

# Adaptation to textured chromatic fields

Qasim Zaidi

*College of Optometry, State University of New York, 100 East 24th Street, New York, New York 10010*

Branka Spehar

*School of Psychology, University of New South Wales, Sydney, NSW 2052, Australia*

Jeremy DeBonet

*Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

Received April 7, 1997; revised manuscript received July 10, 1997; accepted July 22, 1997

Probe-flash threshold curves were used to show that adaptation to textured fields consists not only of adaptation to the steady local constituents but also of a process that is similar to habituation to prolonged temporal modulation, which in this case could be caused by miniature eye movements across element boundaries. The response curves derived from probe-flash thresholds are compressive on both sides of the adaptation level after adaptation to spatially uniform fields but have an accelerating form when they are measured after adaptation to textured backgrounds. This change is suggestive of a response equalization process, which modifies the response function of each mechanism to match the cumulative frequency distribution of its inputs. © 1998 Optical Society of America [S0740-3232(98)01001-1]

*OCIS codes:* 010.1080, 330.2210, 330.7320, 330.1720.

## 1. INTRODUCTION

In this study we examined the nature of adaptation to chromatic textured backgrounds at different stages of the visual system. Models proposed for color adaptation, constancy, or both have generally considered only adaptation to spatially local steady levels or to the space average as a result of eye movements. The results in this paper reveal the functioning of an additional adaptation mechanism whose effects are similar to the effects of exposure to prolonged temporal modulation. For the purposes of this study we used a modified version of the probe-flash paradigm that we had used earlier to distinguish the effects of habituating to prolonged temporal modulation from the effects of adapting to steady uniform fields.<sup>1,2</sup>

Craik<sup>3</sup> introduced notions of efficiency into adaptation research; since then it has been thought to be optimal that when adapted to a certain level, an observer's discrimination should be best at that level. However, this property would be functionally optimal only if it could be assumed that the frequency distribution of stimulation in the near future would have a maximum at or near the adapting level. The situation is quite different when an observer is actively viewing a spatially variegated field. It has been proposed that, as a result of eye movements, the observer should adapt to the space average.<sup>4,5</sup> However, adapting to the average stimulation alone would not be optimal, because it is unlikely that the highest frequency of future stimuli will occur near the average level. In fact, adaptation to any single level is likely to be grossly suboptimal, and it would be more efficient to adapt the range of sensitivity to the range of expected stimulation.

Studies of adaptation to steady spatially uniform fields have shown that maximal sensitivity is limited to a small range of inputs in any state of adaptation and that adaptation shifts the range of maximal sensitivity to coincide with the average stimulus.<sup>1,3,6,7</sup> In one fairly successful class of models of human brightness and color adaptation the limited range of sensitivity has been modeled as being due to a static nonlinear response (input-output) function, and adaptation has been modeled in terms of an early divisive gain control determined by the time-integrated level of stimulation plus a later subtraction of a portion of the time-integrated signal.<sup>8-10</sup> The response function can be estimated from a set of discrimination thresholds measured over an extended range of inputs by using stimuli that are too brief to disturb the state of adaptation during the measurement. One can derive the properties of adaptation mechanisms by repeating the measurements at different states of steady adaptation and by assuming that the response nonlinearity is invariant.<sup>11</sup> This combination of modeling and experiment has also been used to study chromatic adaptation mechanisms in the *S*-cone color system.<sup>6</sup>

Shapiro and Zaidi<sup>1</sup> adapted this experimental paradigm to measure the distribution of thresholds after the observer had habituated to prolonged temporal modulation of lights along a color axis. They found that the change in sensitivity could not be explained by traditional multiplicative or subtractive adaptation combined with an invariant response function but required a change in the shape of the response function. Zaidi and Shapiro<sup>2</sup> proposed a model of "response equalization" to account for changes in the response function. This model postulates that the most efficient use of a limited response range is

to match the shape of the response function to the expected distribution of inputs, so that on average each level of response occurs with equal frequency.<sup>12</sup> If it is assumed that expectations for inputs are set by the recent adaptation history, then the response should be equated at each input level to the cumulative probability distribution of levels in the adapting stimulus.

We tested whether adaptation to a variegated field can be explained in terms of the processes involved in adaptation to spatially uniform steady fields or whether it also includes a component of habituation to prolonged temporal modulation. We measured observers' probe-threshold versus flash-amplitude functions when the observers were adapted to binary random textures and compared them with functions measured when the observers were adapted to spatially uniform fields metameric to each of the constituents of the texture or to their space-averaged level. We isolated the L - M chromatic mechanism<sup>13,14</sup> by restricting stimuli to an equiluminant color axis passing through the achromatic point  $W$  and parallel to the  $L/(L + M)$ , i.e.,  $RG$ , axis of the MacLeod-Boynton chromaticity diagram.<sup>15</sup>

## 2. EQUIPMENT

All stimulus presentations and data acquisition were computer controlled. Stimuli were displayed on the  $14.14^\circ \times 10.67^\circ$  screen of a Barco 7651 color monitor with a refresh rate of 100 frames/s. Images were generated by a Cambridge Research Systems Video Stimulus Generator (CRS VSG2/3) running in a 90-MHz Pentium-based system. Through the use of 12-bit digital-to-analog converters, after gamma correction the VSG2/3 generator can generate 2861 linear levels for each gun. Any 256 combinations of levels of the 3 guns can be displayed during a single frame. By cycling through precomputed look-up tables we were able to update the entire display for each frame. Phosphor chromaticity specifications supplied by Barco and gamma-corrected linearities of the guns were verified with a Spectra Research Spectra-Scan PR-650

photospectroradiometer. Calibration and specification of colors were performed according to the methods detailed by Zaidi and Halevy.<sup>16</sup> Since all lights used in this study were on the  $RG$  axis through  $W$ , only the  $L/(L + M)$  coordinates are relevant. The  $L/(L + M)$  coordinates of the displayed achromatic point and the end points were  $W = 0.652$ ,  $R = 0.706$ , and  $G = 0.598$ . For all lights used in this study,  $S/(L + M)$  was 0.017. The screen was kept at a constant luminance of 30 cd/m<sup>2</sup>. The mean chromaticity of the screen was set to  $W$ .

## 3. EXPERIMENT 1: PROBE-FLASH EXPERIMENTS WITH TEXTURED AND UNIFORM ADAPTING FIELDS

In Experiment 1 we used four different adapting backgrounds: an  $RG$  random binary texture consisting of equal numbers of randomly intermixed squares set to  $R$  and  $G$ , the extreme chromaticities attainable on the  $RG$  axis, and three spatially uniform fields set to  $R$ ,  $G$ , or the space average of the texture  $W$ . Binary textures consisted of 8.52 squares/deg<sup>2</sup>, with a different random arrangement presented on each trial. For each background type, all the data were collected in a single session. The experimental paradigm is shown in Fig. 1. At the initiation of the session the observer adapted to the background for 120 s and readapted after each trial for 2 s. In each trial the adaptation interval was followed by a test interval of 0.05 s. During the test the screen was divided, either vertically or horizontally, into spatially uniform halves that differed in color. The colors were the end points of a segment on the  $RG$  axis, centered at a fixed level. We will use the term "flash" for the color  $F$  of the mean test level and the term "probe" for the difference in chromaticities between the two halves of the test, i.e., the length  $P$  of the chromatic segment. The  $L/(L + M)$  coordinates of the adaptation levels  $G$ ,  $W$ , and  $R$  were 0.598, 0.652, and 0.706 respectively, and those of the flashes were 0.618, 0.629, 0.640, 0.652, 0.663, 0.674, and 0.685. On each trial the observer indicated whether the division was horizontal or vertical. This two-alternative

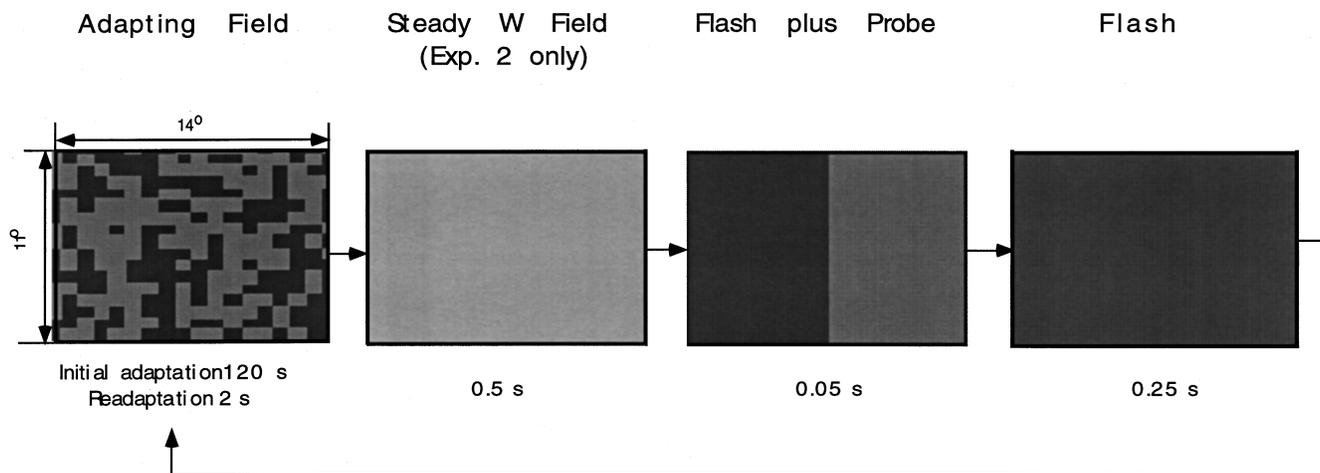


Fig. 1. Spatial configuration and temporal sequence of stimuli for Experiments 1 and 2. The initial adaptation period was 120 s. Each trial was followed by a 2-s period of readaptation. In Experiment 2 only, each trial was preceded by a 0.5-s pause in which the screen was set to  $W$ . In each trial, for 0.05 s the screen was divided into two vertical or horizontal halves,  $F + 0.5P$  and  $F - 0.5P$ . For an additional 0.25 s the screen was set to  $F$ .

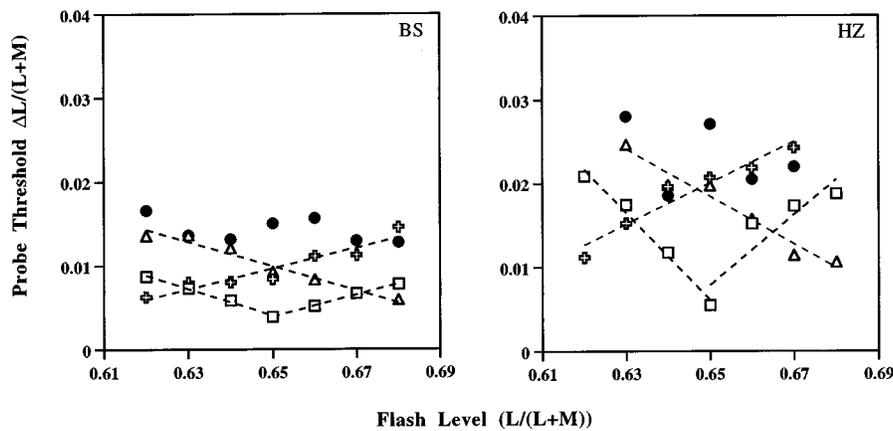


Fig. 2. Results of Experiment 1 for observers BS and HZ. Probe thresholds in  $\Delta L/(L + M)$  units for each of the four adapting backgrounds are plotted versus the  $L/(L + M)$  coordinates of the flash levels. The open squares represent thresholds measured under adaptation to  $W$ . The filled circles represent thresholds measured under adaptation to the textured background. The sets of open triangles and crosses represent thresholds measured after adaptation to the spatially uniform  $R$  and  $G$  fields, respectively. The dashed lines are the best regression fits to the data for the corresponding adaptation conditions.

forced-choice procedure was incorporated into a double-random staircase procedure for each flash level. Trials from all 14 staircases for the 7 flash levels were randomly interleaved. Threshold was estimated as the difference in chromaticity between the two halves at which, with a probability of 71%, the observer could correctly identify the orientation of the division. The probe-flash interval was kept short so as not to disturb the state of adaptation. The flash alone was presented for an additional 0.25 s to rule out the possibility of afterimages providing a clue to the spatial division inside the test. The observers included BS, one of the authors, and a second color-normal female observer, HZ.

The results are shown in Fig. 2. For each of the four adapting backgrounds, the probe thresholds in  $\Delta L/(L + M)$  units are plotted versus the  $L/(L + M)$  coordinates of the flashes. The open squares represent thresholds measured under adaptation to  $W$ . These thresholds are lowest when the flash is identical to the adapting background and rise linearly as a function of flash amplitude, toward both  $R$  and  $G$ . The best-fitting regression lines on each side of  $W$  are shown passing through the points. For observer BS the intercepts of both lines at  $W$  were equal to 0.004; the slopes were  $-0.160$  and  $0.131$  ( $r^2 > 0.99$ ). For observer HZ the intercepts of the lines at  $W$  were equal to 0.006 and 0.008; the slopes were  $-0.519$  and  $0.419$  ( $r^2 = 0.98, 0.82$ ). These V-shaped curves replicate earlier results.<sup>1,2,7</sup> The filled circles represent thresholds measured under adaptation to the textured background. At every flash level the filled circles are considerably higher than the open squares, indicating that after adaptation to the textured background it is considerably more difficult to detect changes at every input level of the  $L - M$  mechanism than during adaptation to the space average.

The sets of open triangles and crosses represent thresholds measured after adaptation to the spatially uniform  $R$  and  $G$  fields, respectively. The dashed lines are the best-fitting regression lines to each of the sets. For observer BS, after adaptation to  $R$  and  $G$ , the slopes were  $-0.160$  and  $0.123$  and the intercepts at the adaptation levels

were 0.004 and 0.005, respectively ( $r^2 = 0.96$  and  $0.93$ ). For observer HZ, after adaptation to  $R$  and  $G$ , the slopes were  $-0.283$  and  $0.246$  and the intercepts at the adaptation levels were 0.007 and 0.01, respectively ( $r^2 = 0.99$  and  $0.90$ ). These thresholds, coupled with the appropriate combination rule, should be sufficient for predicting the data represented by the filled circles if adaptation to the textured field is entirely equivalent to spatially local adaptation to the steady level of the constituent squares. Every decision rule that combines the responses of independently adapted spatially distributed units will predict that thresholds after adaptation to the textured field be no higher than the maxima of the triangles and crosses. At most flash levels the filled circles are close to the maximum of the crosses and triangles, but at the flash level equal to  $W$ , where the thresholds after  $R$  and  $G$  adaptation are approximately equal, thresholds after adaptation to the  $RG$  texture were significantly higher. Clearly, spatially local adaptation to the steady level of the constituent patches is not a sufficient explanation for adaptation to textured fields. For observer HZ, after adaptation to  $R$ ,  $G$ , or  $RG$  texture, some of the thresholds at the two extreme flash levels (0.618 and 0.685) were out of the range measurable by the equipment but only if their magnitudes were higher than the thresholds at 0.629 and 0.674.

The results in Fig. 2 show that adaptation to a textured field consists of more than adaptation to the space average, or independently to the constituents, or to any combination thereof. The observers in these experiments were instructed to fixate the center of the screen, but miniature eye movements are unavoidable even when observers are trying to maintain fixation.<sup>17</sup> If the main effect of eye movements were integration over space within receptive fields,<sup>4,5</sup> the adaptation level would be set by the mean chromaticity and luminance. Because the space-averaged color of the textured background was identical to  $W$ , the elevated thresholds rule out spatial integration as a major factor. We propose that the unexplained component of threshold elevations is due to eye movements causing transient stimulation of receptive fields at the

borders of the squares, thus creating temporal modulation of stimulation to individual neurons, which has been shown to cause chromatically selective elevation of thresholds.<sup>13</sup> In addition, an increase in threshold at the mean level concomitant with a decrease in slopes of threshold curves is reminiscent of the effects of adapting to prolonged temporal modulation.<sup>1,2</sup>

#### 4. EXPERIMENT 2: PROBE-FLASH EXPERIMENTS WITH TEMPORAL DELAYS

In making measurements in Experiment 1, observers noted that the transient afterimage of the textured field also appeared textured, thus indicating that adaptation to the textured field involves multiple adaptation mechanisms, including spatially local adaptation to steady stimulus levels. To distinguish habituation to modulations from adaptation to steady levels, we repeated Experiment 1 but with a 0.5-s presentation of a uniform screen at  $W$  following the adaptation and readaptation intervals (Fig. 1). By using thresholds at the mean level of the adapting stimulus as a measure, Krauskopf *et al.*<sup>13</sup> showed that habituation to prolonged temporal modulation is a relatively long-lasting effect, whereas Shapiro *et al.*<sup>18</sup> showed that within 0.5 s there is a measurable change in probe-flash curves when adaptation is shifted between two spatially uniform chromatic fields.

The results are shown in Fig. 3. For each of the four adapting backgrounds the probe thresholds in  $\Delta L/(L+M)$  units are plotted versus the  $L/(L+M)$  coordinates of the flashes, on the same scales as in Fig. 2. The open squares replot thresholds measured under adaptation to  $W$  in Experiment 1, since these would remain unchanged with the additional short pause at the same adaptation level. The filled circles represent thresholds measured under adaptation to the textured background. Thresholds on the  $W$  (0.652) flash are almost identical in Figs. 2 and 3 for each of the observers; therefore, if these values are taken as a measure of the magnitude of habituation, it can be inferred that habituation to the variegated field does not decay to any appreciable extent after a 0.5-s ex-

posure to the space-averaged spatially uniform field. The sets of open triangles and crosses represent thresholds measured after adaptation to the spatially uniform  $R$  and  $G$  fields, respectively, and, when compared with the sets of open triangles and crosses in Fig. 2, reveal that a 0.5-s exposure to the  $W$  field caused a significant change in adaptation state as reflected by the shapes of the probe-flash curves. The sets in Fig. 3 show systematic departures from linearity. For observer BS the slopes of the best-fitting lines were  $-0.043$  and  $0.054$  ( $r^2 = 0.55$  and  $0.44$ ). For HZ the slopes were  $-0.101$  and  $0.002$  ( $r^2 = 0.62$  and  $0.00$ ). These considerations indicate that the processes involved in habituation to variegated fields have different temporal properties from the processes involved in chromatic adaptation to spatially uniform fields. The shapes of the threshold curves formed by the filled circles also change from Fig. 2 to Fig. 3, but, as discussed in Section 5, this change seems to reflect entirely the restriction that the filled circles remain above or close to the maxima of the triangles and crosses. In Section 5 we will assume that the filled circles that form an inverted V shape are close to representing the state of the system while it is habituated to the textured background, after more-transient effects have mostly subsided.

The solid lines in Fig. 3 are the best regression fits to the filled circles, which represent habituation to the textured background. For observer BS the intercepts at  $W$  were  $0.017$  and  $0.016$  and the slopes were  $0.162$  and  $-0.107$ , respectively ( $r^2 = 0.95$  and  $0.99$ ), for flash levels on the  $G$  and  $R$  sides of the space-averaged mean. For observer HZ the intercepts at  $W$  were  $0.027$  and  $0.028$  and the slopes were  $0.285$  and  $-0.415$ , respectively ( $r^2 = 0.91$  and  $0.89$ ). Thresholds are highest at the space-averaged level  $W$  and decrease linearly toward the coordinates of the constituent colors. Remarkably, at some neural stage the visual system is simultaneously most sensitive to both colors that are present in the texture. Thresholds on the flashes metameric to  $W$  are higher by factors of 4 for observer BS and 5 for observer HZ than the comparable thresholds for adaptation to  $W$ . The marked difference between the inverted V-shaped curve

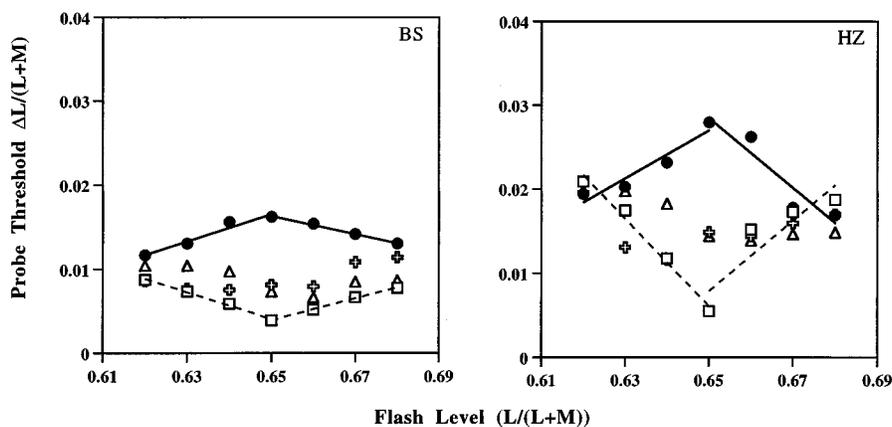


Fig. 3. Results of Experiment 2 for observers BS and HZ. Probe thresholds in  $\Delta L/(L+M)$  units for each of the four adapting backgrounds are plotted versus the  $L/(L+M)$  coordinates of the flashed judgment levels. The open squares represent thresholds measured under adaptation to  $W$ . The filled circles represent thresholds measured under adaptation to the textured background. The dashed and solid lines are the best regression fits to the data for the corresponding adaptation conditions. The sets of open triangles and crosses represent thresholds measured after adaptation to the spatially uniform  $R$  and  $G$  fields, respectively.

and the upright V-shaped curve again makes it obvious that habituation to the textured scene is qualitatively different from adaptation to the space-averaged color.

## 5. MECHANISMS OF ADAPTATION TO CHROMATIC TEXTURES

A model for the effects of adapting to the textured fields not only should take into account adaptation and habituation mechanisms at different stages of the visual system but should also incorporate spatial inhomogeneities in adaptation state across the field and a decision process that involves pooling spatially distributed responses. In this section we first present a plausible model for the steady-state response of the L – M system when the system is adapted to spatially uniform steady  $W$ ,  $R$ , or  $G$  fields. We then consider the possibility that, after adaptation to chromatic textures, L – M neural units are in various states of adaptation across the test field, and we derive predictions for the texture habituation results based on various response combination rules. To accomplish this purpose with a minimal set of mechanistic assumptions we derive the overall response curve of the system in different steady states and infer properties of adaptation and habituation processes from changes in the response curves. We assume that the squares, triangles, and crosses in Fig. 2 reflect the steady state of the system adapted to  $W$ ,  $R$ , and  $G$  fields respectively, and that the inverted V shape formed by the filled circles in Fig. 3 reflects the state of the system habituated to the  $RG$  texture.

In every steady state, if two assumptions are satisfied, the overall input–output (response) function of a visual system can be derived from the kind of probe-threshold versus flash-amplitude curves measured in Experiments 1 and 2. The first assumption is that the perturbation by the probe does not disturb the steady state of the system and that the readaptation interval restores the system to the steady state after the disturbance caused by the flash. This assumption is justified if the probe interval is short compared with the time constants of the adaptation mechanisms in the system. The second assumption is that the difference between the responses of the system to two lights should be a constant amount for all pairs of lights that are at the threshold for discrimination. This assumption provides a linking theory between the response function of the system and empirical thresholds.

For the L – M system we chose to describe the response  $Z$  as separate logarithmic functions of positive and negative opponent signals  $Q$ , with parameters  $\alpha$ ,  $\beta^+$ , and  $\beta^-$ :

$$Z = \frac{1}{\beta^+} \log(\alpha + \beta^+ Q) - \frac{1}{\beta^+} \log(\alpha) \quad \text{for } Q \geq 0,$$

$$Z = \frac{1}{\beta^-} \log(\alpha + \beta^- Q) - \frac{1}{\beta^-} \log(\alpha) \quad \text{for } Q < 0. \quad (1)$$

This response curve is sigmoidal, smooth, roughly linear through  $Q = 0$  and can allow for different amounts of compression or acceleration in the two limbs. The vir-

tues of this form of response function are explicated by Zaidi and Shapiro<sup>2</sup> and Wiegand *et al.*<sup>19</sup>

We used the second assumption to estimate the parameters for each set of probe–flash data. For every flash level, the difference between the responses to  $F + 0.5P$  and  $F - 0.5P$  can be approximated by the probe amplitude  $P$  multiplied by the derivative of the response function,  $1/(\alpha + \beta Q)$ , for  $Q$  equal to the level of the flash  $F$ , i.e.,

$$|Z(F + 0.5P) - Z(F - 0.5P)| \approx \left| \frac{P}{\alpha + \beta F} \right|. \quad (2)$$

For each flash level, threshold is defined as the probe amplitude for which the right-hand side is equal to some constant  $\kappa$ . When terms are transposed, relation (2) predicts that the probe amplitude at threshold will be a linear function of the flash level:

$$P = \kappa(\alpha + \beta^+ F) \quad \text{for } F \geq 0,$$

$$P = \kappa(\alpha + \beta^- F) \quad \text{for } F < 0. \quad (3)$$

Therefore the intercepts and the slopes of the best-fitting regression lines (as in Figs. 2 and 3) provide estimates of the  $\alpha$  and the  $\beta$  and hence of the shapes of the response functions.

To a large extent, the shifts in the threshold curves after adaptation to steady uniform fields (Fig. 2) are explainable simply by insertion of a high-pass temporal filter (HPTF) subsequent to opponent combination of L and M signals but before a static response nonlinearity, as in Fig. 4. Filtering out steady signals at this stage will cause the opponent signal from every steady uniform adapting field to be zero. Therefore an invariant response curve predicts that thresholds will be a function solely of the magnitude and the sign of the  $\Delta L/(L + M)$  difference between the flash and the adapting field, regardless of the adapting background. When thresholds for  $W$ ,  $R$ , and  $G$  adaptation were replotted against flash amplitudes in  $\Delta L/(L + M)$  units, this prediction was perfectly consistent with observer BS's data and was close for observer HZ. For observer BS the best-fitting regression line for all negative flash amplitudes ( $r^2 = 0.974$ ) had an intercept equal to 0.004 and a slope of  $-0.145$ ; for all positive flash amplitudes the intercept was 0.004 and the slope was 0.134 ( $r^2 = 0.943$ ). For observer HZ the best-fitting regression line for all negative flash amplitudes ( $r^2 = 0.830$ ) had an intercept equal to 0.008 and a slope of  $-0.282$ ; for all positive flash amplitudes the intercept was 0.010 and the slope was 0.269 ( $r^2 = 0.852$ ). Consequently, the extrapolated thresholds on the  $R$  and  $G$  adapting fields were equal to the threshold at  $W$ , and the results showed that on prolonged exposure to a steady

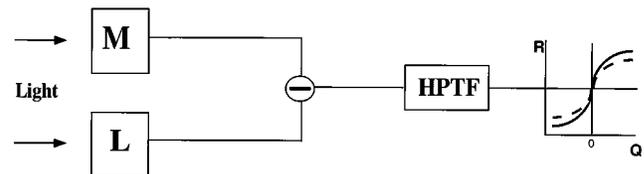


Fig. 4. Schematic of the L – M color mechanism, showing a postopponent HPTF and an adaptable nonlinear response function.

uniform field the L - M system adapts so that sensitivity is maximal at the adapting level and that the differential response at each perturbation level remains unchanged. Similar results were reported by Shapiro and Zaidi<sup>1</sup> and Krauskopf and Gegenfurtner.<sup>7</sup> Parenthetically, since the opponent combination is assumed to be linear the postopponent HPTF could be replaced by an identical HPTF in each preopponent branch. Figure 4 also allows for a change in the shape of the response function, which we will use below to explain the additional effects of adapting to a variegated scene during active viewing.

The estimated response functions for adaptation to *W* and habituation to the *RG* texture for observers BS and HZ are shown in Fig. 5 (left). The dashed and solid response curves were derived from the dashed and solid threshold curves in Fig. 3. The estimated response curve for adaptation to *W* is compressive on both sides of  $Q = 0$ , whereas the response curve for habituation to the *RG* texture has small amounts of acceleration on both sides of  $Q = 0$ . Because the relevant threshold curves were steeper for observer HZ, both compression and acceleration were more exaggerated in her estimated response curves. As was noted earlier,<sup>1,2,20</sup> this type of change in the shapes of the response curves cannot be ex-

plained by any first or second stage multiplicative or subtractive adaptation mechanism.

The main task of this section is to model the effects of adapting to chromatic textured fields, as represented by the filled circles in Fig. 2. For simplicity we assume that, after adapting to the *RG* texture, every L - M unit across the test field is in one of three states: (i) adapted to *R* with the response function ( $Z_R$ ) derived from the regression line through the triangles in Fig. 2, which depict the  $P_R$  thresholds, (ii) adapted to *G* with the response function ( $Z_G$ ) derived from the regression line through the crosses in Fig. 2, which depict the  $P_G$  thresholds, or (iii) habituated to temporal modulation at the texture element boundaries with the response function ( $Z_H$ ) derived from the inverted *V* in Fig. 3, which depicts the  $P_H$  thresholds. Since this temporal modulation is caused by miniature eye movements during fixation, it may affect only a subset of units over the test field. The proportions of these classes of units are assumed to be  $n_R$ ,  $n_G$ , and  $n_H$ . In terms of Fig. 4,  $Z_R$  and  $Z_G$  result from the action of the HPTF without altering  $Z$ , whereas  $Z_H$  requires a change in shape of the response function.

When the same linking assumption as above is used, after adaptation to texture *T*, at threshold the difference

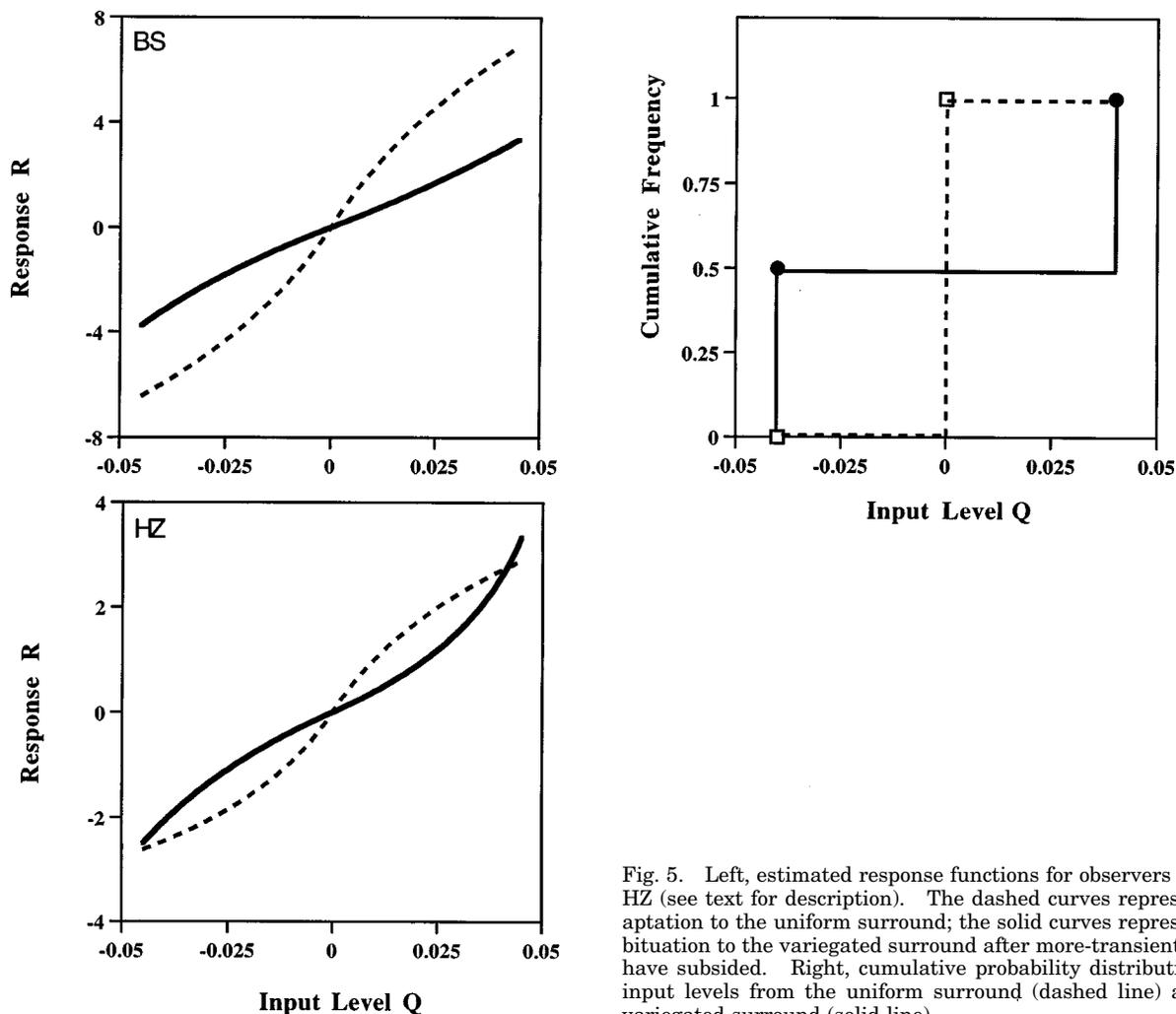


Fig. 5. Left, estimated response functions for observers BS and HZ (see text for description). The dashed curves represent adaptation to the uniform surround; the solid curves represent habituation to the variegated surround after more-transient effects have subsided. Right, cumulative probability distributions for input levels from the uniform surround (dashed line) and the variegated surround (solid line).

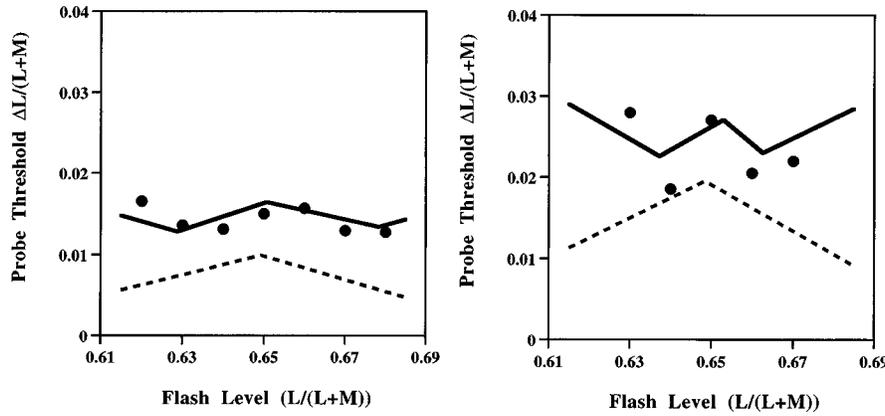


Fig. 6. Probe thresholds measured under adaptation to the textured background (Experiment 1) plotted in  $\Delta L/(L+M)$  versus the  $L/(L+M)$  coordinates of the flashed judgment levels. The dashed lines represent the prediction from Eq. (5). The solid lines represent the prediction from Eq. (6).

between the combined response of the  $L-M$  system to the two halves of the test field,  $|Z_T(F+0.5P) - Z_T(F-0.5P)|$ , should be equal to the same criterion level  $\kappa$ , where  $Z_T$  is the combined response to the texture  $T$ . The responses of the three classes of  $L-M$  units can be combined in a general fashion by use of a Minkowski norm:

$$Z_T = [(n_R Z_R)^\phi + (n_G Z_G)^\phi + (n_H Z_H)^\phi]^{1/\phi}, \quad (4)$$

where  $\phi$  can have any value between 1 and  $\infty$ . When  $\phi = 1$ , the right-hand side of Eq. (4) represents a weighted sum of the responses of all  $L-M$  units across the test field. As  $\phi$  approaches infinity, in the limit  $Z_T$  is equal to the maximum of the constituent responses.<sup>21</sup> As is discussed below, the response function shown in Eqs. (1) leads to fairly straightforward expressions for many different values of  $\phi$ .

However, perceptual considerations of the decision process can be more instructive for our purposes. As stated above, tests that appear immediately after adaptation or readaptation to the textured field had a textured appearance that was due to the after-image of the texture. Therefore two points in the test field could appear different, even if they were on the same half of the probe field, depending on the adaptation state of the underlying units. The decision about the orientation of the spatial division in the test therefore has to involve comparisons of spatially extended combinations. Two combinations are especially worth considering, as their predictions bracket predictions from all other combination rules. Thresholds will be lowest after adaptation to the texture if observers can reliably discriminate between the two halves whenever the responses of any one of the three classes of  $L-M$  units differ between the two halves by the criterion level. This decision rule predicts that, for each flash level  $F_i$ , thresholds  $P_T$  after adaptation to the texture  $T$  should be the least of the thresholds on the  $R$  or  $G$  adaptation curves in Fig. 2 or the inverted V curves in Fig. 3:

$$P_{T_i} = \min(P_{R_i}, P_{G_i}, P_{H_i}). \quad (5)$$

In Fig. 6 the dashed curves show this prediction, and the filled circles replotted from Fig. 2 are clearly well above the predicted levels. On the other hand, thresholds will

be highest if, because of the variegated appearances of the probe halves, observers can reliably distinguish between them only when the responses of all three classes of  $L-M$  units differ between the halves by the criterion value. This rule predicts that, after adapting to the texture, thresholds should be at the maximum of the three curves listed above:

$$P_{T_i} = \max(P_{R_i}, P_{G_i}, P_{H_i}). \quad (6)$$

This prediction is represented by the solid curves in Fig. 6. The data fall close to this prediction and approximate a W shape similar to that of the prediction.

Every other combination rule from Eq. (4) predicts thresholds that fall between the dashed and the solid curves in Fig. 5. As an example, we considered the case when  $\phi = 1$ , which represents a weighted summation of responses from the three classes of unit. For each flash level  $F_i$ , by substituting from relation (2) and Eq. (3) into Eq. (4) and from straightforward algebraic manipulations, we derived the following equation:

$$\frac{P_{R_i} P_{G_i} P_{H_i}}{P_{T_i}} = n_R P_{G_i} P_{H_i} + n_G P_{R_i} P_{H_i} + n_H P_{R_i} P_{G_i}. \quad (7)$$

Using the measured magnitudes of the  $P$ 's, we simultaneously estimated the best  $n$ 's and tested the adequacy of the combination rule by multiple linear regression. The  $r^2$  for the best fit of the two sides of Eq. (7) was only 0.24, thus indicating that no weighted sum of responses from the three classes of units provides an adequate model for the texture adaptation data.

We can now return to the question raised in Section 1 as to the optimal adaptation level for an observer during viewing of a variegated field. Zaidi and Shapiro<sup>2</sup> showed that the most efficient use of the limited dynamic range of a mechanism occurs when the sensitivity of the mechanism at each input level is adjusted to be proportional to the expected relative frequency of that input level. If a mechanism can be assumed to expect the same frequency of input levels as in the adapting stimulus, then the threshold curves after adaptation to  $W$  and habituation to the  $RG$  texture are both qualitatively consistent with re-

sponse equalization: In both cases, maximum sensitivity coincides with the levels of the input present in the adapting stimulus.

According to relation (2), thresholds are inversely proportional to the derivative of the response curve; therefore the degree to which the estimated response curves resemble the cumulative probability distribution of the inputs is a measure of the extent of response equalization in the system.<sup>2</sup> The cumulative probability curves for the uniform  $W$  field and the  $RG$  texture are shown in Fig. 5 (right). The dashed and solid cumulative distributions should be compared with the dashed and solid response curves, respectively. The response curves resemble smoothed versions of the cumulative distribution functions and are, therefore, only qualitatively consistent with a response-equalization process. It is not entirely surprising that the response curves are not as sharp as the cumulative distributions of physical inputs. Given that any input to the visual system passes through cascades of band-limited spatiotemporal filters, the inputs to the locus of habituation to the textured field can have a smoother cumulative distribution than the physical stimulus. Also, if the reason that the effects of habituation to variegated fields are similar to the effects of habituation to modulated fields is that miniature eye movements across texture elements create temporal modulation of stimulation to individual neurons, then spatial integration within each receptive field that lies across a texture boundary smears the effective input distribution. However, these considerations cannot apply to spatially uniform fields; therefore the cumulative distribution that is the input to the habituation stage may be smoothed out by neural noise at earlier stages, or there may be inherent limits to the malleability of the response function. On the basis of output noise considerations, it was proposed recently that the response function be equal to the integral of the cube root of the probability-density function of the input and hence smoother than the cumulative response function.<sup>22</sup>

## 6. DISCUSSION

This paper was motivated by results from a companion paper<sup>23</sup> showing that slow, large-field changes in chromaticity could be masked by spatial patterns that are as much as an order of magnitude higher in spatial frequency. The magnitude of adaptation as a function of background spatial frequency was bandpass, peaking at a frequency roughly 40–120 times the fundamental frequency of the test, which made it obvious that this effect was different from the extensively studied spatial-frequency-specific adaptation.<sup>24</sup> Our best guess was that the results were due to habituation to the temporal modulations caused by involuntary eye movements. In principle, the critical test would have been a repeat of the experiment with perfect image stabilization, but neither did we have the equipment to stabilize miniature eye movements nor was it clear whether the stabilization achievable by commercial equipment would be good enough for this purpose. As an alternative, we examined whether after adaptation to chromatic texture, the change in

threshold curves is similar to the change caused by habituation to chromatic temporal modulation.<sup>1,2</sup>

To identify the types of adaptive mechanism that function during the viewing of variegated scenes, we used a probe-flash experimental technique and simple mathematical considerations. The analyses indicated that, during adaptation to uniform fields, probe-flash threshold curves have unique minima at or close to the adaptation level, whereas, after habituation to a variegated field, threshold curves have multiple minima corresponding to the levels present in the field. By quantitative comparisons of threshold curves measured after adaptation to variegated or uniform fields we showed that adaptation to a variegated field can be adequately explained neither in terms of adaptation to the space-averaged level nor in terms of spatially local adaptation to its constituents. The effects of viewing variegated fields are reminiscent of the desensitizing effects of prolonged temporal modulations. As a result of eye movements across spatial variations, the stimuli imaged onto the receptive fields of individual neurons are modulated in time, resulting in habituation.

In this study the size of the texture squares was chosen to be equal to the  $RG$  texture that maximally masked large field chromaticity changes along the same color axis in the companion study.<sup>23</sup> The result of that study, that textures with larger or smaller elements mask to a lesser extent, provides corroboration for the class of model in Eq. (4) consisting of three classes of adapted units. If the squares are too large, there will be fewer neurons whose receptive fields oscillate across boundaries, and if the squares are too small, there may be too much integration within receptive fields for there to be substantial modulation of responses. In both cases the magnitude of habituation will be less, leading to lower thresholds. Therefore, receptive field sizes and amplitudes of eye movements will jointly determine the sizes of the texture elements that elevate thresholds the most.

The results of this study are related to the phenomenon termed “contrast adaptation”.<sup>25</sup> In particular, Greenlee and Heitger showed that contrast-discrimination curves are flatter after adaptation to drifting sinusoidal gratings than when they are adapted to a spatially uniform field; i.e., the detectability of low-contrast gratings is reduced but the discriminability between high-contrast gratings is improved. Changes in shapes of discrimination curves are consistent with electrophysiological results that show that, in cortical neurons, prolonged exposure to high-luminance-contrast drifting sinusoidal gratings of optimal spatial frequency and orientation changes the parameters of the contrast-response function so that the slope is shallower but the operating range is extended.<sup>26,27</sup> Even though eye movements do not seem to have been explicitly implicated in contrast adaptation to stationary stimuli, they must be necessary because responses of contrast-sensitive neurons decrease to resting levels in the absence of temporal variations of stimulation. The term “contrast adaptation” may be suitable for adaptation to stimuli that contain luminance variations, because postretinal visual neurons seem to carry spatial-contrast information in the luminance domain.<sup>24</sup> However, whether this term is appropriate for adaptation to chromatically

variegated fields<sup>28,29</sup> remains to be justified. Most psychophysical<sup>23</sup> and physiological<sup>14</sup> evidence suggests that visual neurons are low pass in spatial frequency for chromatic variations, so chromatic contrast *per se* may not be an important variable. In fact, Zaidi<sup>30</sup> showed that habituating to different temporal distributions, such as square and triangular waves, led to changes in probe-flash threshold curves that were qualitatively consistent with response equalization to the distribution of inputs.

The response-equalization process proposed here for habituation to a chromatically variegated field matches the system's response function to the cumulative distribution of inputs and is akin to the widely used probability integral transform,<sup>31</sup> yielding a uniform frequency distribution for response levels. Consequently, this process is optimal for maximizing the information capacity of a mechanism with a limited response range.<sup>12,1,32</sup> A nonlinear response function is in some ways a limitation on the information capacity of a system. On the other hand, the Taylor series expansion of this nonlinearity contains higher-order terms. Therefore a nonlinear response function that can be tuned to the cumulative frequency of inputs will contain information about higher-order moments of the input distribution. This information could potentially be used by neural mechanisms in the signal processing of higher-order statistics for separating out mixtures of independent sources or reversing the effect of an unknown filter.<sup>33</sup> Because response equalization requires information about the complete frequency distribution of inputs, it may seem to be computationally unwieldy and physiologically unlikely. However, based on maximizing the information that the response contains about its input, Bell and Sejnowski<sup>34</sup> devised unsupervised gradient ascent learning rules for tuning ogive-shaped response functions to single-peaked input distributions. One can easily adapt their methods to the case of input distributions with peaks on both sides of the zero-opponent signal by applying the gradient rule to each of the response functions in Eq. (1) separately. The training set required for the unsupervised learning rule is provided by the frequency of levels present in the adapting stimulus.

In summary, this study shows that, during adaptation to a variegated chromatic field, thresholds at every input level are close to the maximum of those to be expected from adaptation to the steady local components of the stimulus and from habituation caused by eye movements across chromatic variations. These results have implications for detecting illuminant-caused appearance changes in complex scenes. Changes in illumination cause correlated large-field chromaticity shifts across variegated scenes.<sup>23,35,36</sup> The decision rule for detecting large-field chromaticity changes depicted in Fig. 6 explains why observers are relatively insensitive to the transient changes in object chromaticities caused by changes in illumination over variegated scenes.<sup>23,37</sup>

## ACKNOWLEDGMENTS

We thank Hong Zhan for participating as an observer and John Krauskopf, Andrea Li, Dean Yager, and Mike Brill for discussions. A portion of this study was presented at

the 1996 meeting of the Association for Research in Vision and Ophthalmology and the 1995 European Conference for Visual Perception. This research was supported by National Eye Institute grant EY07556 to Q. Zaidi.

## REFERENCES

1. A. G. Shapiro and Q. Zaidi, "The effect of prolonged temporal modulation on the differential response of color mechanisms," *Vision Res.* **32**, 2065–2075 (1992).
2. Q. Zaidi and A. G. Shapiro, "Adaptive orthogonalization of opponent-color signals," *Biol. Cybern.* **69**, 415–428 (1993).
3. K. J. W. Craik, "The effect of adaptation on differential brightness discrimination," *J. Physiol. (London)* **92**, 406–421 (1938).
4. M. D'Zmura and P. Lennie, "Mechanisms of color constancy," *J. Opt. Soc. Am. A* **3**, 1662–1672 (1986).
5. M. D. Fairchild and P. Lennie, "Chromatic adaptation to natural and incandescent illuminants," *Vision Res.* **32**, 2077–2085 (1992).
6. Q. Zaidi, A. G. Shapiro, and D. C. Hood, "The effect of adaptation on the differential sensitivity of the S-cone color system," *Vision Res.* **32**, 1297–1318 (1992).
7. J. Krauskopf and K. Gegenfurtner, "Adaptation and color discrimination," *Vision Res.* **32**, 2165–2175 (1992).
8. E. N. Pugh and J. D. Mollon, "A theory of the  $\pi_1$  and  $\pi_3$  color mechanisms of Stiles," *Vision Res.* **19**, 293–312 (1979).
9. W. S. Geisler, "Initial image and after-image discrimination in the human rod and cone system," *J. Physiol. (London)* **294**, 165–179 (1979).
10. E. H. Adelson, "Saturation and adaptation of the rod system," *Vision Res.* **22**, 1299–1312 (1982).
11. M. M. Hayhoe, N. I. Benimoff, and D. C. Hood, "The time course of multiplicative and subtractive adaptation processes," *Vision Res.* **27**, 1981–1996 (1987).
12. S. Laughlin, "A simple coding procedure enhances a neuron's information capacity," *Z. Naturforsch.* **36**, 910–912 (1981).
13. J. Krauskopf, D. R. Williams, and D. Heeley, "Cardinal directions of color space," *Vision Res.* **22**, 1123–1131 (1982).
14. A. M. Derrington, J. Krauskopf, and P. Lennie, "Chromatic mechanisms in lateral geniculate nucleus of macaque," *J. Physiol. (London)* **357**, 241–265 (1984).
15. D. I. A. MacLeod and R. M. Boynton, "Chromaticity diagram showing cone excitation by stimuli of equal luminance," *J. Opt. Soc. Am. A* **69**, 1183–1186 (1979).
16. Q. Zaidi and D. Halevy, "Visual mechanisms that signal the direction of color changes," *Vision Res.* **33**, 1037–1051 (1993).
17. R. H. S. Carpenter, *Movements of the Eyes* (Pion, London, 1988).
18. A. Shapiro, Q. Zaidi, and D. Hood, "Adaptation in the red-green (L – M) color system," *Invest. Ophthalmol. Visual Sci.* **S31**, 262 (1990).
19. T. E. Weigand, N. Graham, and D. C. Hood, "Testing a computational model of light-adaptation dynamics," *Vision Res.* **35**, 3037–3051 (1995).
20. H. R. Wilson and R. Humanski, "Spatial frequency adaptation and contrast gain control," *Vision Res.* **33**, 1133–1149 (1993).
21. N. B. Haaser and J. A. Sullivan, *Real Analysis* (Van Nostrand Reinhold, New York, 1971).
22. D. I. A. MacLeod and T. von der Twer, "Optimal nonlinear codes," *Invest. Ophthalmol. Visual Sci.* **38**, S254 (1997).
23. Q. Zaidi, B. Spehar, and J. S. DeBonet, "Color constancy in variegated scenes: the role of low-level mechanisms in discounting illumination changes," *J. Opt. Soc. Am. A* **14**, 2608–2621.
24. N. Graham, *Visual Pattern Analyzers* (Oxford U. Press, New York, 1989).
25. M. W. Greenlee and F. Heitger, "The functional role of contrast adaptation," *Vision Res.* **28**, 791–797 (1988).

26. J. A. Movshon and P. Lennie, "Pattern selective adaptation in visual cortical neurones," *Nature (London)* **278**, 850–852 (1979).
27. G. Sclar, P. Lennie, and D. D. DePriest, "Contrast adaptation in striate cortex of macaque," *Vision Res.* **29**, 747–755 (1989).
28. M. A. Webster and J. D. Mollon, "The influence of contrast adaptation on color appearance," *Vision Res.* **34**, 1993–2020 (1994).
29. M. A. Webster and J. D. Mollon, "Colour constancy influenced by contrast adaptation," *Nature (London)* **373**, 694–698 (1995).
30. Q. Zaidi, "Adaptation processes governed by the distribution of inputs," *European Conference on Visual Perception, Perception* **22**, Suppl. 2, 62 (1993).
31. B. Harris, *Theory of Probability* (Addison-Wesley, Reading, Mass., 1966).
32. N. Brady and D. J. Field, "Early nonlinearities in visual coding and natural image statistics," *Invest. Ophthalmol. Visual Sci.* **38**, S633 (1997).
33. S. Haykin, *Blind Deconvolution* (Prentice-Hall, Englewood Cliffs, N.J., 1994).
34. A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Comput.* **7**, 1129–1159 (1995).
35. J. L. Dannemiller, "Rank ordering of photoreceptors catches from objects are nearly illumination invariant," *Vision Res.* **33**, 131–137 (1993).
36. D. H. Foster and S. M. C. Nascimento, "Relational colour constancy from invariant cone-excitation ratios," *Proc. R. Soc. London, Ser. B* **250**, 116–121 (1994).
37. K. J. Linnell and D. H. Foster, "Dependence of relational colour constancy on the extraction of a transient signal," *Perception* **25**, 221–228 (1996).