

Spatiotemporal templates for detecting orientation-defined targets

Department of Psychology, Neuroscience & Behaviour,
McMaster University, Hamilton, ON, Canada, &
Institute for Human Science and Biomedical Engineering,
National Institute of Advanced Industrial Science and
Technology (AIST), Tsukuba, Ibaraki, Japan



Masayoshi Nagai

Department of Psychology, Neuroscience & Behaviour,
McMaster University, Hamilton, ON, Canada



Patrick J. Bennett

Department of Psychology, Neuroscience & Behaviour,
McMaster University, Hamilton, ON, Canada



Allison B. Sekuler

Using the classification image technique, the present experiments revealed several characteristics of human observers' spatiotemporal templates for the detection of orientation-defined targets. The stimulus consisted of a spatial 5×5 array of elements displayed in 5 (Experiments 1 and 2) or 15 (Experiment 3) temporal frames. A target was defined by the first- or the second-order characteristics of the textures. In Experiment 1, a target signal was presented across all five frames, and observers typically relied on the most reasonable cues in all five frames for detecting targets. In other words, they used the first-order cue for detecting the first-order target and used the second-order cue for detecting the second-order target. Moreover, the spatial profile for detecting the first-order sustained target was localized at the border of the target area, but that for the second-order sustained target showed broader spatial tuning. Presenting the target in just the third temporal frame, as was done in Experiment 2, changed the temporal profile of the observers' templates in the expected manner: Observers used the first-order cue for the first-order target detection and the second-order cue for the second-order target detection only in the third frame. However, changing the temporal characteristics also affected the kinds of spatial cues that were used to detect a target. For example, the classification images revealed that observers used second-order cues (as well as first-order cues) to detect a first-order target, and there was a trend toward increasing the extent of spatial information used when the temporal information was restricted. In Experiment 3, we found similar results for detecting the first-order flashed target with finer, 15-temporal-frame presentation. Lastly, we showed that the classification image is a useful way to reveal individual differences that are not shown with traditional psychophysical techniques.

Keywords: response classification, classification image, texture segregation, first- and second-order cues, spatiotemporal templates, orientation

Citation: Nagai, M., Bennett, P. J., & Sekuler, A. B. (2007). Spatiotemporal templates for detecting orientation-defined targets. *Journal of Vision*, 7(8):11, 1–16, <http://journalofvision.org/7/8/11/>, doi:10.1167/7.8.11.

Introduction

The segregation of visual scenes is a critical process in early vision. In natural scenes, segregation is achieved mainly by extracting luminance-defined edges separating different objects. However, the human visual system can segregate scenes into different regions even when no luminance cues are available. For example, texture patterns with line arrays at different orientations segregate from each other (e.g., Beck, 1966a, 1966b, 1967; Nothdurft, 1985, 1992, 1993a, 1993b), as do regions that differ in color (Moller & Hurlbert, 1996; Nothdurft, 1993b), motion (Nothdurft, 1993a, 1993b; Sekuler, 1990), and temporal synchrony (Kandil & Fahle, 2001; Lee & Blake, 1999; Morgan & Castet, 2002; Sekuler &

Bennett, 2001). Moreover, Nothdurft found that local differences in elements' orientation, luminance, color, and direction of motion must be increased at the border for the segmentation to occur when the overall variation of features within each region is raised (e.g., Nothdurft, 1992, 1993a, 1993b; for review, see Nothdurft, 1994). These psychophysical studies have described the visual attributes that can serve as cues for visual segregation. However, with traditional psychophysical techniques, it is difficult to show how each local texture element contributes to the visual segregation. For example, although Nothdurft demonstrated that local differences in visual attributes were important cues for segregation, it remains unclear whether observers use all local differences or just one element (or several elements) to segregate a texture into distinct regions.

The present experiments were performed to examine how individual texture elements influence judgments in a texture discrimination task. Specifically, we asked the following: Which texture elements are critical for segmentation? At which times during the stimulus presentation do those elements affect perception? And, how consistent are processing strategies across individuals and conditions? We used the response classification technique (Ahumada & Lovell, 1971; Beard & Ahumada, 1998) to address these questions.

In this technique, in each trial, a unique external noise is added to a stimulus that an observer must classify (e.g., A or B). In some trials, the observer's classifications will be correct. However, in other trials, the noise may make one stimulus (e.g., Stimulus A) look more like the other stimulus (e.g., Stimulus B), leading to incorrect classifications. After many trials, the noise fields presented on each trial are sorted into four stimulus–response classes (N_{AA} , N_{AB} , N_{BA} , and N_{BB}). Here, N_{AB} represents all samples of noise fields where Stimulus A was presented and the observer classified it as Stimulus B. The mean classification image (C_{mean}) is calculated as follows:

$$C_{\text{mean}} = [\text{Mean}(N_{AA}) + \text{Mean}(N_{BA})] - [\text{Mean}(N_{BB}) + \text{Mean}(N_{AB})]. \quad (1)$$

The classification image is a map that shows the stimulus locations where the values of the noise affected an observer's responses; thus, it can be thought of as a “behavioral receptive field” (Gold, Murray, Bennett, & Sekuler, 2000). The mean classification image (Equation 1) is an estimate of the linear template used by an observer (Murray, Bennett, & Sekuler, 2002) to perform a particular task, but the technique can also be used to estimate the effects of nonlinear mechanisms (Neri & Heeger, 2002). The variance classification image (C_{var}), which estimates one kind of nonlinear template, is calculated as follows:

$$C_{\text{var}} = [\text{Var}(N_{AA}) + \text{Var}(N_{BA})] - [\text{Var}(N_{BB}) + \text{Var}(N_{AB})]. \quad (2)$$

This second-order image represents the association between the observer's response and the squared noise value at each location in the stimulus.

The standard classification image technique, or variations thereof, has been used to study a variety of visual phenomena, including vernier acuity (Beard & Ahumada, 1998), face discrimination (e.g., Gold, Sekuler, & Bennett, 2004; Gosselin & Schyns, 2003; Mangini & Biederman, 2004; Sekuler, Gaspar, Gold, & Bennett, 2004), perceptual organization (Gold et al., 2000), attention (e.g., Eckstein, Shimozaki, & Abby, 2002; Neri & Heeger, 2002; Solomon,

2002; Tse, Sheinberg, & Logothetis, 2003), perceptual learning (Gold et al., 2004), and stereo vision (Gosselin, Bacon, & Mamassian, 2004; Neri, Parker, & Blakemore, 1999). In the following experiments, we used the classification image technique to examine how observers discriminate textures composed of differently oriented line segments.

Experiment 1: Detection of orientation-defined first- and second-order sustained targets

The textures used in all of the following experiments consisted of a spatial 5×5 array of elements displayed in five temporal frames. In the first experiment, the task was to detect a target texture that was presented in the center three rows of the array on all five temporal frames. The target and background textures differed either in mean orientation of the elements (first-order target) or in orientation variance (second-order target).

Methods

Observers

Four undergraduate students and one graduate student (age range = 19–27 years; mean age = 21.33 years) at McMaster University participated in Experiment 1. All had normal or corrected-to-normal visual acuity and were naive regarding the purpose of the experiment. Observers were paid \$10 for each session.

Apparatus

Stimuli were displayed on a 21-in. AppleVision monitor (resolution: 1152×870 pixels; screen size: 38.0×28.5 cm; refresh rate: 75 Hz), controlled by an Apple G3 computer. Observers viewed the stimuli binocularly from a distance of 100 cm, and head position was stabilized with a chin-and-forehead rest.

Stimuli

The stimulus was a dynamic texture that consisted of a five-frame movie, with each frame containing one 5×5 array of oriented, Gaussian-blurred lines (Figure 1A). Each movie frame was presented for 80 ms (i.e., six video refreshes). Each 0.264° line was centered on one intersection of an invisible grid made of lines separated by approximately 0.34° . Thus, the total stimulus size was approximately $1.62^\circ \times 1.62^\circ$. The orientation of a line ranged from 90° (vertical) to 180° (horizontal) in steps of 0.5° (see Figure 1B). After the presentation of the movie,

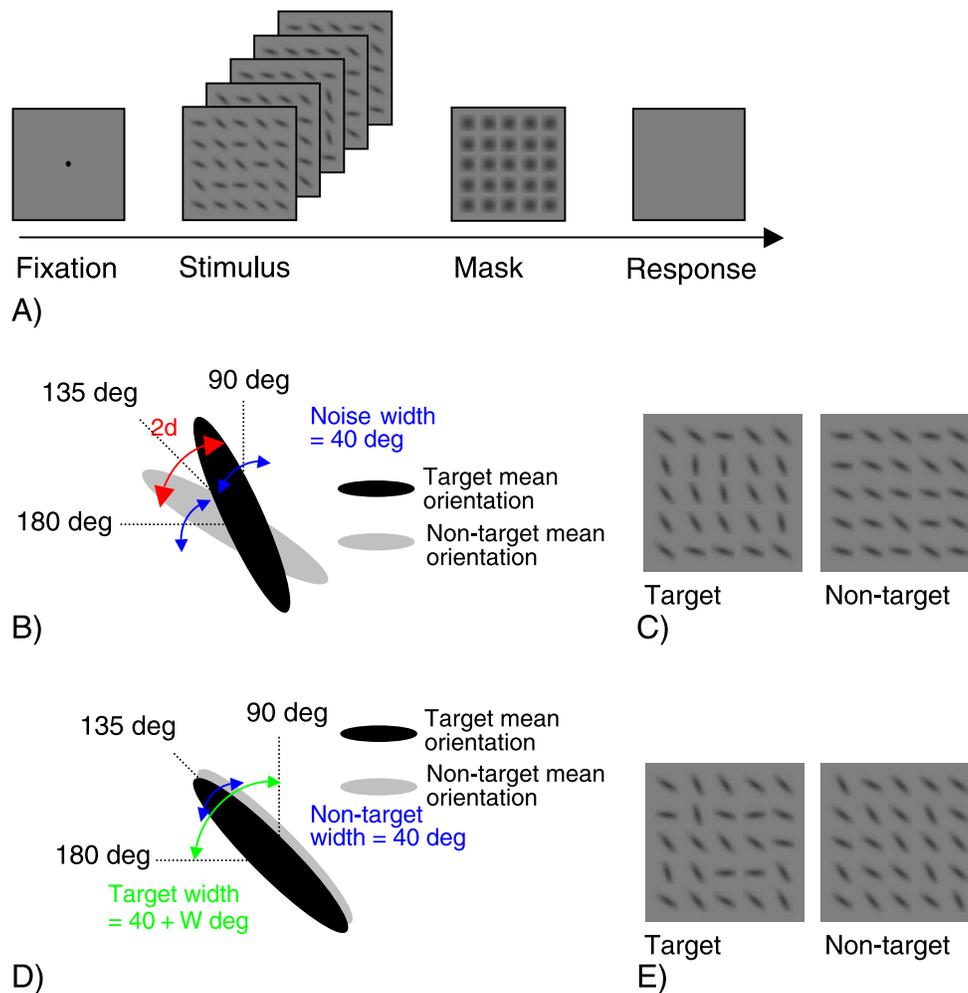


Figure 1. The stimuli used in this study. (A) The time course of a trial. (B) A representation of line blob at the target and nontarget regions in the mean orientation condition. (C) An example of the target and nontarget textures in the mean orientation condition. (D) A representation of line blob at the target and nontarget regions in the orientation variance condition. (E) An example of target and nontarget textures in the orientation variance condition. See text for details.

a 5×5 array of circular blobs was presented as a mask. The lines and circular blobs had a negative contrast of 50% against the background luminance (46.6 cd/m^2). The spatial contrast profiles of the lines and blobs were defined as follows:

$$\text{Contrast}_{\text{line}}(x, y) = -0.5 \cdot \exp\left\{-\pi \cdot [x'/10.5]^2\right\} \times \exp\left\{-\pi \cdot [y'/3.5]^2\right\} \quad (3)$$

and

$$\text{Contrast}_{\text{blob}}(x, y) = -0.5 \cdot \exp\left\{-\pi \cdot [x'/10.5]^2\right\} \times \exp\left\{-\pi \cdot [y'/10.5]^2\right\}, \quad (4)$$

where $x' = [(x - 7)\cos(\theta) + (y - 7)\sin(\theta)]$, $y' = [-(x - 7)\sin(\theta) + (y - 7)\cos(\theta)]$, θ is orientation (in radians), and x

and y are integers ranging from 0 to 13. Observers attempted to discriminate textures that contained a target from textures that did not. In a nontarget texture, the orientations of all elements were drawn randomly from the same distribution. In a target texture, the orientations of all elements in the middle three rows were selected from one distribution, whereas the orientations of the remaining elements were drawn from a different distribution. In the first-order, or mean orientation, condition (Figures 1B and 1C), the distributions of the target and nontarget orientations were uniform distributions (width = 40° , minimum step size = 0.5°) that differed only in mean orientation. Specifically, the mean target and nontarget orientations were $135^\circ + d$ and $135^\circ - d$, respectively. The difference between means, $2d$, was adjusted for each observer to produce correct responses on approximately 75% of the trials. For the second-order, or orientation variance, condition (Figures 1D and 1E), the distributions of the target and nontarget orientations were uniform

Observer	75% Threshold (deg)	% Correct	d'	β
I.A.	5.0	73.3	1.250	0.875
J.M.	5.0	72.9	1.225	1.127
Y.Y.	12.0	75.8	1.399	1.035
V.A.	8.0	75.7	1.388	0.993

Table 1. Seventy-five percent threshold, percentage correct, d' , and bias measure (β) for the sustained target detection in the mean orientation condition in [Experiment 1](#).

distributions that had the same mean (135°) but different widths (i.e., variances). Specifically, the nontarget uniform distribution had a width of 40° and the target distribution had a width of $40^\circ + w$. The difference in distribution width, w , was adjusted for each observer to ensure correct responses on approximately 75% of the trials. In both conditions, the orientations of the texture elements were constrained to be within $135^\circ \pm 45^\circ$ (i.e., $90\text{--}180^\circ$). In other words, the maximum values of d and w were 25° and 50° , respectively.

Procedure

Each trial began with the presentation of a black or white (contrast of -0.39 or 0.39 , respectively) fixation point at the center of the screen. After the fixation point appeared, the observer pressed the space bar to continue. The fixation point was extinguished 547 ms after the key press, then a uniform blank screen was presented for 80 ms, followed by the five frames of texture stimuli (duration = $5 \times 80 = 400$ ms) and 507 ms of mask. After the removal of the mask, observers were required to judge whether or not the target was presented. Observers were aware that the probability of a target being presented on any given trial was 0.5. Auditory feedback indicated whether the observer's response was correct or incorrect, and the fixation point was redrawn 1,000 ms after the response, to signal the beginning of the next trial ([Figure 1A](#)).

Mean orientation and orientation variance conditions were tested in separate blocks, and the order of conditions was counterbalanced across observers. Before the experimental sessions, 75% correct thresholds for each condition were determined based on three to seven training sessions that used the method of constant stimuli. The first experimental session started after making sure that each observer's performance had reached a stable asymptote. We calculated the 75% correct threshold for each training session, and the averaged threshold over the last two training sessions was used as the fixed stimulus level in the experimental sessions. In some cases, however, these stimulus values produced performance that was significantly different from the desired rate of 75% correct. In those cases, we adjusted the stimulus signal-to-noise level and restarted the experimental sessions. Each observer

then participated in three 1-hr experimental sessions of 1,200 trials, with the signal level fixed.

Results

Threshold in the mean orientation condition, defined as the deviation angle, $2d$, yielding 75% correct responses, was 5.0° for observers I.A. and J.M., 12.0° for observer Y.Y., and 8.0° for observer V.A. Threshold in the orientation variance condition, defined as the increment in noise range, w , was 8.0° for observer I.A., 12.0° for observer J.M., 14.0° for observer Y.Y., and 31.0° for observer V.A. The correct response rates in the mean orientation condition in the experimental sessions were close to 75% for all observers: 73.3% for observer I.A., 72.9% for observer J.M., 75.8% for observer Y.Y., and 75.6% for observer V.A. Percentage correct in the orientation variance condition was also close to 75%: 75.1% for observer I.A., 75.6% for observer J.M., 75.3% for observer Y.Y., and 73.3% for observer V.A. Hence, our estimates of threshold produced roughly equal performance across tasks and observers. Moreover, d' was not different between observers, and the response bias measure, β , did not show strong bias in this experiment (see [Tables 1](#) and [2](#)).

Classification images for the two conditions are shown in [Figures 2](#) and [3](#). In these figures, the numbers represent the temporal order of the movie frames: classification images for Frames 1, 2, 3, 4, and 5. The classification image denoted by "1–5" was calculated by first averaging the noise arrays on each trial across temporal frames and then sorting the averaged arrays according to [Equation 1](#) or [2](#). Note that we calculated classification images based on orientations, not contrast. In other words, each frame in the classification image was calculated from the 5×5 arrays of orientation noise presented on that temporal frame. The mean classification images indicate the extent to which the probability of an observer responding "target present" was associated with the orientation of a texture elements. The variance classification images indicate the extent to which the probability of an observer responding "target present" was associated with *deviations* (positive or negative) away from the mean orientation.

Colored pixels in [Figures 2](#) and [3](#) represent spatiotemporal loci where the orientations of texture elements were

Observer	75% Threshold (deg)	% Correct	d'	β
I.A.	8.0	75.1	1.363	0.866
J.M.	12.0	75.6	1.395	0.839
Y.Y.	14.0	75.3	1.369	1.091
V.A.	31.0	73.3	1.247	0.957

Table 2. Seventy-five percent threshold, percentage correct, d' , and bias measure (β) for the sustained target detection in the orientation variance condition in [Experiment 1](#).

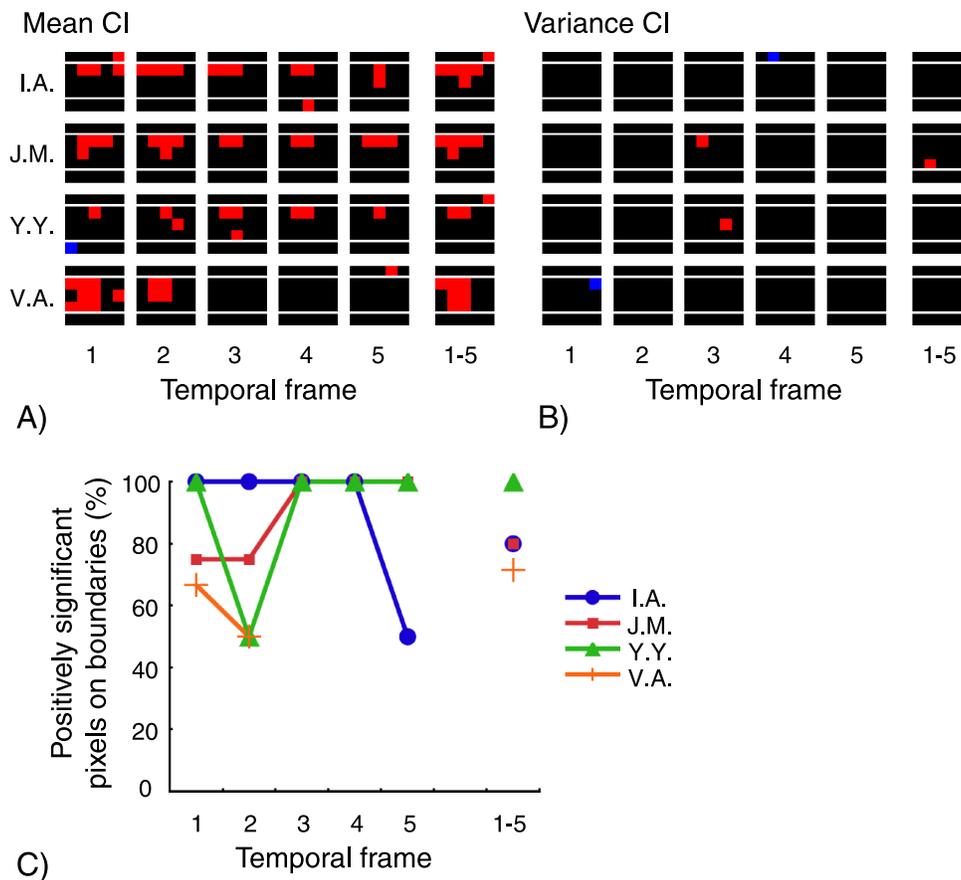


Figure 2. The results for the sustained target detection in the mean orientation condition in Experiment 1. (A) The mean classification images for each time frame: the first to the fifth frames. The rightmost column shows the mean classification image (1–5) averaging across the five temporal frames. (B) The variance classification images for each time frame: the first to the fifth frames. The rightmost column shows the variance classification image (1–5) based on the variance across the five temporal frames. In Panels A and B, the horizontal white lines represent the border between the target and nontarget regions. (C) The proportion of positively significant pixels (red) on the upper and lower rows in the target region out of all the target region in Frames 1, 2, 3, 4, 5, and 1–5. See text for details.

associated significantly with an observer's response. In mean classification images, red and blue pixels indicate significant positive and negative associations, respectively, between an element's steepness (i.e., steepness = $135^\circ - \theta$, where θ is an element's orientation) and the probability of an observer responding "target present". In variance classification images, red pixels indicate spatiotemporal locations where deviations (positive or negative) away from the mean orientation were significantly and positively correlated with the probability of an observer responding "target present", whereas blue pixels indicate spatiotemporal locations where large deviations from the mean orientation were significantly and negatively correlated with "target present" responses. The significance levels were calculated by a permutation test: The responses of each subject were randomly shuffled, and the classification images (mean and variance) were then recalculated for this random permutation of responses. This process was repeated 1,000 times, and the resulting set of classification images was used to estimate the distribution of orientations at each pixel under the null

hypothesis of no association between the observer's response and the element's orientation (Efron & Tibshirani, 1993). These distributions were then used to assess the statistical significance of each pixel in the original classification images. We used the Bonferroni method for controlling Type I error: The alpha level for each test was .002, which corresponds to a Type I error rate that is equal to or less than .05 for each 5×5 texture.

Figure 2A shows that for targets defined by differences in mean orientation, observers I.A., J.M., and Y.Y. were influenced primarily by the orientation of elements in the upper and/or lower rows of the target region (i.e., rows 2 and 4) across all temporal frames, and the upper row is used much more consistently than either the lower or middle row for these three subjects. Figure 2C represents the proportion of positively significant pixels (red) on the upper and lower rows in the target region out of all the target regions at Frames 1, 2, 3, 4, 5, and 1–5 (no symbol means no positively significant pixel at that frame). This representation shows that observers tended to use the boundaries in the target region. In fact, the proportion of significant

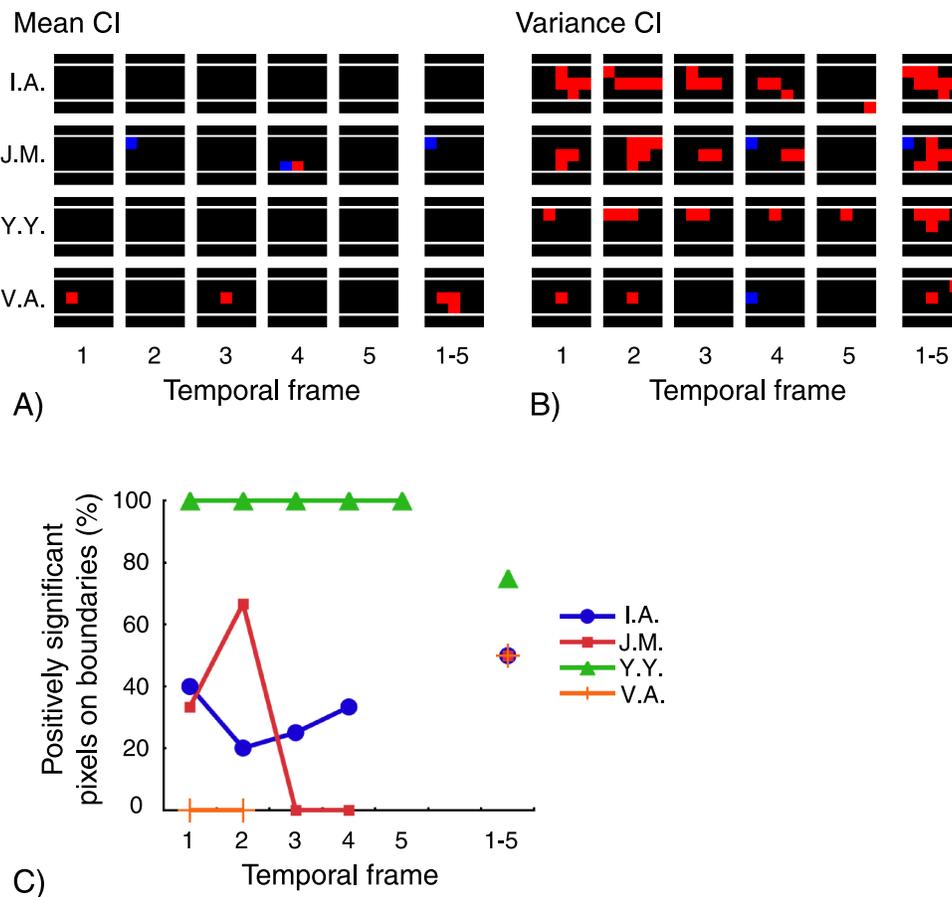


Figure 3. The results for the sustained target detection in the orientation variance condition in [Experiment 1](#). (A) The mean classification images for each time frame: the first to the fifth frames. The rightmost column shows the mean classification image (1–5) averaging across the five temporal frames. (B) The variance classification images for each time frame: the first to the fifth frames. The rightmost column shows the variance classification image (1–5) based on the variance across the five temporal frames. In Panels A and B, the horizontal white lines represent the border between the target and nontarget regions. (C) The proportion of positively significant pixels (red) on the upper and lower rows in the target region out of all the target region in Frames 1, 2, 3, 4, 5, and 1–5. See text for details.

elements within the texture that fell along a boundary was 0.89 (17/19), 0.85 (17/20), and 0.91 (10/11) for observers I.A., J.M., and Y.Y., respectively. There is little evidence that these observers were influenced by the orientation of elements in the background (i.e., rows 1 and 5 in [Figure 2A](#)). Also, the variance classification images suggest that second-order orientation cues did not influence responses in the mean orientation condition ([Figure 2B](#)). [Figure 3B](#) shows that for targets defined by differences in orientation variance, these same observers, I.A., J.M., and Y.Y., were influenced primarily by second-order orientation cues. Moreover, observers I.A. and J.M. were influenced by elements at several locations within the target region (i.e., rows 2–4) on Stimulus Frames 1–4, but observer Y.Y. was influenced only by elements at the upper row. [Figure 3C](#) shows that observers I.A. and J.M. did not have a tendency to use just the boundaries in the target region for second-order stimuli; they used the center row in the target region as well. There is little evidence that these observers were influenced by second-order cues in the background, nor

were they influenced strongly by first-order orientation cues ([Figure 3A](#)) in the orientation variance condition.

Classification images obtained from observer V.A. differed from those obtained from the other observers. In the mean orientation condition, for example, V.A. was influenced by the orientations of elements throughout the target region (i.e., rows 2–4) rather than just the top and bottom target rows ([Figure 2A](#)). Also, unlike the other observers, V.A.'s response in that condition was influenced only by elements displayed in the first two stimulus frames. It is important to note that these differences in these classification images were not related to changes in sensitivity: V.A.'s threshold in the mean orientation condition was approximately equal to the mean threshold of the entire group. In the orientation variance condition, however, observer V.A.'s threshold was much higher than the average threshold obtained from the remaining observers ($w = 31^\circ$ vs. $w_{\text{avg}} = 11.3^\circ$). Interestingly, the classification images measured in this condition do possess structure that may be related to this observer's

poor performance. For example, V.A. used a second-order cue at only a single location and only during the first two movie frames (Figure 3B). Moreover, observer V.A. seems to have relied on a first-order cue to detect the second-order target (Figure 3A). In general, this observer relied on the first-order information and early temporal frames to detect both kinds of targets.

Discussion

When detecting targets in the mean orientation condition, one reasonable strategy is to base responses on the difference between the average orientations of the texture elements in the target and nontarget regions. Alternatively, one might base responses on local comparisons, for example, the difference between orientations of adjacent texture elements on opposite sides of the border separating the target and nontarget regions (Nothdurft, 1992, 1993a, 1993b; for review, see Nothdurft, 1994). Some aspects of our results are consistent with the idea that comparisons are done on elements near the texture border. For example, the mean classification images indicated that three of the four subjects relied on elements at the upper and/or lower border of the target. However, elements in the background did not have a strong influence on observers' decisions, as might be expected if decisions were based on local computations of orientation contrast. Also, observer V.A. was influenced by elements throughout the target region and not just near the target–background border.

It should be noted that a plausible alternative strategy would be to use measures of orientation variance to detect targets in the mean orientation condition. For example, the observer could monitor rows of texture elements on both sides of the target–background edge. The variance of the orientations within this set of elements is larger on target trials than on nontarget trials and could, therefore, serve as a detection cue. Such a strategy would allow the observer to detect the presence, but not the polarity, of a texture-defined edge. Observers seem not to have used such a strategy: The variance classification images do not provide any evidence of an association between the squared deviation of a texture element and an observer's response in the mean orientation condition (Figure 2B).

In contrast, most observers did make effective use of second-order cues in the orientation variance condition. It is interesting to note that the spatial characteristics of the classification images differed slightly across conditions: Specifically, observers I.A. and J.M. were influenced by more pixels in the interior of the target in the orientation variance condition than in the mean orientation condition. This finding suggests that the significance of elements near the target–background border may differ for orientation textures that differ in first-order (mean) or second-order (variance) cues.

Experiment 2: Detection of orientation-defined first- and second-order “flashed” target

Although the classification images in Experiment 1 provide us with a clear indication of the spatial tuning of target detection processes, they do not provide any evidence of temporal tuning: Three of the four observers used information in nearly all stimulus frames. This absence of temporal tuning likely reflects the fact that the stimulus was presented on every temporal frame. In this experiment, therefore, we used a flashed target presentation in which the target signal was presented only in the third frame.

Methods

Observers

Three observers (I.A., Y.Y., and V.A.) who took part in Experiment 1 also participated in Experiment 2. All observers were naive regarding the purpose of the experiment, and they were paid \$10 for each session.

Stimuli and procedures

The stimuli were identical to those in the previous experiment, except that the targets were presented only during the third stimulus frame. In other words, on target trials, the target was presented just on the third frame and nontargets were presented on all other frames. Observers participated in five to eight training sessions prior to the experimental sessions to stabilize performance. As in Experiment 1, detection thresholds for each observer were determined based on performance in these training sessions. Threshold in the mean orientation condition, defined as the deviation angle ($2d$) yielding 75% correct responses, was 13.0° for observer I.A., 21.0° for observer Y.Y., and 31.0° for observer V.A. Threshold in the orientation variance condition, defined as the increment added to the noise range (w), was 20.0° for observer I.A. and 37.0° for observer Y.Y. Observer V.A. was unable to attain 75% correct performance, even with the maximum increment used in our experiment (i.e., $w = 50^\circ$), and was, therefore, not tested further in that condition.

Results

The correct response rates in the mean orientation condition in the experimental sessions for all observers were close to 75%: 74.5% for observer I.A., 75.5% for observer Y.Y., and 74.1% for observer V.A. Percentage correct in the orientation variance condition was 72.8% for observer I.A. and 74.4% for observer Y.Y. Hence, our esti-

Observer	75% Threshold (deg)	% Correct	d'	β
I.A.	13.0	74.5	1.132	0.985
Y.Y.	21.0	75.5	1.390	1.180
V.A.	31.0	74.1	1.314	1.272

Table 3. Seventy-five percent threshold, percentage correct, d' , and bias measure (β) for the flashed target detection in the mean orientation condition in [Experiment 2](#).

mates of threshold produced roughly equal performance across tasks and observers. The threshold, percentage correct, d' , and β values are summarized in [Tables 3](#) and [4](#). As in [Experiment 1](#), the three observers were relatively unbiased and exhibited similar sensitivity.

[Figures 4A](#) and [4B](#) show classification images obtained in the mean orientation condition. As found in [Experiment 1](#), there was little evidence that the orientation of the background elements affected behavior. However, in other respects, the classification images were markedly different from those obtained in [Experiment 1](#). For example, unlike what was found in [Experiment 1](#), the responses of observers I.A. and Y.Y. in the mean orientation condition were influenced by second-order ([Figure 4B](#)) and first-order ([Figure 4A](#)) cues. Although the spatial extent of the influence of second-order cues differed significantly across observers (being much greater in I.A.), the temporal characteristics of the classification images were similar: For both I.A. and Y.Y., the influence of second-order cues extended from Stimulus Frames 1 through 4. Another difference between the current results and those obtained in [Experiment 1](#) is that observers I.A. and V.A. were influenced by first-order cues almost exclusively on the third stimulus frame. In other words, the mean classification images from I.A. and V.A. exhibited very precise temporal tuning. Observer I.A.'s results are particularly striking in this regard: The orientations of 7 of the 15 elements within the target region on Frame 3 were correlated significantly with I.A.'s response. In contrast, observer Y.Y. was influenced by a much smaller set of elements in four of the five frames. It is important to note, however, that even observer Y.Y. exhibited temporal tuning to first-order cues. [Figure 4C](#) shows the temporal dynamics of one element of Y.Y.'s classification image. The horizontal axis in [Figure 4C](#) represents time (in movie frames), and the vertical axis represents noise values (in degrees of orientation) in one pixel (row 2, column 3) in the mean classification images (the only pixel that reaches significance in Frames 2–5 for observer Y.Y.). Here, the positive values represent steeper noise orientations in the mean classification image. The plot peaks at the third frame, showing that observer Y.Y. was influenced most strongly by that element in the third frame. [Figure 4D](#) shows the same analysis for the variance classification image, and one can see that the temporal tuning is much

broader in that condition. Finally, it is interesting to note that V.A., unlike the other two observers, was not influenced by second-order cues in the mean orientation condition ([Figure 4B](#)). This lack of influence, together with the fact that V.A. was not able to obtain 75% correct in the orientation variance condition, is consistent with the idea that this observer was very insensitive to second-order orientation cues.

[Figures 5A](#) and [5B](#) show the results from observers I.A. and Y.Y. for the orientation variance condition. For both observers, significant pixels were found primarily in the variance classification images ([Figure 5B](#)). These pixels were clustered in time, occurring predominantly during the one frame that contained the target (i.e., Frame 3), but were distributed spatially over the entire target. Individual differences were less pronounced as compared with the first-order flashed target condition.

Discussion

The two observers who were sensitive to second-order information (I.A. and Y.Y.) used second-order cues to detect targets in *both* the mean orientation and orientation variance conditions. This result differs from [Experiment 1](#), where observers did not seem to use second-order cues to detect a first-order sustained target. How would a second-order cue aid the detection of a flashed target in the mean orientation condition? As noted in the [Discussion](#) section of [Experiment 1](#), a second-order spatial cue exists for detecting the presence, but not the polarity, of the border between the target and the background. The flashed target presentation used in [Experiment 2](#) also introduced a temporal cue that was not present in [Experiment 1](#). More specifically, because the target was transient, the orientation variance measured across temporal frames at each target element's location was greater on target-present trials than on target-absent trials. Thus, there are two different second-order cues, that is, spatial and temporal, that might help observers detect the first-order flashed target. The influence of the latter second-order cue—temporal—is illustrated nicely by the collapsed variance classification image (far right columns, [Figures 4B](#) and [5B](#)). These classification images are consistent with the idea that observers were influenced by the variation of orientation across stimulus frames.

Observer	75% Threshold (deg)	% Correct	d'	β
I.A.	20.0	72.8	1.212	1.002
Y.Y.	37.0	74.4	1.318	1.147
V.A.	N/A			

Table 4. Seventy-five percent threshold, percentage correct, d' , and bias measure (β) for the flashed target detection in the orientation variance condition in [Experiment 2](#).

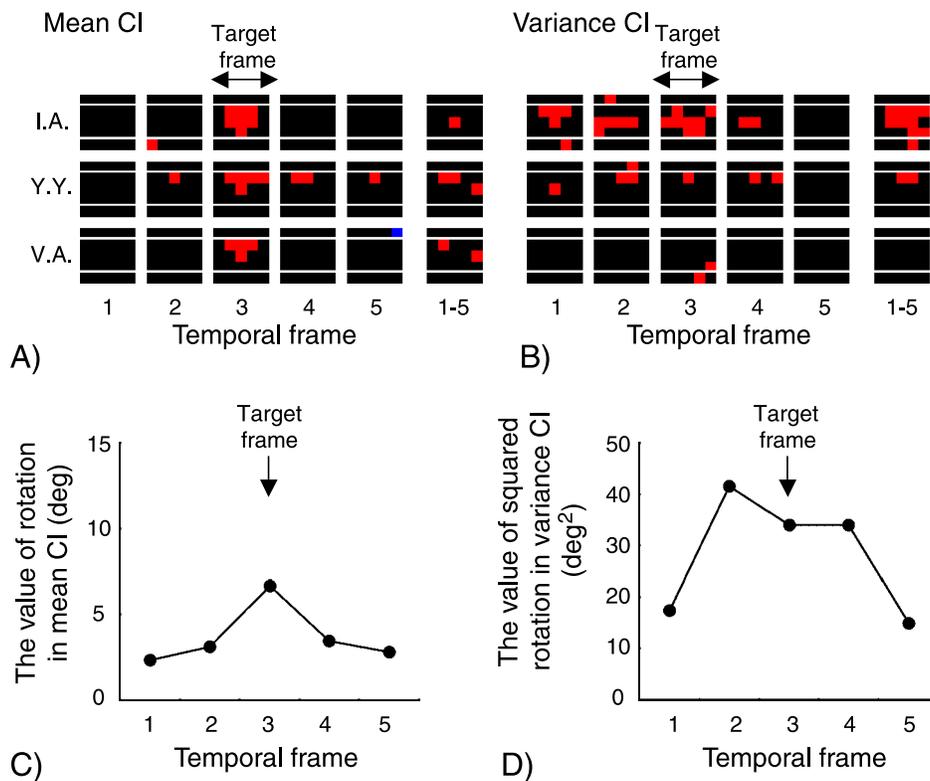


Figure 4. The results for the flashed target detection in the mean orientation condition in Experiment 2. (A) The mean classification images for each time frame: the first to the fifth frames. The rightmost column shows the mean classification image (1–5) averaging across the five temporal frames. (B) The variance classification images for each time frame: the first to the fifth frames. The rightmost column shows the variance classification image (1–5) based on the variance across the five temporal frames. In Panels A and B, the horizontal white lines represent the border between the target and nontarget regions. (C) The time course of the magnitude of the center of the second-row element in the mean classification image from observer Y.Y. The unit in the vertical axis is degree of rotation. (D) The time course of the magnitude of the center of the second-row element in the variance classification image from observer Y.Y. The unit in the vertical axis is squared degree of rotation. See text for details.

Although second-order cues were used to detect signals in both the mean orientation and orientation variance conditions, the variance classification images differed in the two conditions. The influence of second-order cues was spread approximately equally across several stimulus

frames in the mean orientation condition but was concentrated mostly in the third stimulus frame in the orientation variance condition. This difference is particularly striking in observer Y.Y.: The variance classification images contained significant elements in Frames 1–4

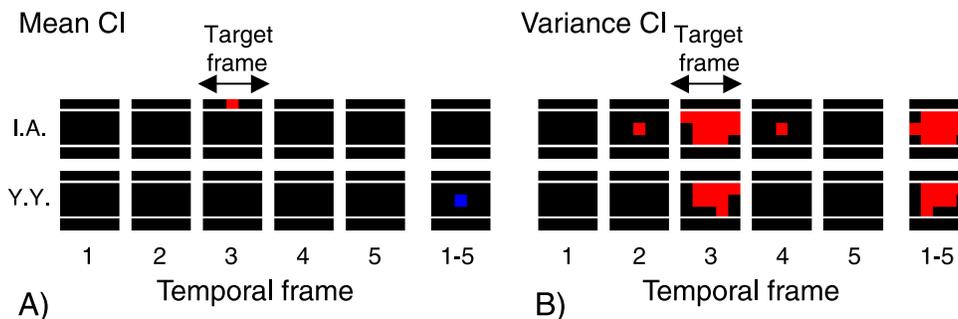


Figure 5. The results for the flashed target detection in the orientation variance condition in Experiment 2. (A) The mean classification images for each time frame: the first to the fifth frames. The rightmost column shows the mean classification image (1–5) averaging across the five temporal frames. (B) The variance classification images for each time frame: the first to the fifth frames. The rightmost column shows the variance classification image (1–5) based on the variance across the five temporal frames. In Panels A and B, the horizontal white lines represent the border between the target and nontarget regions. See text for details.

in the mean orientation condition but only in Frame 3 in the orientation variance condition. We suggest that this difference in the classification images' temporal structure reflects the fact that observers integrated second-order cues across time to detect the mean orientation target but integrated cues across space to detect the orientation variance target. In addition, the fact that we found little evidence that decisions were influenced by background elements is inconsistent with the idea that variance was computed across elements on either side of the target–background border.

A comparison of Experiments 1 and 2 shows that changing from a sustained to a transient target presentation affected observers differently. In Experiment 1, the mean classification image revealed that observer I.A. used a small number of elements distributed across all five temporal frames to detect a first-order target (Figure 2A). In Experiment 2, observer I.A. used a larger number of texture elements, but only during one frame. Thus, changing from a sustained to a transient temporal presentation seems to have resulted in a space–time trade-off in the use of first-order cues by observer I.A. For observers Y.Y. and V.A., however, no such space–time trade-off was found (although observer V.A. did switch from relying on first-order cues in Frame 1 to detect the sustained target to relying on first-order cues in Frame 3 to detect the transient target; see Figures 2A and 4A). Regarding the use of second-order cues, it is harder to discern individual differences in the effect of temporal presentation because none of the observers used second-order cues in Experiment 1 to detect the first-order target, and all but V.A. did so in Experiment 2.

Experiment 3: Flashed target detection with 15 frames

When flashed targets were presented, observers' classification images exhibited clear temporal structure. However, the temporal resolution of the stimulus presentation was relatively coarse; therefore, we conducted an additional experiment with a finer temporal scale to probe the temporal development of the classification image in more detail. Here, we focus only on the mean orientation condition because observers tended to make use of both first- and second-order cues when detecting targets in this condition.

Methods

Observers

Six observers (age range = 20–27 years; mean age = 22.83 years) took part in this experiment. Observer Y.Y. had taken part in Experiments 1 and 2; A.K., K.H., and

B.B. were students at McMaster University, Canada; and I.K. and D.M. were students at the University of Tsukuba, Japan. Observers had normal or corrected-to-normal visual acuity, were naive regarding the purpose of the experiment, and were paid \$10 or its equivalent for each session.

Stimuli and procedures

The stimuli and procedures were the same as those used in the mean orientation condition in Experiment 2, with the following exceptions. The number of temporal frames increased from 5 to 15. Note that the total duration of stimulus presentation was the same as in Experiment 2 but that frame duration was decreased from 80 to 26.67 ms (i.e., $26.67 \text{ ms} \times 15 \text{ frames} = 400 \text{ ms}$ total duration). The target was presented on Frames 7–9; thus, target duration was also the same as in Experiment 2. The data were collected across 10,800 trials distributed across nine test sessions (nine 1,200-trial sessions) after four training sessions. (Due to a computer malfunction, the first 200 trials for observer A.K. were not recorded; hence, the total number of trials for this observer was 10,600.)

Results

Threshold in the first-order task, estimated during the training sessions and defined as the deviation angle, $2d$, yielding 75% correct responses, was 29.0° for observer A.K.; 16.0° for observer Y.Y.; 11.0° for observers K.H., I.K., and D.M.; and 15.0° for observer B.B. The percentage correct during the test sessions was 78.8% for observer A.K., 75.6% for observer Y.Y., 73.3% for observer K.H., 79.7% for observer B.B., 73.5% for observer I.K., and 76.8% for observer D.M. Estimates of threshold, percentage correct, d' , and β are summarized in Table 5.

Classification images are shown in Figure 6. Temporal modulation of the influence of first-order information is indicated by significant variation in the number of significant pixels in the mean classification images for observers B.B., I.K., and D.M. (Figure 6A). However, temporal modulation is not obvious for observers A.K., Y.Y., and K.H. because these observers consistently used a small number of elements in mean classification images, typically in the upper row of the target region, across most temporal frames. However, a closer examination of the data indicates that all observers exhibited temporal tuning. Thus, Figure 7 shows the temporal modulation of the value of one representative element that reached the statistically significant level most often for each observer (row = 2, column = 3 for observers A.K., Y.Y., and K.H.; row = 3, column = 3 for observers B.B. and D.M.; and row = 3, column = 2 for observer I.K.) in the mean (Panel A) and variance (Panel B) classification images. The horizontal axes in Figures 7A and 7B represent

Observer	75% Threshold (deg)	% Correct	d'	β
A.K.	29.0	78.8	1.642	1.438
Y.Y.	16.0	75.6	1.385	1.041
K.H.	11.0	73.3	1.248	0.898
B.B.	15.0	79.7	1.661	1.107
I.K.	11.0	73.5	1.259	0.993
D.M.	11.0	76.8	1.465	0.923

Table 5. Seventy-five percent threshold, percentage correct, d' , and bias measure (β) for the flashed target detection in the mean orientation condition in Experiment 3.

temporal frames, and the vertical axes represent the value of rotation (in degrees) of the mean classification image and the value of squared rotation (in degrees squared) in

the variance classification image, respectively. The use of the first-order cue peaked between the seventh and ninth frames for all observers, coinciding with the actual timing of the target presentation (Figure 7A). Permutation tests confirmed that most values in the first-order image differed from chance ($p < .0034$; family-wise Type I error rate $\leq .05$). The statistically significant frames are shown as circles in Figure 7A. Permutation tests were also performed to evaluate the significance of all pairwise comparisons among the values in the first-order image (family-wise Type I error rate = .05). These tests confirmed that the values on Frames 7, 8, and 9 were different from those on other frames. For observers A.K., Y.Y., and K.H., the values on Frames 7, 8, and 9 differed from values on all other frames. For observer B.B., the values on Frames 7, 8, and 9 differed from those on Frames 1, 5, and 12 and that on Frame 8 differed from that on

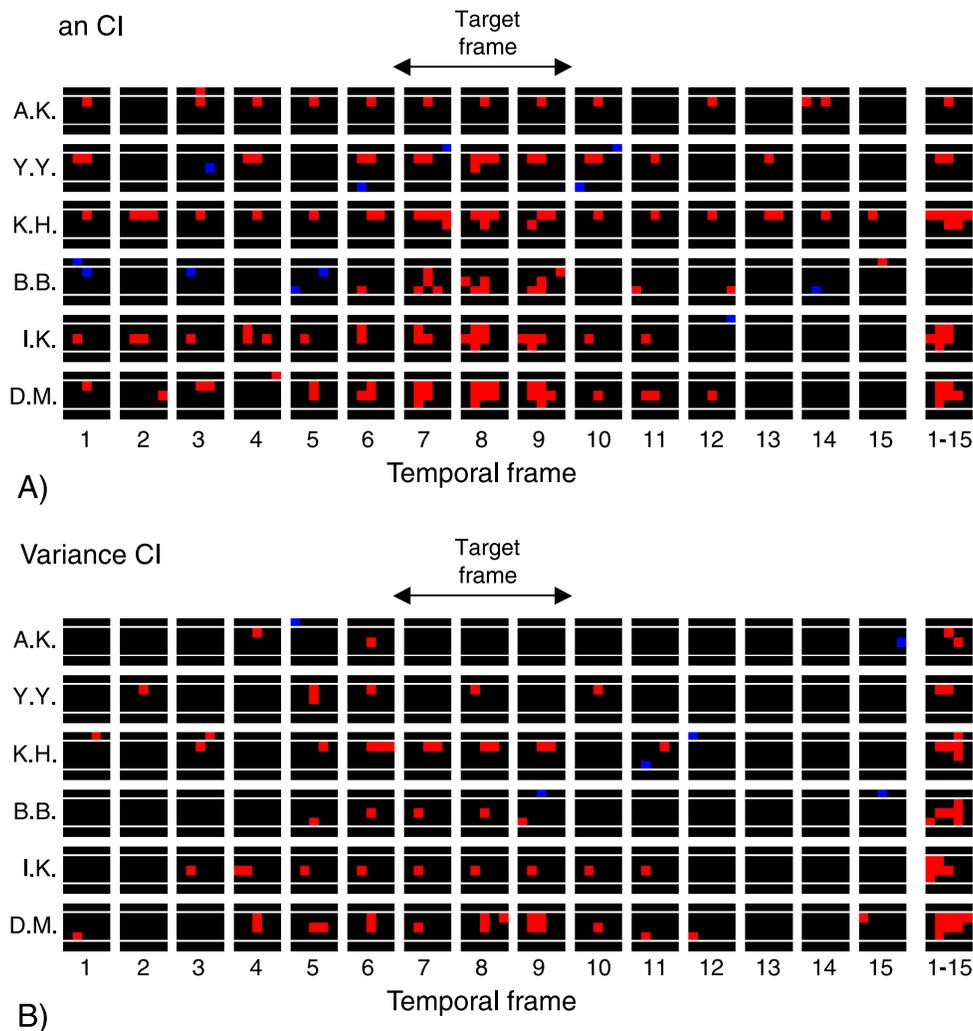


Figure 6. The results for the flashed target detection in the mean orientation condition in Experiment 3. (A) The mean classification images for each time frame: the 1st to the 15th frames. The rightmost column shows the mean classification image (1–15) averaging across the five temporal frames. (B) The variance classification images for each time frame: the 1st to the 15th frames. The rightmost column shows the variance classification image (1–15) based on the variance across the five temporal frames. In Panels A and B, the horizontal white lines represent the border between the target and nontarget regions. See text for details.

Frame 4. For observer I.K., the values on Frames 7 and 8 differed from those on Frames 1, 2, 3, 5, 10, 11, 12, 13, 14, and 15 and that on Frame 9 differed from those on Frames 1, 2, 11, 12, 13, 14, and 15. For observer D.M., the values on Frames 7, 8, and 9 differed from those on Frames 1, 2, 3, 4, 13, 14, and 15; that on Frame 7 differed from those on Frames 5, 10, 11, and 12; and that on Frame 9 differed from those on Frames 5 and 11. None of the other comparisons were significant. Thus, all observers exhibited very clear temporal tuning for the use of first-order cues.

As in [Experiment 2](#), the temporal tuning of second-order cues was different from that of the first-order cues. There was no clear peak between the seventh and ninth frames in variance classification images, except for observer I.K. ([Figure 7B](#)). Instead, the influences of second-order cues were shown as several small peaks prior to and during the target presentation (on Frame 4 for observer A.K., on Frames 2 and 5 for observer Y.Y., on Frames 3 and 9 for observer K.H., on Frames 6 and 8 for observer B.B., on Frames 4 and 7 for observer I.K., and on Frames 5 and 9 for observer D.M.). For observers A.K., Y.Y., and I.K., the influence of second-order cues seemed greater prior to than during the target presentation. Statistically significant frames are shown as circles in [Figure 7B](#) based on permutation tests ($p < .0034$; family-wise Type I error rate $\leq .05$). However, pairwise comparisons among values in the variance classification images did not confirm these peaks prior to and during the target presentation, except that the values on Frame 2 differed from those on Frames 8 and 14 for observer Y.Y. (family-wise Type I error rate $\leq .05$). Thus, we did not obtain clear evidence of strong temporal tuning in the use of second-order cues even for the representative element before the target timing for most of the observers.

Discussion

As in [Experiment 2](#), the variance classification image measured in [Experiment 3](#) suggested that the second-order properties of the noise influenced the observer's decision regarding the presence of a target, defined by a change in mean texture orientation, but that the influence of second-order cues was not closely linked to the temporal properties of the target. In this regard, our results are similar to those reported by Neri and Heeger (2002). In their experiments, observers were required to detect a bright or dark bar flashed briefly against a gray background, when the target and background were embedded in a dynamic visual noise. Neri and Heeger found significant structure in the mean classification image during the presentation of the target and in the variance classification image for 100 ms prior to the presentation of the target. They suggested that the presentation of very bright or very dark noise elements prior to target onset could capture visual attention and thereby influence target detection and that the spatiotemporal structure in the variance classification image reflected the operation of these attentional mechanisms. Interestingly, we also found an influence of second-order cues at 133.35–26.67 ms prior to the presentation of the target, and for three observers, this early influence seemed greater than the influence of second-order cues during the target presentation ([Figure 7B](#)). However, we found no statistical evidence (pairwise comparisons among values in the variance classification images) of clear temporal tuning prior to and during the target presentation in the variance classification image.

Lastly, one interesting point about the individual differences was that two Japanese observers (I.K. and D.M.) seemed to use more central clusters of elements, whereas four Canadian observers (of whom two were of

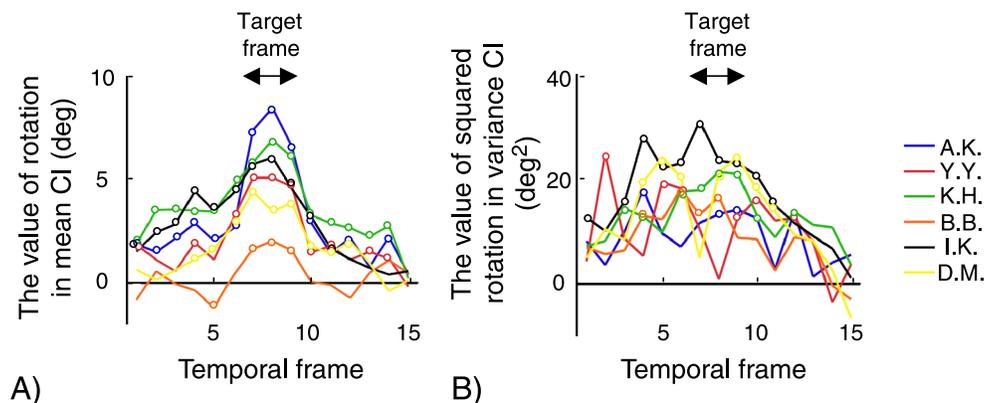


Figure 7. The value of one representative element in classification images. (A) The value of one element in each temporal frame mean classification image. The unit in the vertical axis is degree of rotation. (B) The value of one element in each temporal frame variance classification image. The unit in the vertical axis is squared degree of rotation. The statistically significant frames are shown as circles. (Permutation tests confirmed that most values in the first-order image differed from chance; $p < .0034$; family-wise Type I error rate $\approx .05$.) See text for details.

Asian descent, who have been residing in Canada for more than 3 years) tended to use elements at the edge of the target region. Although it was not the intent of our experiment to investigate cross-cultural differences in visual perception, our results are consistent with recent research in that field (for a review, see Nisbett & Masuda, 2003). People immersed in North American cultures are assumed to be relatively more attuned to a focal object and less sensitive to context. Thus, they might focus on the most critical region—information right at the border in the target region. In contrast, people immersed in Asian cultures are thought to be more attuned to contextual information—namely, information that surrounds the focal object, a broader region than the border. Because of the small number of observers for cross-cultural studies, this finding should be taken as tentative, to be addressed more directly in future research.

General discussion

The salience of local texture differences

Nothdurft (1992, 1993a, 1993b, 1994) has shown that local differences in texture attributes are especially important for segregating regions that differ in first-order textural properties. For example, Nothdurft (1991) proposed a two-stage model in which localized filters first evaluated the difference between the orientations of adjacent elements, and then the local differences were pooled across space. This type of model predicts that elements near the texture–background border should have high weight in the classification image. Some aspects of the present results obtained with first-order orientation textures are consistent with this idea. For example, in [Experiment 1](#), when detecting a sustained target, the decisions of three of four observers were influenced primarily by elements along the texture–background border ([Figure 2A](#)). In [Experiment 2](#), observers Y.Y. and V.A. detected a transient first-order target primarily based on elements along the upper texture–background border ([Figure 4A](#)). However, other aspects of our results are not consistent with Nothdurft’s idea. If observers based their decisions on local computations of orientation contrasts, then background texture elements along the target–background border should have influenced behavior, but we found very little evidence for this in the mean classification images. Furthermore, the decisions of one observer in [Experiment 1](#) (i.e., V.A.) and another in [Experiment 2](#) (i.e., I.A.) were influenced by elements within the interior of the target, and therefore the salience of local orientation contrast in first-order targets varied significantly across individuals. Finally, classification images obtained with second-order textures also indicated that observers were influenced by elements within the

interior of the target, not just along the texture–background border, in both sustained ([Figure 3B](#)) and transient ([Figure 5B](#)) conditions.

Overall, our results do not support the notion that elements near the texture–background border were unusually salient in our task. Of course, our results do not imply that such locations will not be important in other contexts. There are several differences between our stimuli and task and those used by Nothdurft, which may account for the different findings. For instance, Nothdurft used larger arrays of textures (e.g., 12×12) than the ones used in the present experiments. This stimulus difference may be significant because the size of the targets used by Nothdurft was a much smaller fraction of the entire stimulus than that of the targets used here. In other words, the targets used by Nothdurft were embedded in a larger array of background elements. Another potentially important difference between the studies lies in the nature of the task. In this study, observers detected a target. In Nothdurft’s experiments, however, observers typically had to localize a target or discriminate its orientation, which may have led to a greater emphasis being placed on accurately encoding the texture–background border.

We did find that the spatial structure in the classification images depended strongly on the nature of the target. For example, when detecting a first-order sustained target, three of four observers were influenced most by texture elements along the edge of the target region ([Figure 2A](#)). When detecting second-order texture targets, on the other hand, three of four observers were influenced by elements along the border and in the interior of the target region ([Figure 3B](#)). Thus, observers pooled information across many more elements when estimating orientation variance than when estimating mean orientation. This difference between conditions is not inconsistent with previous reports that second-order mechanisms are sensitive to a lower range of spatial frequencies than are first-order mechanisms (e.g., Solomon & Sperling, 1995).

Our findings suggest that the spatial cues used to detect first-order texture differences may depend on the temporal properties of the stimulus. In [Experiment 1](#), which used a sustained target, three of the four observers relied on a few elements near the texture–background border during the entire temporal presentation of the stimulus. In [Experiment 2](#), which used a flashed target, observers tended to use more elements in the interior of the target region during the one frame on which the target was presented. This space–time trade-off is consistent with the idea that observers used different spatiotemporal channels to detect the different first-order targets: Channels with sustained temporal characteristics tend to have small spatial receptive fields, whereas mechanisms with transient temporal characteristics tend to have large spatial receptive fields (Kulikowski & Tolhurst, 1973; Tolhurst, 1975).

Advantages of classification images over other methods

To date, many different techniques have been used to investigate the texture segregation process. Standard psychophysical methods have informed us of the many aspects of the segregation process at the global scale, but these methods are not optimized to reveal local contributions to that process (e.g., the contribution of individual elements in a texture). Moreover, with standard psychophysical methods, even if there were many potential factors other than the manipulated factor, only the effect of the manipulated factor can be interpreted. For example, Nothdurft (e.g., Nothdurft, 1992, 1993b; for review, see Nothdurft, 1994) manipulated the magnitude of local differences in visual attributes (e.g., orientation, color, motion, and luminance) at the border between distinct regions and showed that the border cue is important for visual segmentation. However, these studies did not manipulate nonborder factors and, thus, could not speak of the influence from nonborder regions.

The response classification technique has several advantages over standard psychophysical methods. This technique can show both the local elements' contribution and their global contribution (e.g., in the spatiotemporal distribution of local element contributions). Moreover, this method is not limited by manipulated factors. In response classification, the experimenter simply selects the alternative stimuli and requires observers to discriminate these stimuli. As a result, response classification can visualize any of the local elements' contributions or any factor that is related to the task. These are advantages over the standard psychophysical methods and may explain the seemingly discrepant results between our study and Nothdurft's studies.

Classification images can also show individual differences in the spatiotemporal detection templates that could not be revealed by traditional psychophysical measures of sensitivity (e.g., threshold, percentage correct rate, and reaction time). For example, V.A. showed a relatively high threshold (31.0°) to detect the second-order sustained target in [Experiment 1](#). However, just knowing the value of V.A.'s threshold does not tell us anything about this observer's strategy. V.A.'s elevated threshold may have been due to the use of a different and ineffective strategy, or she may have used the same strategy as other observers, but just less efficiently (cf. Sekuler et al., 2004, for an example of the latter). Response classification enables us to distinguish between these two possibilities: The classification image clearly showed that this observer used an ineffective strategy for detecting the second-order sustained target. Rather than using information in the second-order (variance) template, like other observers did, V.A. relied on less useful information in the first-order (mean) template when detecting the second-order sustained target. Moreover, in [Experiments 2](#) and [3](#), the

classification images for detecting the targets showed individual differences in spatiotemporal tuning that would not have been detected based simply on standard threshold or percentage correct measures.

Previous researchers have taken neurophysiological approaches to the study of texture perception, as well as psychophysical approaches. Neurophysiological studies have been useful in characterizing spatial and temporal processing at the level of individual neurons. For example, Lamme (1995) investigated contextual modulations of V1 neurons in monkeys using motion- and orientation-defined textures. V1 cells responded more strongly when the surrounding area of their receptive fields contained orthogonally moving elements or oriented lines than in the case where a uniform texture was presented. This finding has also been replicated with disparity-, color-, and luminance-defined textures (Zipser, Lamme, & Schiller, 1996). Moreover, the firing rate of V1 neurons is stronger to the border between two texture regions than to homogenous elements within a single region (Nothdurft, Gallant, & Van Essen, 2000). Lastly, Lamme (1995) and Lamme, Rodriguez-Rodriguez, and Spekreijse (1999) have suggested that there are three temporal stages in the responses of V1 neurons. The first stage is a basic orientation tuning at a latency of approximately 60 ms. The second is a border detection stage at a latency around 80 ms. The third is a surface representation stage at about 120 ms.

In the visual system, neurons at different levels/loci have, in most cases, different roles of visual information processing. For example, some V1 neurons show highly specialized receptive field properties: the combination of orientation and spatial frequency (De Valois, Yund, & Helper, 1982) or the combination of color and/or luminance and spatial frequency (Thorell, De Valois, & Albrecht, 1984). In contrast, other neurons in V1 show general modulation characteristics or similar contextual modulation by texture segregation, irrespective of kinds of textures or cues (Zipser et al., 1996). Although the neurophysiological method is suitable for describing local neurons' properties, it is not easy to determine how strongly each neuron's activity contributes to the final decision regarding target detection. For example, one higher visual system function, attention, selects which visual region or object should be processed faster and more profoundly than others. This means that higher visual stages could use each V1 cell's activity differently, depending on its attentional weights (Hopf et al., 2006). Some V1 cells would contribute strongly to the final decision, but others would not. It also remains unclear what the time course of the whole visual segregation processing system is. The response classification technique can clarify some of these issues. In other words, the classification image can help us visualize systematically organized spatiotemporal templates with fine resolution for a specific task.

Conclusion

The classification image technique successfully shows how texture elements localized in space and time contribute to the detection of texture-defined targets and complements studies of figure–ground segregation that use standard psychophysical and neurophysiological methods. Moreover, this technique enables us to clearly visualize individual differences that may not be apparent in more global measures of performance (e.g., thresholds).

Acknowledgments

This study was supported by JSPS-CIHR Joint Health Research Program, NSERC Discovery grants to PJB and ABS, and the Canada Research Chair Program.

Commercial relationships: none.

Corresponding author: Masayoshi Nagai.

Email: masayoshi-nagai@aist.go.jp.

Address: Institute for Human Science and Biomedical Engineering, National Institute of Advanced Industrial Science and Technology, AIST Tsukuba Central 6, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan.

Footnote

¹Classification image calculation (Equation 1) is based on unbiased observers (Murray et al., 2002).

References

- Ahumada, A. J., Jr., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America*, *49*, 1751–1756.
- Beard, B. L., & Ahumada, A., Jr. (1998). A technique to extract relevant image features for visual tasks. *Proceedings of SPIE*, *3299*, 79–85.
- Beck, J. (1966a). Effects of orientation and of shape similarity on perceptual grouping. *Perception & Psychophysics*, *1*, 300–302.
- Beck, J. (1966b). Perceptual grouping produced by changes in orientation and shapes. *Science*, *154*, 538–540. [PubMed]
- Beck, J. (1967). Perceptual grouping produced by line figures. *Perception & Psychophysics*, *2*, 491–495.
- De Valois, R. L., Yund, E. W., & Hepler, N. (1982). The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research*, *22*, 531–544. [PubMed]
- Eckstein, M. P., Shimozaki, S. S., & Abbey, C. K. (2002). The footprints of visual attention in the Posner cueing paradigm revealed by classification images. *Journal of Vision*, *2*(1):3, 25–45, <http://journalofvision.org/2/1/3/>, doi:10.1167/2.1.3. [PubMed] [Article]
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.
- Gold, J. M., Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2000). Deriving behavioural receptive fields for visually completed contours. *Current Biology*, *10*, 636–666. [PubMed] [Article]
- Gold, J. M., Sekuler, A. B., & Bennett, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science*, *28*, 167–207.
- Gosselin, F., Bacon, B. A., & Mamassian, P. (2004). Internal surface representations approximated by reverse correlation. *Vision Research*, *44*, 2515–2520. [PubMed]
- Gosselin, F., & Schyns, P. G. (2003). Superstitious perceptions reveal properties of internal memory representations. *Psychological Science*, *14*, 505–509. [PubMed]
- Hopf, J. M., Boehler, C. N., Luck, S. J., Tsotsos, J. K., Heinze, H. J., & Schoenfeld, M. A. (2006). Direct neurophysiological evidence for spatial suppression surrounding the focus of attention in vision. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 1053–1058. [PubMed] [Article]
- Kandil, F. I., & Fahle, M. (2001). Purely temporal figure-ground segregation. *European Journal of Neuroscience*, *13*, 2004–2008. [PubMed]
- Kulikowski, J. J., & Tolhurst, D. J. (1973). Psychophysical evidence for sustained and transient detectors in human vision. *The Journal of Physiology*, *232*, 149–162. [PubMed] [Article]
- Lamme, V. A. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience*, *15*, 1605–1615. [PubMed] [Article]
- Lamme, V. A., Rodriguez-Rodriguez, V., & Spekreijse, H. (1999). Separate processing dynamics for the texture elements, boundaries and surfaces in primary visual cortex of the macaque monkey. *Cerebral Cortex*, *9*, 406–413. [PubMed] [Article]
- Lee, S. H., & Blake, R. (1999). Visual form created solely from the temporal structure. *Science*, *284*, 1165–1168. [PubMed]

- Mangini, M., & Biederman, I. (2004). Making the ineffable explicit: Estimating the information employed for face classification. *Cognitive Science*, 28, 209–226.
- Moller, P., & Hurlbert, A. C. (1996). Psychophysical evidence for fast region-based segmentation in motion and color. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 7421–7426. [PubMed] [Article]
- Morgan, M., & Castet, E. (2002). High temporal frequency synchrony is insufficient for perceptual grouping. *Proceedings of the Royal Society B: Biological Sciences*, 269, 513–516. [PubMed] [Article]
- Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2002). Optimal methods for calculating classification images: Weighted sums. *Journal of Vision*, 2(1):6, 79–104, <http://journalofvision.org/2/1/6/>, doi:10.1167/2.1.6. [PubMed] [Article]
- Neri, P., & Heeger, D. J. (2002). Spatiotemporal mechanisms for detecting and identifying image features in human vision. *Nature Neuroscience*, 5, 812–816. [PubMed] [Article]
- Neri, P., Parker, A. J., & Blakemore, C. (1999). Probing the human stereoscopic system with reverse correlation. *Nature*, 401, 695–698. [PubMed]
- Nisbett, R. E., & Masuda, T. (2003). Culture and point of view. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 11163–11170. [PubMed] [Article]
- Nothdurft, H. C. (1985). Orientation sensitivity and texture segmentation in patterns with different line orientation. *Vision Research*, 25, 551–560. [PubMed]
- Nothdurft, H. C. (1991). Texture segmentation and pop-out from orientation contrast. *Vision Research*, 31, 1073–1078. [PubMed]
- Nothdurft, H. C. (1992). Feature analysis and the role of similarity in preattentive vision. *Perception & Psychophysics*, 52, 355–375. [PubMed]
- Nothdurft, H. C. (1993a). The conspicuousness of orientation and motion contrast. *Spatial Vision*, 7, 341–363. [PubMed]
- Nothdurft, H. C. (1993b). The role of features in preattentive vision: Comparison of orientation, motion, and color cues. *Vision Research*, 33, 1937–1958. [PubMed]
- Nothdurft, H. C. (1994). Common properties of visual segmentation. In G. R. Bock & J. A. Goode (Eds.), *Higher-order processing in the visual system, CIBA Foundation Symposium, 184* (pp. 272–286). England: Wiley. [PubMed]
- Nothdurft, H. C., Gallant, J. L., & Van Essen, D. C. (2000). Response profiles to texture border patterns in area V1. *Visual Neuroscience*, 17, 421–436. [PubMed]
- Sekuler, A. B. (1990). Motion segregation from speed differences: Evidence for nonlinear processing. *Vision Research*, 30, 785–795. [PubMed]
- Sekuler, A. B., & Bennett, P. J. (2001). Generalized common fate: Grouping by common luminance changes. *Psychological Science*, 12, 437–444.
- Sekuler, A. B., Gaspar, C. M., Gold, J. M., & Bennett, P. J. (2004). Inversion leads to quantitative, not qualitative, changes in face processing. *Current Biology*, 14, 391–396. [PubMed] [Article]
- Solomon, J. A. (2002). Noise reveals visual mechanisms of detection and discrimination. *Journal of Vision*, 2(1):7, 105–120, <http://journalofvision.org/2/1/7/>, doi:10.1167/2.1.7. [PubMed] [Article]
- Solomon, J. A., & Sperling, G. (1995). 1st- and 2nd-order motion and texture resolution in central and peripheral vision. *Vision Research*, 35, 59–64. [PubMed]
- Thorell, L. G., De Valois, R. L., & Albrecht, D. G. (1984). Spatial mapping of monkey V1 cells with pure color and luminance stimuli. *Vision Research*, 24, 751–769. [PubMed]
- Tolhurst, D. J. (1975). Reaction times in the detection of gratings by human observers: A probabilistic mechanism. *Vision Research*, 15, 1143–1149. [PubMed]
- Tse, P. U., Sheinberg, D. L., & Logothetis, N. K. (2003). Attentional enhancement opposite a peripheral flash revealed using change blindness. *Psychological Science*, 14, 91–99. [PubMed]
- Zipser, K., Lamme, V. A., & Schiller, P. H. (1996). Contextual modulation in primary visual cortex. *Journal of Neuroscience*, 15, 7376–7389. [PubMed] [Article]