Spatial profiles of local and nonlocal effects upon contrast detection/discrimination from classification images

Steven S. Shimozaki
Department of Psychology, University of California, Santa Barbara, CA, USA

Miguel P. Eckstein
Department of Psychology, University of California, Santa Barbara, CA, USA

Craig K. Abbey
Department of Psychology, University of California, Santa Barbara, CA, USA

We used classification images (A. J. Ahumada, Jr., & J. Lovell, 1971) to estimate the perceptual filter in a task designed to assess both local and nonlocal effects upon contrast detection/discrimination. Three observers performed a yes/no detection or discrimination task of a uniform circular decrement (radius = 0.68 deg) near threshold presented for 100 to 400 ms. Stimuli were presented in ring image noise that either covered the signal and an annular surrounding area (out to 1.36 deg), or only the surrounding annular area (out to 1.36 deg). Both the signal and the annular surround appeared on a uniform background. With ring noise over both the signal and surround, the amplitudes of the classification images in the signal area decreased as radial distance increased from the signal/surround border, and no effect of the surround was found. With ring noise only in the surround, classification images indicated noncontiguous effects at both the signal/surround border (local) and the surround/background border (nonlocal). The spatial extents of the nonlocal effects (< 0.07 deg) were smaller than local effects (0.25 deg), whereas the peak amplitudes of the local and nonlocal effects were comparable. These results suggest that the nonlocal effects were smaller than the local effects, and that the smaller effects would be due to smaller effective areas, as opposed to smaller amplitudes over the same area. Little or no change was found in the classification images across stimulus duration, suggesting that both the local and nonlocal processes found in this study were completed within 100 ms.

Keywords: contrast detection, contrast discrimination, classification images

Introduction

Local and nonlocal brightness induction effects

Brightness can be defined generally as the perceived luminance of a stimulus, but it is clear that human observers, when judging brightness, do not operate as simple radiometric devices. The percept of brightness may be induced to change by stimuli other than the target stimulus to be judged. A well-known influence is local contrast, in which a stimulus appears darker when it is adjacent to a brighter stimulus, and vice versa. (e.g., Wallach, 1948; Heinemann, 1955, 1972; Shapley, 1986; Whittle & Challands, 1969). Another example that could be described as a contrast-driven brightness effect is the classic Craik-Cornsweet-O’Brien (CCO) illusion (see Figure 1) (Craik, 1966; Cornsweet, 1970; O’Brien, 1959). Generally, a luminance border determines the brightness of both regions, such that the region nearest the negative contrast border (left) appears darker than that the region nearest the posi-
tive contrast border (right), extending to the parts of the regions with equal luminance across the border.

Aside from local effects upon brightness, there are also nonlocal effects from stimuli not directly adjacent to the stimulus to be judged. Figure 2 is an example of what is typically known as brightness assimilation from Shapley and Reid (1985) (Helson, 1963; Arend, Buehler, & Lockhead, 1971; Shapley, 1986; Reid & Shapley, 1988; Rudd & Arrington, 2001; Hong & Shevell, 2004). Both the disks and the surrounds on the left and right sides have the same luminance; and, therefore, the local contrasts are the same on both sides. Thus, any difference in brightness in the disks is due to the variation in the background luminances, with the left disk typically appearing brighter (however, for counter-examples, see Hong & Shevell, 2004). Classically, assimilation describes a process in which an object's color shifts toward (‘takes on’ the appearance of its surroundings; in this case, the disks appearing more like their adjacent surrounds) (Helson, 1963; Kanizsa, 1979; Hurvich, 1981). Reid and Shapley (1988) choose a different definition, in which assimilation describes "the long-distance interactions which tend to add the apparent brightness of a region to the brightness of regions adjacent to it."

As implied by their definition, Reid and Shapley (1988) suggest that assimilation can be modeled by a weighted linear integration of local contrast effects, with decreasing influence with increasing distance. Rudd and Arrington (2001) propose a somewhat different model in which more distal effects are modulated (‘blocked’) by the amplitudes of intervening contrast borders. Also, a general class of models of brightness perception have been proposed based upon the convolution of the stimulus with a single filter, such as a difference-of-Gaussian (Heinemann & Chase, 1995), or several filters across a range of spatial scales, with the shapes of a difference-of-Gaussians (Moulden & Kingdom, 1991; Blakselee & McCourt, 1997), the 2nd derivative of Gaussians (Kingdom & Moulden, 1992), and an oriented difference-of-Gaussians (Blakselee & McCourt, 1999, 2001).

Studies of the dynamics of brightness induction effects across paradigms tend to support a time course of about 50 to 200 ms (e.g., DeValois, Webster, De Valois, & Lingelbach, 1986; Paradiso & Nakayama, 1991; Rossi & Paradiso, 1996; Paradiso & Hahn, 1996; McCourt & Foxe, 2004). Several authors have suggested a retinotopic ‘filling-in’ mechanism for brightness perception (e.g., Grossberg & Todorovic, 1988; Pessoa, Mingolla, & Neumann 1995), and specifically for both the CCO illusion (Paradiso & Nakayama, 1991; Rossi & Paradiso, 1996; Paradiso & Hahn, 1996) and assimilation (Rudd & Arrington, 2001) effects. In these cases, it is assumed that the brightness percept begins at the contrast borders, and spreads retinotopically to the adjoining areas.

In this study, we assess local and nonlocal effects for a relatively simple contrast detection/discrimination, which may or may not reflect brightness judgments. While possibly suggested by local contrast and CCO illusions, a priori local contrast judgments need not have any direct relation to judgments of perceived luminance (i.e., brightness). Also, contrast judgments plausibly might begin with center-surround or retinal gain control mechanisms in the retina (Enroth-Cugell & Robson, 1966; Shapley & Enroth-Cugell, 1984; Shapley, 1986; Fiorentini, Baumgartner, Magnusson, Schiller, & Thomas 1990), whereas brightness judgments appear to arise from processing in early visual cortical areas (e.g., DeValois et al., 1986; Blakselee & McCourt, 2003; McCourt & Foxe, 2004). Aside from a possible physiological differentiation, Arend and Spehar (1993a, 1993b) have demonstrated that contrast, brightness, and lightness (generally, perceived achromatic surface reflectance) judgments are psychophysically separable, given certain viewing (usually more complex) conditions and observer instructions (also, Cornsweet & Teller, 1965; Sparrock, 1969; Guth, 1973). The interaction and separability of these three perceptions is a topical issue, as well as the question of which percept might be dominant in any given situation (Blakselee & McCourt, 2003; Gilchrist & Economou, 2003; Kingdom, 2003).

Assessing spatial profiles with classification images

Furthermore, in this study, we attempted to directly assess the spatial profile and time course of effects upon contrast detection/discrimination with a technique known as classification images. Classification images were first developed by Ahumada and Lovell (1971) in the field of audition, and have only recently been applied to vision research (Ahumada, 1996). Generally, classification images are generated in tasks with image noise by correlating response outcomes with the particular image noise samples leading to those outcomes. This technique has been shown to be uniquely effective in directly determining the spatial weighting of information that the observer uses to make a
particular judgment, or roughly speaking, the observer's perceptual 'template' or 'filter.' For example, in recent years, classification images have been used to assess several different visual tasks, such as Vernier acuity (Ahumada, 1996; Beard & Ahumada, 1998) and other position discrimination judgments (Levi & Klein, 2002; Li, Levi, & Klein, 2004), depth (Neri, Parker, & Blakemore, 1999), Kanizsa squares and other illusory contours (Gold, Murray, Bennett, & Sekuler, 2000), and visual attention (Eckstein, Shimozaki, & Abbey, 2002; Eckstein, Pham, & Shimozaki, 2004).

In a typical classification image study, stimuli are presented in image noise, and the image noise fields associated with each possible stimulus/response outcome are then pooled and averaged. For example, in a yes-no detection or discrimination task of an arbitrary signal, there are four possible outcomes for a trial: hit (responding "signal present" when the signal is present), miss (responding "signal absent" when the signal is present), false alarm (responding "signal present" when the signal is absent), and correct rejection (responding "signal absent" when the signal is absent). When an observer makes a false alarm, it is assumed that the error is partly due to some random perturbation of the sample of image noise in that trial leading the observer to say "signal" when the signal was not present. Thus, the classification image for false alarms (the averaged noise starting from 0.68 and extending radially out to 1.36 deg from the center) was followed immediately by a 100-ms high-contrast Gaussian white noise mask (SD = 9.38 cd/m²). After responding

General methods

Three observers (female, aged 20 to 22 years, with normal or corrected-to-normal visual acuity) participated in a yes/no detection or discrimination task of a signal presented for 100 ms, 200 ms, or 400 ms. The signal was a uniform circular decrement disk near threshold (radius = 0.68 deg) on a uniform background (luminance = 30.0 cd/m²) presented on 50% of the trials. Decrements were chosen to avoid any possible percept of the signal as being self-luminous (Bonato & Gilchrist, 1994; Heggelund, 1974). Also, previous research has shown that decrements have stronger local contrast effects (Heinemann, 1955), and other studies have shown differences in detection (Boynton, Ikeda, & Stiles, 1964; Patel & Jones, 1968; Short, 1966) and discrimination (Whittle, 1986) thresholds for decrements and increments, with decrements being slightly better.

The signal was presented in rotationally invariant (i.e., circularly symmetric) ring image noise, extending radially 1.36 deg from the center. In Experiment 1, the noise covered both the central signal region and an annular region extending 0.68 deg (1.36 - 0.68 deg) beyond the center signal region. In Experiment 2, the noise only covered the annular region extending 0.68 deg (1.36 - 0.68 deg) beyond the center signal area. Each uniform ring of noise was approximately one pixel (0.034 deg) in width. The stimulus was followed immediately by a 100-ms high-contrast Gaussian white noise mask. After responding

Figure 3. The stimuli for Experiment 1, a yes/no contrast decrement discrimination task with ring noise over signal and surround. The pedestal with no signal appears on the left, and the pedestal with the signal appears on the right. The contrast difference depicted here is larger than that used in the experiment.

Figure 4. Stimuli for Experiment 2, a contrast decrement detection task with ring noise over the surrounding area only. The no signal stimulus appears on the left, and the signal stimulus appears on the right. The contrast difference depicted here is larger than that used in the experiment.
about signal presence with a key press, observers were given feedback on whether their responses were correct or incorrect. Contrasts of the pedestal (only in Experiment 1) and of the signals (in both Experiments 1 and 2) were chosen through pilot studies to achieve d’ between 1.0 and 2.0 for all observers and conditions.

Each observer performed in blocks of 100 trials with the same stimulus duration, and the order within a set of three blocks (one for each stimulus duration) was randomized. Behavioral results from each block were converted from hit and false alarm rates to values of d’ using a standard signal detection theory transformation assuming equal variances (Green & Swets, 1974). Classification images were calculated from taking the mean of the noise fields of the trials generating a particular outcome (hit, miss, false alarm, or correct rejection). For Experiment 1, only the classification images for the false alarm trials are presented; for Experiment 2, the classification images were combined across each outcome using a simple combination rule (Beard & Ahumada, 1998; Murray et al., 2002; Ahumada, 2002). This was done for technical reasons regarding potential biases in Experiment 1 due to the presence of the noise over the signal. Statistical tests on the overall classification images were performed with the Hotelling T² statistic, which is the multivariate equivalent of a t statistic (Harris, 1985; Abbey & Eckstein, 2002; Eckstein et al., 2002; Eckstein et al., 2004). Single-sample Hotelling T² tests were done to assess the existence of significant (nonzero) classification images; independent two-sample Hotelling T² tests (Bonferroni-corrected) were performed to assess differences in the classification images for different stimulus durations (see Appendix A for details). Single-sample t tests (Bonferroni-corrected for two-tailed α = 0.5) were used to test individual radii (pixels) against values of zero.

Stimuli were presented in a darkened room on a monochrome monitor with a viewing size of 32.51 x 24.38 cm and a resolution of 1024 x 768 pixels (Image Systems Corp., Minnetonka, MN), sitting 50 cm from the observer. At this distance, each pixel subtended 0.034 deg of visual angle. Luminance calibrations were performed with software and equipment from Dome Imaging Systems, Inc. (Luminance Calibration System, Waltham, MA).

### Experiment 1: Noise over signal and surround

In this experiment, the ring image noise covered both the signal area and the annular area surrounding the signal (see Figure 3).

### Methods

The three observers participated in a yes/no contrast decrement discrimination task; the signal was a circular contrast decrement (contrast = 7.81%) on a pedestal (decrement, contrast = 31.2%). The extent of the ring image noise was 1.36 deg, twice the radius of the signal/pedestal (0.68 deg), which was equal to a radius of 40 pixels (20 for the surround, 20 for the signal/pedestal). The distribution of luminances for the cross-section of ring noise was Gaussian (SD = 3.52 cd/m²). Each observer performed in 3100 (OC) or 4000 (KC, AH) trials for each stimulus duration. As mentioned earlier, only the false alarm trials were used to construct the classification images. The number of false alarm trials for the 100-, 200-, and 400-ms stimulus durations, respectively, were AH - 606, 628, and 560; KC - 664, 751, and 536; OC - 407, 449, and 442.

### Results and discussion

Figure 5 depicts the d’ values in Experiment 1 for the three observers. For KC, d’ increased slightly with increasing stimulus duration [F(1, 225) = 7.683, p = .006, MSE = 23.71]; the other observers’ performances were not significantly changed with stimulus duration. Also, both AH [F(1, 225) = 13.38, p < .0001, MSE = 40.23] and KC [F(1, 225) = 24.43, p < .0001, MSE = 40.23] had slightly better overall performance than OC.

![Figure 5](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933507/) Behavio ral performance for Experiment 1. Values of d’ for three observers and three stimulus durations (100 ms, 200 ms, and 400 ms). Error bars are SEM.

Table 1 summarizes the statistical results by Hotelling T² for the overall classification images. The first three sections give the results for tests of nonzero classification images, and the last three sections summarize the results for differences between classification images. The graphs on the left of Figure 6 show the classification images in Experiment 1 from the false alarm trials only for the three observers. The classification images are presented along a single dimension, with the x-axis representing the radial distance from the center of the signal. The position of the signal (starting in the center and extending to 0.68 deg) is indicated by the red line segment below the x-axis. The y-axis represents the amplitude of the classification image, which roughly corresponds to the weighting of information at each radius for the observer. The graphs on the right of Figure 6 summarize the single-sample t statistics for nonzero values of the individual radii of the classification images. The red dashed lines indicate criterion t values for significant differences from zero, Bonferroni-corrected.
For all observers, the classification images were all significantly different from zero, and there was no effect of stimulus duration (Table 1). Their classification images all showed a similar pattern; larger amplitudes were found at the signal/surround border, and the amplitudes dropped to nearly zero closer to the center (Figure 6, left). As shown in the right graphs of Figure 6, the individual radii with non-zero values in the signal area extended from the signal/surround border inward to a radial distance of about 0.30 deg from the center. Except for KC, and OC with the 100-ms stimulus duration, there is no evidence of any effect of the surrounding area. For KC, there is a clear local contrast effect, in which an increment adjacent to the signal led to a percept of a decrement signal.

In this situation with the ring noise, an ideal observer uses only the one-dimensional values for each radius, and ignores the irrelevant information in the surrounding area. Therefore, the ideal observer would be represented in

<table>
<thead>
<tr>
<th>100 ms vs. zero</th>
<th>200 ms vs. zero</th>
<th>400 ms vs. zero</th>
<th>100 ms vs. 200 ms</th>
<th>100 ms vs. 400 ms</th>
<th>200 ms vs. 400 ms</th>
</tr>
</thead>
<tbody>
<tr>
<td>df num</td>
<td>df dem</td>
<td>F</td>
<td>p value</td>
<td>df num</td>
<td>df dem</td>
</tr>
<tr>
<td>AH</td>
<td>40</td>
<td>566</td>
<td>9.129</td>
<td>&lt;.0001</td>
<td>AH</td>
</tr>
<tr>
<td>KC</td>
<td>40</td>
<td>624</td>
<td>8.326</td>
<td>&lt;.0001</td>
<td>KC</td>
</tr>
<tr>
<td>OC</td>
<td>40</td>
<td>367</td>
<td>3.716</td>
<td>&lt;.0001</td>
<td>oc</td>
</tr>
<tr>
<td>AH</td>
<td>40</td>
<td>1193</td>
<td>0.752</td>
<td>0.870</td>
<td>AH</td>
</tr>
<tr>
<td>KC</td>
<td>40</td>
<td>1125</td>
<td>1.052</td>
<td>0.384</td>
<td>KC</td>
</tr>
<tr>
<td>OC</td>
<td>40</td>
<td>850</td>
<td>0.784</td>
<td>0.853</td>
<td>OC</td>
</tr>
<tr>
<td>AH</td>
<td>40</td>
<td>1147</td>
<td>0.738</td>
<td>0.885</td>
<td>AH</td>
</tr>
<tr>
<td>KC</td>
<td>40</td>
<td>1066</td>
<td>0.534</td>
<td>0.992</td>
<td>KC</td>
</tr>
</tbody>
</table>

Table 1. F values and p values for overall classification images from Hotelling $T^2$ statistic for Experiment 1. Tests versus zero: single-sample Hotelling $T^2$ tests against the null classification image (all zeroes). Tests of differences for different stimulus durations: independent two-sample Hotelling $T^2$ tests. See Appendix A for details.
Figure 6 (left) as a step function, with a constant negative value in the signal area, and a zero value in the surrounding area. The human observers, however, did not have a flat function in the signal area, and tended to weight information more heavily at the edge at the signal. One possibility is that the observers were using a suboptimal strategy of weighting of information by the two-dimensional image (i.e., the increasing area, or equivalently, the number of pixels with increasing radius), even though only the single dimension of radial distance was relevant.

**Experiment 2: Noise surrounding target only**

In Experiment 1, little or no effect of the surround was found. In Experiment 2, to emphasize effects of the surround, the image noise appeared only in the surround, and not over the signal (see Figure 4). Also, the high signal/surround border contrast in Experiment 1 might have masked effects of the surround. Therefore, to reduce the signal/surround border contrast, the task was changed from a discrimination to a detection task. Another potential factor was that only false alarm trials were used, for technical concerns related to calculating classification images.

As in this study, these concerns were not relevant, allowing the use of all trials in calculating classification images. Lastly, it was thought that the variations in the energy of the image noise might have masked surround effects; in Experiment 2, the image noise energy was normalized across all trials.

**Methods**

The three observers performed a yes/no contrast decrement detection task in which the rings of image noise surrounded only the signal, and did not appear over the signal. Because this was a detection task without noise over the signal, the contrast of the signal needed was less than the discrimination task in Experiment 1 (contrast = 3.1%). As in Experiment 1, the image noise extended out to a radial distance of 1.36 deg (twice the radius of the signal). The cross-section of the radial noise was first Gaussian-distributed; then, each noise field’s cross-sectional contrast energy was normalized [to $E = 191.7$ (cd/m^2)]. Analysis by Nykamp and Ringach (2002) suggest that such a deviation from Gaussian noise does not bias the expected values derived from classification images; also, we performed simulations under these conditions that confirmed that the classification images were unbiased by the normalized image noise. Each observer performed 4000 trials for each stimulus duration. As mentioned in the General methods, the classification images were constructed by combining across all outcomes with a simple combination rule. Some trials (mostly by OC) were lost due to uncoded (inappropriate) keypresses. The actual number of total trials for the 100-, 200-, and 400-ms stimulus durations, respectively, were AH-4000, 4000, and 4000; KC - 3996, 3999, and 4000; and OC - 3948, 3963, and 3917.

**Results and discussion**

Figure 7 summarizes the $d'$ values in Experiment 2 for the three observers. For KC and OC, the results did not improve significantly with stimulus duration. For AH, a slight improvement [$F(1, 225) = 9.55$, $p = .0022$, MSE = 23.71] was found with stimulus duration, and her performance was better overall than the other observers [AH vs. KC, $F(1, 225) = 35.05$, $p < .0001$, MSE = 40.23; AH vs. OC, $F(1, 225) = 124.6$, $p < .0001$, MSE = 40.23]. Also, KC's overall performance was better than OC’s overall performance [$F(1, 225) = 27.48$, $p < .0001$, MSE = 40.23].

Table 2 summarizes the Hotelling $T^2$ statistics for the overall classification images. The graphs on the left side of Figure 8 show the classification images averaged across all trials for the three stimulus durations in Experiment 2, and the graphs on the right side of Figure 8 summarize the single-sample $t$ statistics for nonzero values for individual radii in the classification images. The red dashed lines in these graphs indicate criterion $t$ values for significant differences from zero, Bonferroni-corrected.

All classification images were significantly different from zero, and there were no significant changes in the classification images with stimulus duration (Table 2). For AH, the difference in classification images for the 100-ms and 400-ms stimulus durations was near significance at $p = .036$, but did not meet the Bonferroni-corrected $\alpha$ for 3 comparisons ($\alpha = .05/3 = .0167$). As shown in Figure 8, all observers had a local effect at the signal/surround border, in which an increment in the surround led to signal present (decrement) responses. Also, all observers had a nonlocal effect at the signal/background border, in which an increment in the surround also led to target present responses. Classification image amplitudes in an intermediate band of the surround (between the signal/surround and the surround/background borders) did...
nonlocal
local
Figure 7. Tests versus zero:

(36x57)

round/background border was smaller, less than 0.07 deg spatial extent of the nonlocal effect at the surround approximately 0.25 deg across the observers. Conversely, the values in the classification images at that border, was estimated by the range of contiguous significant nonzero extent of the local effect at the signal/surround border, as not differ significantly from zero. In general, the spatial extent of the local effect at the signal/surround border, as estimated by the range of contiguous significant nonzero values in the classification images at that border, was approximately 0.25 deg across the observers. Conversely, the spatial extent of the nonlocal effect at the surround/background border was smaller, less than 0.07 deg (Figure 8, right). For OC, smaller spatial extents for signifi-
cant local effects were found (particularly for the 400-ms stimulus duration, in which no significant local effect was found). This difference in results appears to be due to lower amplitudes in her classification images, which could be a reflection of her lower performance in the task (Figure 7).4

Figure 8. Classification images (left) and single-sample t values for the classification images (right) for Experiment 2. The x-axes represent the radial distance from the center of the signal, in degrees of visual angle; the bars labeled “Signal” represent the signal area. Classification images (left). The y-axis represents the amplitude of the classification images, which corresponds to the weighting of information for a particular radial distance. The error bars are SEM. t values (right). The y-axis represents the single-sample t values against zero for individual radii of the classification images. The dashed red lines indicate significant differences from zero at an uncorrected two-tailed α = .0025, giving a two-tailed α = .05 across all comparisons, Bonferroni-corrected for 20 comparisons (total radius of noise in the surround in number of pixels = 20). Degrees of freedom for the 100-, 200-, and 400-ms stimulus durations, respectively, were AH – 3999, 3999, and 3999; KC – 3995, 3998, and 3999; and OC – 3947, 3962, and 3916.

not differ significantly from zero. In general, the spatial extent of the local effect at the signal/surround border, as estimated by the range of contiguous significant nonzero values in the classification images at that border, was approximately 0.25 deg across the observers. Conversely, the spatial extent of the nonlocal effect at the surround/background border was smaller, less than 0.07 deg (Figure 8, right). For OC, smaller spatial extents for signifi-

cant local effects were found (particularly for the 400-ms stimulus duration, in which no significant local effect was found). This difference in results appears to be due to lower amplitudes in her classification images, which could be a reflection of her lower performance in the task (Figure 7).4

All observers’ classification images had near-zero or negative values for the first radius in the surround at the signal/surround border. This was likely caused by location

<table>
<thead>
<tr>
<th>Stimulus Duration (ms)</th>
<th>100</th>
<th>200</th>
<th>400</th>
</tr>
</thead>
<tbody>
<tr>
<td>OC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AH</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KC</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. F values and p values for overall classification images from Hotelling $T^2$ statistic for Experiment 2. Tests versus zero: single-sample Hotelling $T^2$ tests against the null classification image (all zeroes). Tests of differences for different stimulus durations: independent two-sample Hotelling $T^2$ tests. See Appendix A for details.
uncertainty regarding the position of the border, in which part of the surround was confused as part of the signal. OC seemed particularly affected, with highly significant negative values for the first radius pixel across all stimulus durations.

In comparing the individual radii with the largest amplitudes for the local and nonlocal effects, it was found that the peak amplitudes for the local effects for AH were significantly greater than those for her nonlocal effects, by a factor of about 2-3 (two-sample t tests: 100 ms, t(7998) = 3.22, p = .0013; 200 ms, t(7998) = 2.23, p = .026; 400 ms, t(7998) = 2.32, p = .020). For OC, the peak amplitudes for the local contrast effect tended to be slightly greater (by a factor of 1.03 to 1.45), but were not significantly different from the peak amplitudes for the nonlocal effect. For OC, the peak amplitudes of the local and nonlocal effects did not differ significantly, with the local effects having about equal (100 ms, 200 ms) or tending to smaller (400 ms) peak amplitudes than the nonlocal effects.

An ideal observer would not use the information in the surrounding area, as it was irrelevant to the task; therefore, an ideal observer would have no classification image in this task (represented as all zeroes in the graphs on the left side of Figure 8). It is clear that this was not the case for the human observers, as both local and nonlocal information in the surround affected their judgments.

### General discussion

When the ring image noise was placed only in the surrounding area (Experiment 2), clear evidence of effects upon contrast detection was found at the surround/background border, such that an increment in the surround (and 0.68 deg away from the signal) led the observers to judge that the signal was a decrement. Also in Experiment 2, strong contrast effects were found at the signal/surround border, such that an increment in the surround at this border led to judgments of a decrement signal. Conversely, the intermediary band of the surround between the two borders (signal/surround and surround/background) had little effect. As the two effects at the borders appear to be noncontiguous, with an intermediary band with no effect, we classify the two effects as local and nonlocal, even though the actual retinal distance of the nonlocal effect was modest (0.68 deg), relative to several studies of brightness and brightness induction (e.g., McCourt, 1982; Moulden & Kingdom, 1991; Kingdom & Moulden, 1992; Zaidi, Yoshimi, Flanagan, & Canova, 1992; McCourt, 1994; Blakemore & McCourt, 1997, 1999, 2001). Presumably, the classification images were based upon two separate border effects, one at the signal/surround border and one at the surround/background border, such that the observed luminance profile led to a percept of an increment in the surround, which in turn led to a percept of a decrement signal in the center.

Little evidence of effects of the surround was found for Experiment 1, in which the ring image noise covered both the signal and the surround. This was possibly due to the classification images assessing both observers' weighting of relevant information in the signal area, as well as the irrelevant information of the surround. In this task, the observers should have favored the relevant information and ignored the surround. (Note: See the text above, here and here, and below concerning the use of information by an ideal observer.) This strategy might have masked any effect of the surround in Experiment 1. In Experiment 2, with the image noise only in the surround, it follows that we could only assess information use in the surround, as well. While the observers again should have used the information only in the signal area (and they did to some extent, given their greater than chance performance), the information was constant throughout the experiment. With the variability in the stimulus concentrated in the surround, we effectively reduced the degrees of freedom of the estimated classification images, thereby increasing the statistical power of Experiment 2 to find effects in the surround. Also, in Experiment 1, the higher signal/surround contrasts (with the presence of the pedestal in a discrimination task), relatively few number of trials, and/or the variations in image noise energy might have masked the effects of the surround. In Experiment 2, with the lower signal/surround contrasts (with the detection task), the higher number of trials, and the normalized image noise energy (as well as placing the noise only in the surround), the effects of the surround were revealed.

In Experiment 2, the nonlocal effect had a smaller spatial extent, approximately 0.07 deg, compared to the local effect, which was about half the surround annulus width (0.25 deg). Also, the amplitudes of the nonlocal and local effects were approximately the same, except for AH (whose amplitudes of the nonlocal effects were about half that of the amplitude of the local effects). Assuming that judgments were based upon a linear operation or response from the classification images (i.e., the common assumption of a correlation of a template, such as the classification image, with the stimulus; see Beard & Ahumada, 1998; Abbey, Eckstein, & Bochud, 1999; Abbey & Eckstein, 2002; Ahumada, 2002; Murray et al., 2002; Solomon, 2002), the integrals of the classification images (mass) for the local and nonlocal effects are indicative of their overall strengths. Thus, these results are consistent with smaller effects from the nonlocal border. In addition, generally the results imply that smaller nonlocal effects relative to local effects would be due to a smaller effective area (as opposed to smaller amplitudes for the same effective area).

It seems likely that any effects upon contrast judgments found in this study were completed by 400 ms, well beyond most previous estimates for the time course of contrast judgments (perhaps tens of milliseconds; Enroth-Cugell & Robson, 1966; Shapley & Enroth-Cugell, 1984; Fiorentini, 1990) and brightness induction effects (as mentioned earlier, 50-200 ms; see DeValois et al., 1986; Paradiso & Na-
Coupled with the results of no difference in the classification images from 100 to 400 ms in Experiment 2, the results imply that both the local and nonlocal effects were probably completed by 100 ms in this task. Some may interpret the lack of change in the classification images with increasing duration as lack of evidence for a filling-in mechanism of brightness contrast perception. This filling-in presumably begins at a border, and propagates to the rest of the region in question (DeValois et al., 1986; Grossberg & Todorovic, 1988; Paradiso & Nakayama, 1991; Pessoa et al., 1995; Rossi & Paradiso, 1996; Paradiso & Hahn, 1996; Rudd & Arrington, 2001).

The classification image technique, however, assesses the information being used for a particular task, in this case, a contrast detection/discrimination. It does not assess the spatiotemporal profile of the percept of the central area, which still might be governed by a filling-in from the center/surround border. Also, either shorter stimulus durations or larger surrounds might have revealed the temporal characteristics (such as filling-in) of the effects with this task. For example, if these effects are indeed distance-dependent, as implied by the filling-in metaphor, then a larger surround should lead to slower developing effects.

Thus, we feel that the results of this study are agnostic to the debate of a filling-in brightness mechanism.

For these tasks, an ideal observer predicts that information at each radius (and not each individual pixel) should be weighted equally in the signal area, and ignored in the surrounding area. Thus, a specific classification image is predicted for each experiment. Perhaps not surprisingly, human observers deviated substantially from the ideal observer predictions in these tasks. In Experiment 1, in the presentation format on the left sides of Figures 6 and 8 with the x-axes representing radial distance, an ideal observer would yield a step function with a constant negative value in the signal area, and zero values in the surrounding area. For the human observers, the influence of information in the signal region upon observers’ judgments decreased as the radial distance from the signal/surround border increased. This would be consistent with the observers suboptimally weighting the information by the number of pixels (the area) comprising each radial distance, rather than the single value for each radius. In Experiment 2, an ideal observer would yield no classification image (zero values throughout). For the human observers, non-zero classification images were found, indicating the use of local and nonlocal borders in their judgments. This suggests that human observers were compelled to use this information by some visual constraint, even though that information was irrelevant.

Classification images, in the context of this task, might seem to bear some similarity to a point-by-point brightness matching technique employed by others (Heinemann, 1972; McCourt, 1994; Blakeslee & McCourt, 1997, 1999), as both methods assess multiple points at a fine spatial resolution. However, unlike point-by-point brightness matching, classification images do not measure brightness perceptions of individual points (or radii), but instead gives estimates of how the observer uses information at each radius to influence their judgment of contrast for the entire signal area.

Beginning with the signal/surround border and going outward, the local effects seen in Experiment 2 tend to rise, and then fall off to zero at about 0.25 deg. As mentioned earlier, the initial zero or negative values at the signal/surround border likely were due to uncertainty regarding the exact location of the border (with O/C being particularly affected), in which observers apparently could not confidently attribute the first radius of the surround as being part of the surround. The decrease of the local effect with distance agrees qualitatively with results by Zaidi et al. (1992). In this study, the authors measured local induction effects of circularly symmetric temporally modulating sinusoidal patterns upon a central circular signal. The spatial frequencies of the modulations were varied, and were presented either singly, or in combination with other patterns. For achromatic modulations, the induction effects were fit well with a model assuming the integration of a spatial weighting function that decreased with distance from the signal/surround border, modeled as a negative exponential (also see Spehar, Debonet, & Zaidi, 1996).

A final question is the relationship of this study to brightness perception. It seems clear from contrast brightness induction (e.g., Wallach, 1948; Heinemann, 1955, 1972; Shapley, 1986; Whittle & Challands, 1969) and the CCO illusion ( Craik, 1966; Cornsweet, 1970; O’Brien, 1959) that contrast judgments are not independent from brightness judgments (see also McCourt & Kingdom, 1996). Also, Arend and Spehar (1993a, 1993b) found that, in simple degraded viewing conditions (similar to the center/surround displays in this experiment), contrast, brightness, and lightness judgments do tend to become the same. Finally, some authors have stated that observers more generally might have difficulty disassociating contrast, lightness, and brightness judgments (Jacobsen & Gilchrist, 1988; Sewall & Wooten, 1991; Whittle, 1991). While it is speculative to suggest (see text in the Introduction), it is possible that the results found here reflect not only a judgment of contrast, but also in part a judgment of brightness.

### Appendix A: Hotelling $T^2$ statistic

Tests for nonzero classification images were performed with the single-sample Hotelling $T^2$ statistic, the multivariate equivalent of the univariate t statistic (Harris, 1985; for previous examples of using this statistic for classification images, see Abbey & Eckstein, 2002; Eckstein et al., 2002, 2004). Equation 1 gives the computation for the single-sample $T^2$, with $N$ = the number of observations (trials), $x$ = a vector containing the observed classification image, $x_0$ = a vector containing the null hypothesis classification...
image (all zeroes in this case), and $K^{-1}$ = the inverse of the sample covariance matrix for the individual noise images used to construct the observed classification image.

$$T^2 = N[x - x_0]^	op K^{-1} [x - x_0] \quad (1)$$

The single-sample Hotelling $T^2$ statistic may be transformed into an $F$ statistic with Equation 2, with $p$ = the length of vectors $x$ and $x_0$ (the number of pixels in the classification images). The resulting $F$ statistic has $p$ degrees of freedom in the numerator, and $N - p$ degrees of freedom in the denominator.

$$F = \frac{N - p}{p(N - 1)} T^2 \quad (2)$$

The independent two-sample Hotelling $T^2$ statistic was used to assess differences between two observed classification images. Equation 3 gives the computation for the independent two-sample Hotelling $T^2$, with $N_1$ = the number of observations for the first observed classification image, $N_2$ = the number of observations for the second observed classification image, $x_1$ = a vector containing the first observed classification image, $x_2$ = a vector containing the second observed classification image, and $K^{-1}$ = the inverse of the pooled covariance matrix combining the sum of square deviations and sum of squared products from both samples.

$$T^2 = \frac{N_1N_2}{(N_1 + N_2)} [x_2 - x_1]^	op K^{-1} [x_2 - x_1] \quad (3)$$

The independent two-sample Hotelling $T^2$ may be transformed into an $F$ statistic with Equation 4, with $p$ = the length of vectors $x_1$ and $x_2$ (the number of pixels in the classification images). The resulting $F$ statistic has $p$ degrees of freedom in the numerator, and $N_1+N_2-p-1$ degrees of freedom in the denominator.

$$F = \frac{N_1+N_2-p-1}{p(N_1+N_2-2)} T^2 \quad (4)$$

**Acknowledgments**

This research was supported by the National Aeronautics and Space Administration (NAG-1157), National Institutes of Health (NIH-53455), and the National Science Foundation (NSF-0135118). The authors thank Audrey Hill-Lindsey, Katherine Chong, and Oriana Chavez for their participation as observers.

Commercial relationships: none.
Corresponding author: Steven S. Shimozaki.
Email: shimozak@psych.ucsb.edu.
Address: Psychology, University of California, Santa Barbara, CA, 93106-9660.

**Footnotes**

1 The assumption of deriving a single perceptual filter from a classification images is most accurate in cases in which the observer is approximated well by a linear operator (the correlation of a single template with the exact stimulus location). Previous studies have shown that simple discrimination/detection tasks in which the signal may be easily localized lead to good approximations of human observers as linear operators (Beard & Ahumada, 1998; Abbey et al., 1999; Abbey & Eckstein, 2002; Ahumada, 2002; Murray et al., 2002; Solomon, 2002).

2 Ahumada (2002) studied the effect of nonlinear detection mechanisms (that deviate from the assumption of a single perceptual template or a linear combination of templates, e.g., have location uncertainty) on classification images. For Gabor patches, Ahumada showed that nonlinear observers may result in classification images from the signal present trials biased in favor of the signal, so that the classification images do not represent the underlying perceptual filter in the task. To avoid this potential issue, we chose to calculate the classification images only from the false alarm trials in Experiment 1.

3 As the image noise did not appear over the signal, the classification images should not be biased (by the presence of the signal) in the hit and miss trials, as in Experiment 1. Also, this was tested quantitatively by comparing the classification images calculated from the signal trials (hits and misses) to those from the noise trials (false alarms and correct rejections). As only differences in shape (and not amplitude) were relevant to this issue, the tests were performed on the scaled classification images (the amplitudes were matched to yield a minimum Hotelling $T^2$ value; see Eckstein et al., 2002). No significant differences were found between the signal (hits and misses) and noise (false alarms and correct rejections) classification images across all observers and durations ($F$ values from Hotelling $T^2$ test: 0.817 - 1.280, $p$ values from Hotelling $T^2$ test: .180 - .700). Therefore, in Experiment 2, we combined across all trial types for these classification images, using a simple combination rule appropriate for unbiased responses [(hit - miss) + (FA - CR)] (Beard & Ahumada, 1998; Ahumada, 2002; Murray et al., 2002).

4 As the use of information in the signal area cannot be assessed in Experiment 2, the lower-amplitude classification images of OC may have two causes. The first is that OC was more efficient (ideal) than the other two observers in ignoring the irrelevant information in the surround. The other is that OC has an overall reduction of the use of information across both the signal and surround area, due to larger internal noise. Ahumada (2002) found that, in general, the amplitude of the classification image is inversely related to the internal noise of the observer. OC's lower behavioral performance (Figure 7), increased location uncertainty (negative amplitude for the first radii of the classification images, see text for discussion), and more numer-


