. Percept. Perform.*, 11(5):583–97.

Farell, B. and Pelli, D. G. (1993). Can we attend to large and small at the same time? *Vision Research*, 33(18):2757–72.

Feldman, J. A. (1985). Four frames suffice: a provisional model of vision and space. *Behavioral and Brain Sciences*, 8:265–89.

Folk, C. L., Remington, R. W., and Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *J. Exp. Psychol. Hum. Percept. Perform.*, 18:1030–44.

Foster, D. H. and Ward, P. A. (1991). Asymmetries in oriented-line detection indicate two orthogonal filters in early vision. *Proc. R Soc. Lond. B*, 243:75–81.

Found, A. and Muller, H. J. (1996). Searching for un-

known feature targets on more than one dimension: investigating a dimension-weighting account. *Perception and Psychophysics*, 58:88–101.

Fuggetta, G., Lanfranchi, S., and Campana, G. (2009). Attention has memory: priming for the size of the attentional focus. *Spatial Vision*, 22(2):147–59.

Geyer, T., Muller, H. J., and Krummenacher, J. (2006). Cross-trial priming in visual search for singleton conjunction targets: Role of repeated target and distractor features. *Perception and Psychophysics*, 68(5):736–49.

Greenwood, P. M. and Parasuraman, R. (1999). Scale of attentional focus in visual search. *Perception and Psychophysics*, 61(5):837–59.

Greenwood, P. M. and Parasuraman, R. (2004). The scaling of spatial attention in visual search and its modification in healthy aging. *Perception and Psychophysics*, 66(1):3–22.

Grimes, J. (1996). On the failure ot detect changes in scenes across saccades. 2.

Hawkins, H. L., Hillyard, S. A., Luck, S. J., Mouloua, M., Downing, C. J., and Woodward, D. P. (1990). Visual attention modulates signal detectability. *J. Exp. Psychol. Hum. Percept. Perform.*, 16:802–11.

Henderson, J. M. (1997). Transsaccadic memory and integration during real-world object perception. *Psychological Science*, 50:243–71.

Henderson, J. M., Pollatsek, A., and Rayner, K. (1989). Covert visual attention and extrafoveal information use during object identification. *Perception and Psychophysics*, 45:196–208.

Hoffman, J. E. and Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception and Psychophysics*, 57(6):787–95.

Hollingworth, A., Williams, C. C., and Henderson, J. M. (2001). To see and remember: visually specific information is retained in memory from previously attended objects in natural scenes. *Psychonomic Bulletin & Review*, 8(4):761–68.

Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology*, 23:420–56.

Irwin, D. E. and Andrews, R. (1996). Integration and accumulation of information across saccadic eye movements. In Inui, T. and McClelland, J. L., editors, *Attention and Performance XVI: Information integration in perception and communication*.

Itti, L. and Koch, C. (2001). Computational modelling of visual attention. *Nature Neuroscience*, pages 194–203.

Kahneman, D., Treisman, A., and Gibbs, B. J. (1992). The reviewing of object files: object-specific integration of information. *Cognitive Psychology*, 24:175–219.

Kristjansson, A. and Campana, G. (2010). Where perception meets memory: A review of repetition priming in visual search tasks. *Atten Percept Psychophys*, 72:5–18.

LaBerge, D., Carlson, R. L., Williams, J. K., and Bunney, B. G. (1997). Shifting attention in visual space: Tests of moving-spotlight models versus an activity-distribution model. *J. Exp. Psychol. Hum. Percept. Perform.*, 23(5):1380–92.

Levin, D. T. and Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin and Review*, 4:501–6.

Lowe, D. G. (2003). Distinctive image features from scale-invariant keypoints.

Malik, J., Belongie, S., Leung, T., and Shi, J. (2001). Contour and texture analysis for image segmentation. *International Journal of Computer Vision*.

Maljkovic, V. and Nakayama, K. (1994). The priming of pop-out I: Role of features. *Memory and Cognition*, 22:657–72.

Maunsell, J. H. R. and Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, 10:363–401.

Moraglia, G. (1989). Visual search: spatial frequency and orientation. *Perceptual and Motor Skills*, 69:675–89.

Muller, H. J., Reimann, B., and Krummenacher, J. (2003). Visual search for singleton feature targets across dimensions: stimulus and expectancy-driven effects in dimension weighting. *J. Exp. Psychol. Hum. Percept. Perform.*, 29(5):1021–35.

Nagy, A. L. and Sanchez, R. R. (1990). Critical color differences determined with a visual search task. *Journal of the Optical Society of America*, 7:1209–17.

O'Regan, J. K., Deubel, H., Clark, J. J., and Rensink, R. A. (2000). Picture changes during blinks: looking without seeing and seeing without looking. *Visual Cognition*, 7:191–211.

O'Regan, J. K., Rensink, R. A., and Clark, J. J. (1999). Change-blindness as a result of 'mudsplashes'. *Nature*, 398:34.

Palmer, J., Verghese, P., and Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, 40:1227–68.

Pollatsek, A. and Rayner, K. (1992). What is integrated across fixations? In Rayner, K., editor, *Eye movements and visual cognition: scene perception and reading*, pages 166–91.

Rensink, R. A. (2000). Seeing, sensing, and scrutinizing. *Vision Research*, 40:1469–87.

Rensink, R. A. (2002). Change detection. *Annual Review of Psychology*, 53:4245–277.

Rensink, R. A., O'Reagan, J. K., and Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Visual Cognition*, 7:127–145.

Rensink, R. A., O'Regan, J. K., and Clark, J. J. (1995).

Image flicker is as good as saccades in making large scene changes invisible. *24*, pages 26–8.

Riesenhuber, M. and Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2(11):1019–25.

Sagi, D. (1988). The combination of spatial frequency and orientation is effortlessly perceived. *Perception and Psychophysics*, 43:601–3.

Serre, T., Oliva, A., and Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. *Proceedings of National Academy of Sciences*, 104(15):6424–9.

Simons, D. J. and Levin, D. T. (1998). Failure to detect changes to people in a real-world interaction. *Psychonomic Bulletin and Review*, 5:644–9.

Theeuwes, J. (1993). Visual selective attention: A theoretical analysis. *Aeta Psychologica*, 83:93–154.

Torralba, A. (2004). Contextual influences on saliency. In Itti, L., Rees, G., and Tsotsos, J., editors, *Neurobiology of Attention*, pages 586–93. Academic Press.

Trehub, A. (1991). *The cognitive brain*.

Treisman, A. (1998). The perception of features and objects. In Wright, R. D., editor, *Visual Attention*. Oxford University Press.

Treisman, A. and Gelade, G. (1980). A feature-integration theory of atttention. *Cognitive Psychology*, 12(1):97–136.

Treisman, A. and Souther, J. (1985). Search assymmetry: a diagnostic for preattentive processing of seperable features. *Journal of Experimental Psychology: General*, 114:285–310.

Verghese, P. (2001). Visual search and attention: a signal detection approach. *Neuron*, 31:523–35.

Verghese, P. and Pelli, D. G. (1994). The scale bandwidth of visual search. *Vision Research*, 34(7):955–62.

Wolfe, J. M. (2001). Asymmetries in visual search: an introduction. *Perception and Psychophysics*, 63:381–9.

Wolfe, J. M. ., Cave, K. R., and Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance*, 15(3):419–33.

Wolfe, J. M., Friedman-Hill, S. R., Stewart, M. I., and O'Connell, K. M. (1992). The role of categorization in visual search for orientation. *J. Exp. Psychol. Hum. Percept. Perform.*, 18:34–49.

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

Wolfe, J. M., Klempen, N. L., and Shulman, E. P. (1999). Which end is up? Two representations of orientation in visual search. *Vision Research*, 39(12):2075–86.

Wolfe, J. M., Reinecke, A., and Brawn, P. (2006). Why don't we see changes? the role of attentional bottlenecks and limited visual memory. *Visual Cognition*, 19(4-8):749–80.

Yeshurun, Y. and Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature*, 396(6706):72–5.

Yeshurun, Y. and Carrasco, M. (2008). The effects of transient attention on spatial resolution and the size of the attentional cue. *Perception and Psychophysics*, 70(1):104–13.

Zelinsky, G. J. (2001). Eye movements during change detection: Implications for search constraints, memory limitations, and scanning strategies. *Perception and Psychophysics*, 63(2):209–25.

Zhang, H., Berg, A. C., Maire, M., and Malik, J. (2006). Svm-knn: Discriminative nearest neighbor classification for visual category recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2126–36.