Pergamon

0042-6989(95)00209-X

Vision Res., Vol. 36, No. 10, pp. 1371–1390, 1996

Printed in Great Britain
0042-6989/96 $15.00 + .00

Effect of Spatial Scale and Background Luminance on the Intensive and Spatial Nonlinearities in Texture Segregation

NORMA GRAHAM,*† ANNE SUTTER‡

Received 10 August 1994; in revised form 24 July 1995

Perceived segregation between element-arrangement textures is affected both by spatial scale and background luminance. The effects on the spatial nonlinearity are consistent with the proposed structure for complex (second-order) channels. The effects on the intensive nonlinearity are *not* consistent with an early, local nonlinearity but are consistent with either (i) a relatively early, local, nonlinearity occurring before the spatial frequency channels but after a sensitivity-setting stage, or (ii) inhibitory interaction among channels modeled as a normalization network. Thus the texture intensive nonlinearity comes *after* sensitivity to spatial frequency and background luminance has been determined. For six of seven observers, the texture intensive nonlinearity was compressive by 10% contrast for both increments and decrements (at high background luminance, large spatial scale). Copyright © 1996 Elsevier Science Ltd.

Adaptation Gain Inhibition Nonlinear Texture

INTRODUCTION

Models incorporating spatial frequency- and orientation-selective channels explain many aspects of perceived segregation among regions distinguished by texture, particularly when known nonlinearities are included. [For an excellent review of the earlier texture literature, see Bergen (1991).]

One such nonlinearity is the *spatial nonlinearity* involved in *complex channels*. A complex channel consists of two stages of linear filtering separated by a rectification-type nonlinearity. A number of investigators have invoked such processes in the study of texture and motion perception, often calling them *non-Fourier* or *second-order* or simply *nonlinear* processes (e.g. Robson, 1980; Grossberg & Mingolla, 1985; Shapley & Gordon, 1985; Chubb & Sperling, 1988; Sutter *et al.*, 1989; Fogel & Sagi, 1989; Sperling, 1989; Turano & Pantle, 1989; Victor & Conte, 1991; Wilson *et al.*, 1992).

A second nonlinearity—the *intensive nonlinearity*—is necessary also. It depends more directly on the intensities (or perhaps contrasts) than on the spatial characteristics of the pattern (e.g. Sperling, 1989; Graham, 1991; Victor & Conte, 1991; Graham *et al.*, 1992a). This intensive nonlinearity might result from an *early, local nonlinearity* preceding the channels (perhaps retinal light adaptation). As diagrammed in Fig. 1, a compressive nonlinear function might be applied to each point of the stimulus, and the output of this nonlinear function might then be the input to the frequency- and orientation-selective channels.

Alternately, as in Fig. 2, the intensive nonlinearity in texture segregation might result from interaction among the channels [e.g. intracortical inhibition (Morrone *et al.*, 1982; DeValois & Tootell, 1983; Bonds, 1989)]. We have modeled such interaction by a normalization network based on the work of Heeger (1991, 1994) and Robson (1988a,b). Other investigators have invoked

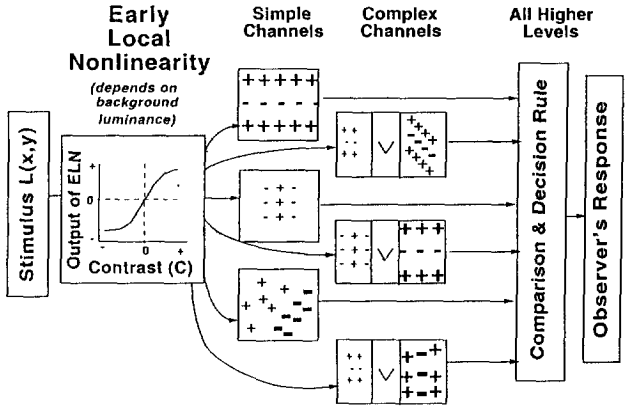

FIGURE 1. Diagram of the early–local nonlinearity hypothesis. See Fig. 10 for a modified version of the early–local nonlinearity hypothesis.

*To whom all correspondence should be addressed [*Email* nvg@psych.columbia.edu].

†Department of Psychology, Columbia University, New York, NY 10027, U.S.A.

‡Department of Psychology, Loyola University, 6525 North Sheridan Rd, Chicago, IL 60626, U.S.A.

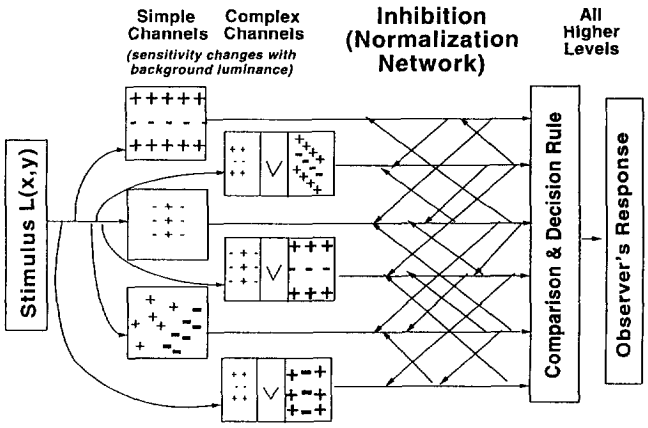

FIGURE 2. Diagram of the normalization hypothesis. The intensive nonlinearity is the result of inhibitory interaction among the channels themselves instantiated as a normalization network.

similar processes in somewhat different forms to explain texture segregation or other similar tasks (e.g. Grossberg & Mingolla, 1985; Sperling, 1989; Lubin & Nachmias, 1990; Malik & Perona, 1990; Wilson, 1990; Bergen & Landy, 1991; Landy & Bergen, 1991; Victor & Conte, 1991; Lubin, 1992; Gorea & Papathomas, 1993).

The primary aim of the study reported here is to explore the intensive nonlinearity at different spatial scales and background luminances in an attempt to more fully characterize its quantitative properties and to begin to understand its source and function. One specific motivation for this study is the following observation: if the intensive nonlinearity were in fact acting point-by-point on the stimulus before any spatial-filtering action occurs (as shown in Fig. 1), it should not depend on the spatial scale of the pattern. In addition, we were interested in the effect of spatial scale and background luminance on complex channels. Finally, other studies (e.g. Cannon & Fullenkamp, 1993; Graham *et al.*, 1993) have found dramatic individual differences in similar situations; another aim of the present study was to collect enough results from individual subjects that individual observers' nonlinearities could be described and compared.

METHODS AND PROCEDURES

Two experiments are reported here. Briefly, both experiments used stimuli like those in Fig. 3. The two types of elements in any one stimulus were always squares of the same size, but stimuli could differ in their overall spatial scale. (When spatial scale varied, both the square widths and the inter-square spaces varied proportionately—while keeping the number of elements the same—as if the same stimulus were viewed from different distances.) At each spatial scale, the contrasts of the two elements were varied to allow the fitting of models as in Graham *et al.* (1992a). In Expt 2, background luminance was also varied. Seven observers were studied (one in both experiments, four others in Expt 1, and two others in Expt 2).

The Stimuli

Element-arrangement Textures and Constant-Difference Series. The stimuli used here (e.g. Fig. 3) are *element-arrangement* texture patterns like those used originally by Beck *et al.* (1983). In a constant-difference series of these stimuli (see Fig. 4), the background luminance and the difference between the luminances of the two element types remains fixed, but the absolute luminances of the two element types vary together.

We will use the word "pattern" to mean a particular spatial arrangement at a particular background luminance without commitment to the contrast of its elements—e.g. for one pattern, both types of elements are 0.5 deg squares, the space between squares is 0.5 deg, and the background luminance is 200 td. Figure 5 shows the full set of 66 contrast combinations used for any particular pattern. The horizontal axis gives the contrast of one element type:

$$c_1 = \Delta L_1 / L_{\text{bkd}},$$

where L_1 is the luminance of the elements of type one, L_{bkd} is the luminance of the background and $\Delta L_1 = L_1 - L_{\text{bkd}}$. The vertical axis gives the contrast of the other element type. Contrasts are shown in arbitrary units called steps, and the size of a step was in general different for different patterns. Any set of stimuli along a positive diagonal in Fig. 5 form a constant-difference series. All stimuli on any line through the origin have the same ratio ($\Delta L_2 / \Delta L_1$) and thus the same contrast ratio (c_2 / c_1). This line through the origin can be represented by its angle, called the *contrast-ratio angle*, which we measure relative to the negative diagonal. Thus, it goes from − 90 deg for same-sign-of-contrast (both elements dark) to + 90 deg for same-sign-of-contrast patterns (both light).

Spatial Characteristics of the Stimuli. The two element types were always squares of the same size. The space between two adjacent squares was the same as the square width. There were always 12 rows and 15 columns of elements with the five central columns being a checkerboard region and each set of five flanking columns a striped region as in Fig. 3. (Logically the seven central rows could be considered to be a checkerboard region. That is not how most observers report the perception, however. This distinction is not of relevance here in any case, as we do not study the boundary.)

For Expt 1, the three different spatial scales had the following characteristics. At the largest scale the squares in the pattern were 0.33 deg wide (16 pixels) as were the spaces between the squares. Thus, the repetition period of either the checkerboard or striped regions, which is two rows and columns of square elements with the associated inter-element spaces, was 1.33×1.33 deg (64×64 pixels); therefore, the fundamental frequency, which is equal to the reciprocal of the repetition period, was 0.75 c/deg both horizontally and vertically. The middle spatial scale was one-quarter that of the largest spatial scale The smallest spatial scale was one-quarter that of the middle. In summary, the three spatial scales used in

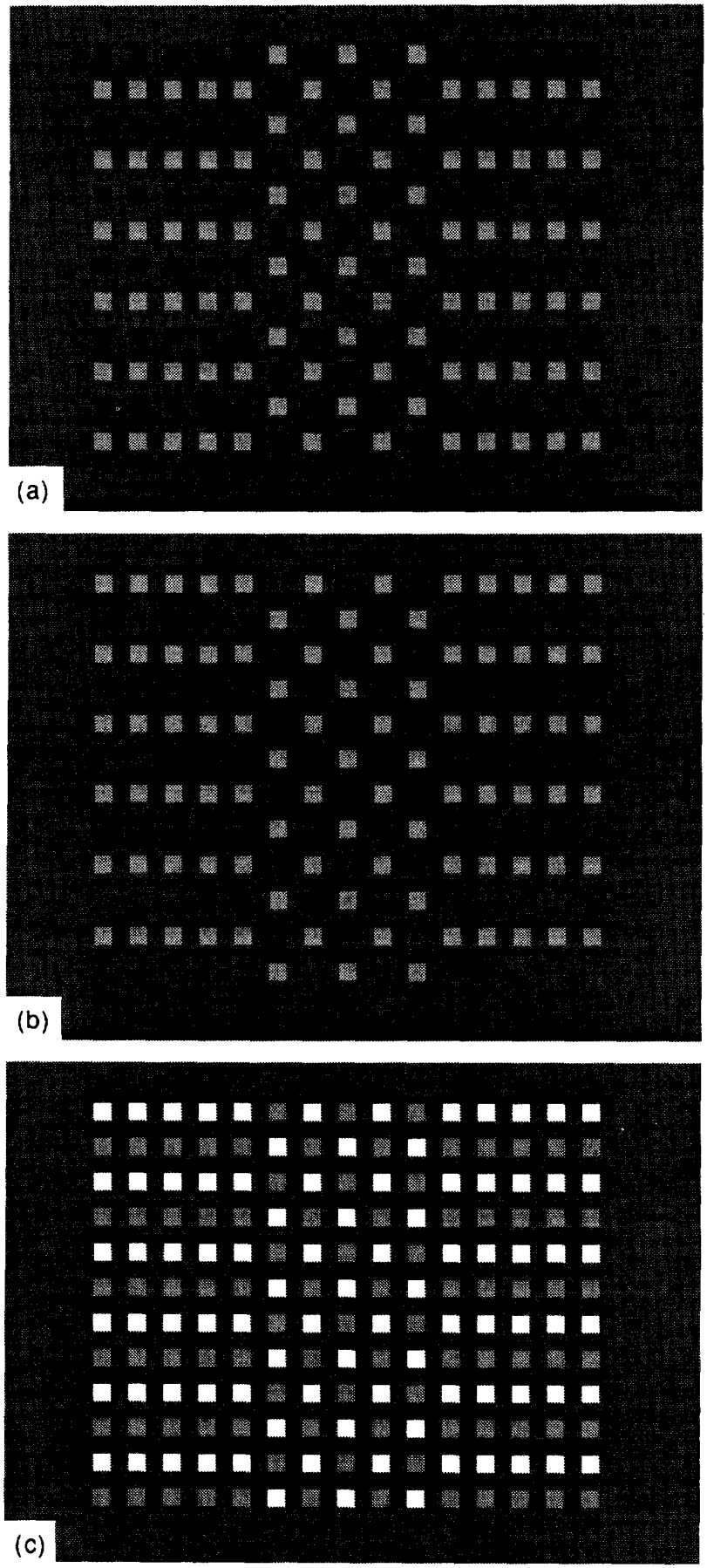

FIGURE 3. Reproductions of several of the stimuli. The two element types are arranged in a checkerboard in the center regions and in stripes in the flanking regions. Illustrated are (a) an *opposite-sign-of-contrast* stimulus, (b) a *one-element-only* stimulus, and (c) a *same-sign-of-contrast* stimulus. The stimuli shown here will have been distorted in reproduction.

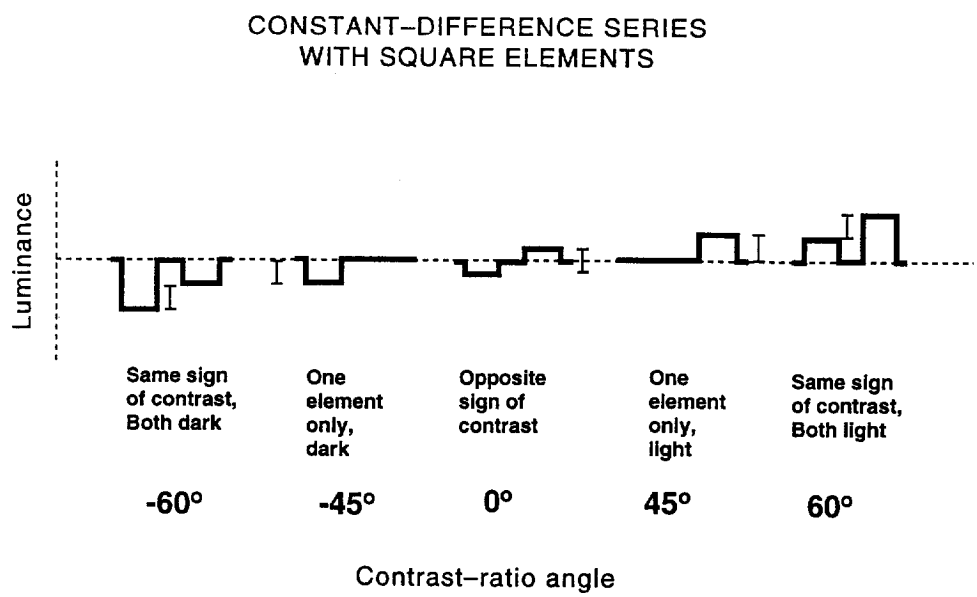

FIGURE 4. Each small diagram shows the luminance profiles of the two element types in a single stimulus. These five stimuli are all in the same *constant-difference series*—i.e., the difference between the luminances (and, consequently, contrasts) of the two element types is the same in all five stimuli. The contrast-ratio angle is defined in the next figure.

Expt 1 had square widths of 0.33, 0.08, and 0.02 deg corresponding to fundamental frequencies of 0.75, 3, and 12 c/deg.

The smallest spatial scale used in the Expt 1 could not be used in Expt 2 as it was invisible at the lower background luminances even at the highest available contrast. The three scales used in Expt. 2 had square widths of 0.33, 0.17, and 0.08 deg, corresponding to fundamental frequencies of 0.75, 1.5, and 3 c/deg.

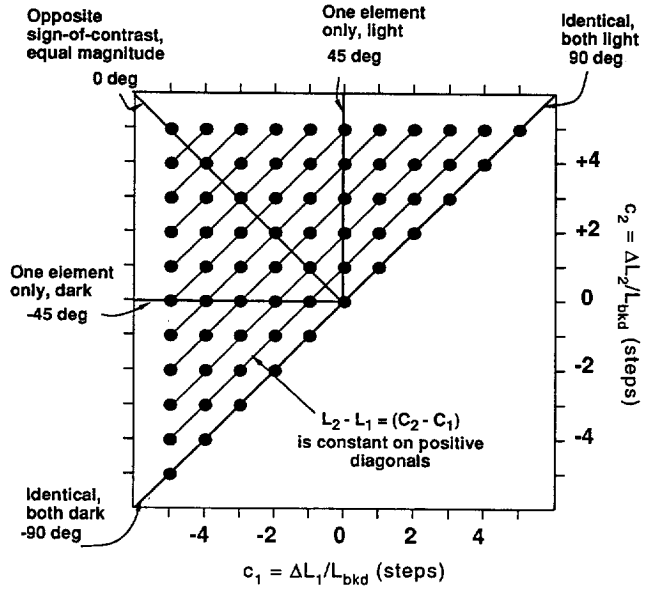

FIGURE 5. Diagram of the half-matrix of 66 different contrast combinations used for each pattern in these experiments. The contrast of one element type is plotted on the horizontal axis and that of the other element type on the vertical axis. Each circle represents one stimulus. The stimuli along any positive diagonal form a constant-difference series. The ratio of the contrasts of the two element types is constant along lines through the origin. The corresponding contrast-ratio angle is labeled outside the diagram.

Viewing Conditions and Background Luminance. The luminous screen was approximately 16 cm high and 21 cm wide which, at a viewing distance of 0.91 m, was 10×13 deg of visual angle. The background luminance of the screen was constant throughout each experiment (during stimuli, fixation points, and inter-stimulus intervals) at 18 ft-L.

In Expt 1, the observer viewed the screen binocularly while sitting in a chair with unrestrained head and natural pupils. [A typical observer's pupil at this light level would be about 2.5 mm dia (e.g. Hood & Finkelstein, 1986, Table 5.2) corresponding to about 300 td. We did not measure our observers' pupils.] The chair was set so that the distance between the eye and the screen was 0.91 m initially (presumably modulated a few centimeters by unintended shifts in head position). There was a small lamp 6 ft behind the observer which—along with the CRT screen itself—provided some ambient illumination in the room.

In Expt 2, the observer sat inside a small booth that shielded the observer's eyes from stray light in the room. The observer's chin rested on a chinrest, and the booth and chinrest were placed so that the distance from the artificial pupil to the center of the CRT screen was 0.91 m. The observer viewed the screen monocularly through a 2 mm dia artificial pupil placed as close as possible to the front of the eye (in the case of MH and KC) or to eyeglasses (in the case of WS). Optical correction for MH was done with a lens placed directly behind the artificial pupil. Observer WS wore eyeglasses; observer KC wore contacts. With a 2 mm dia pupil, 18 ft-L corresponds to 200 phot td. (Since the artificial pupil in this experiment was not at the same effective place in the light beam as the real pupil, especially in the case of WS, the size of the real pupil may have had some minor effect on the resulting illuminance on the retina, an effect not

represented in the troland values given here.) Three lower background luminances were also used in Expt 2. These lower backgrounds were produced by placing neutral density filters between the artificial pupil and the screen. These neutral density filters attenuated the luminance by approximately 0.6, 1.2, and 1.8 log units, thus producing levels of approximately 50, 13, and 3 phot td.

Contrast of the Stimuli. The step-sizes were chosen on preliminary work to crudely equate visibility at the different spatial scales and background luminances.

In Expt 1, the step-size for the largest and middle scales (fundamental frequencies of 0.75 and 3 c/deg) was 4% contrast. Thus, the largest contrast used (5 steps) was 20%. For the smallest scale (fundamental frequency 12 c/deg) the step-size was increased to 8% contrast, and thus the largest contrast used was 40%.

In Expt 2, the step-size was identical for the three scales at each background luminance but varied with background luminance. The step-sizes were 4, 8, 12, and 16% from the highest to the lowest background luminance used.

Equipment and Calibration. The patterns were generated and the experiments run by a Macintosh IIci on a standard Apple monochrome monitor using Pascal programs built upon programs kindly supplied by Hugh Wilson. Based on calibrations with a uniform field, there were available 151 linearized gray levels. The background luminance of all our stimuli was set at the midpoint and hence the smallest contrast step was nominally 1/75. However, the following difficulty should be noted. The luminance of a pixel is not perfectly independent of its neighbors. This is a particular problem with the smallest squares used in Expt 1, since they only contained a single pixel. To obtain a given contrast with these stimuli, we found it was necessary to increase the nominal contrast (the contrast in the array of numbers controlling the pixel-by-pixel luminance of the CRT) by a factor of approximately two. The true contrast for the different sizes of squares was estimated both by some physical calibrations and by comparing psychophysical thresholds at different viewing distances. (For example, the 1-pixel squares at 0.9 m should act like 4-pixel squares at 3.6 m distance when their true contrasts are identical and the observer is able to accommodate to the different viewing distances correctly.) We are not completely satisfied with either type of calibration and feel there is substantial uncertainty in the estimates of contrast at the smallest size. This uncertainty is not as much as a factor of 1.5, however, so one might conservatively say that the step-size is definitely somewhere between 5 and 12% rather than the 8% mentioned above. Uncertainty of this magnitude does not affect the conclusions below.

The Procedures

Structure of the Experiments. Each block of trials in both Expts 1 and 2 contained 198 trials (3 spatial scales of pattern × 66 contrast combinations in each). In any block, the trials were all at the same background luminance. An observer generally participated in one block a day.

In Expt 1, there were 4 blocks of trials run for each of the 5 observers; thus there were 4 repetitions of each stimulus for each observer.

In Expt 2, there were supposed to be 16 blocks for each of the 3 observers (4 blocks at each of the 4 background luminances). Thus, there were again supposed to be 4 repetitions of each stimulus for each observer. But observer KC did not complete the last two blocks (one block at each of the middle two background luminances). The order of the blocks at different background luminances was random with the constraint that one block at each background luminance was done before moving on to the next set of 4 blocks.

Structure of a Trial and the Response Scale. Each trial started when the subject pressed the top inch of a response device (an "Unmouse"). A fixation pattern then appeared for 1 sec. It was a cross located in the middle of the screen, at 10% contrast (8×16 pixels, i.e., 0.17×0.33 deg). Immediately after the fixation pattern, the stimulus pattern was presented for 1 sec with an abrupt onset and offset. After stimulus offset, a 1 sec delay occurred and then a beep signaled that the observer could make a response by pressing the appropriate position within a rectangle (about 10 cm wide × 2.5 cm high) on the response device. Although the responses were actually recorded on a finely divided scale (from 0 to 100, as the observers knew), five equally spaced numerals were written on the face of the response device to guide their responses. A sheet of paper was available whenever they wished to look at it stating that the meaning of these numerals were:

0—No segregation between the regions
1—Barely perceptible segregation between regions
2—Perceptibly segregated regions
3—Moderately segregated regions
4—Highly segregated regions.

After the observer's response, there was a double-beep.

We used the 1 sec delay because some preliminary experiments with computer-recorded responses but without the delay failed to replicate the dip in segregatability for opposite-sign-of-contrast square-element patterns relative to one-element-only stimuli. Several possible reasons for this effect were briefly discussed in Graham *et al.* (1993). Sutter and Graham (1995) present evidence that complex channels have considerably slower processing dynamics than do simple channels; thus, without the imposed delay observers may go ahead and respond on the basis of the simple-channels response before the complex-channels response has been registered.

Observers and instructions. There were five observers in Exp. 1 and three observers in Expt 2; one observer, MH, ran in both experiments. All were undergraduates at Columbia University. The results shown for observers SO, WO, and TH in Expt 1 were the first results collected for these observers. The other observers had participated in several texture-rating experiments before running Expt

2 here. The observers were naive as to the purpose of these experiments, although two of the observers (MH and CV) were co-authors on related papers (Graham *et al.*, 1992b, 1993).

Before participating in their first segregation-rating experiments, the observers all received 15–30 min of instructions including a series of practice patterns. They were told to maintain fixation at the center of the screen (even after the fixation mark had disappeared) and to indicate by their response the degree to which the regions immediately and effortlessly segregated. They were explicitly instructed NOT to focus on the individual squares or any other form of local information and NOT to indicate the result of scrutinizing the patterns for differences. They were asked to ignore factors such as the overall size and the degree of brightness of the pattern. They were also instructed to maintain a focus of attention that was global (while, however, continuing to fixate the center of the screen).

Average Segregation Ratings. For each of the 66 stimuli (as in Fig. 5) for a given pattern (a given spatial scale, background luminance, and observer), the average segregation rating was computed over the four repetitions of the stimulus. The models were fit to these averages.

Segregation Thresholds. A segregation threshold c_0 was also calculated for each pattern. This threshold c_0 is the contrast in the one-element-only version of the pattern that leads to a mid-scale segregation rating (a response scored at 50 out of 100—corresponding to "2-perceptibly segregated regions" on the written scale). In practice, computing the segregation threshold involved interpolation between values of contrast used for one-element-only stimuli in the experiment or, in two cases, extrapolation (for observers TH and WO at the smallest spatial scale in Expt 1, aided by available results for opposite-sign-of-contrast stimuli).

In relating the results here to other work, one might wish to know how the texture-segregation thresholds used here relate to the conventional detection thresholds. Do regions only segregate when one-element-only stimuli are far above detection threshold, or will they segregate whenever the necessary information is visible? We have not done an exhaustive study of detection thresholds, but some measurements (for observers WS and KC at the highest and lowest background luminances and the largest and smallest scales from Expt 2) produced conventional detection thresholds ($d' = 1.5$) that were 2–4 times lower than the segregation thresholds for the same patterns.

THE MODELS

This section briefly reviews previous models (Graham, 1991; Graham *et al.*, 1992a–c) and presents sets of predictions from the two candidates for the intensive nonlinearity. A third model for the intensive nonlinearity will be presented in the Results section.

Simple Models

Models containing only simple channels (channels consisting of a single stage of linear filtering) make a simple prediction for element-arrangement textures where both elements are squares of the same size (as used here)—namely, all the stimuli in a constant-difference series should perceptually segregate to almost exactly the same extent. Here we approximate simple models' predictions by the following equation (as we did in our earlier papers after justification by the predictions of the full model). Let D_S be the contribution of the simple channels to the predicted perceived segregation. Then,

$$D_S = w_S \cdot |L_1 - L_2| = w_S \cdot |\Delta L_1 - \Delta L_2|. \quad (1)$$

The constant w_S reflects the sensitivity of the observer's simple channels and depends on the pattern (i.e., the spatial scale and the background luminance) but not on the contrasts of the elements.

Although these simple models do *not* contain complex channels nor the intensive nonlinearities, they must contain a rule for computing the observer's response from the channels' outputs (the rightmost box in Figs 1 and 2). The rule used in all our models here is effectively this: two regions segregate to the extent that they elicit different amounts of activity in one or more of the channels. D_S in Eqn (1) is a measure of this extent. The observer's rating is assumed to be a monotonic function of D. We routinely use a family of rules to insure that the particular form of the rule makes no difference to the conclusions. See Graham (1991) and Graham *et al.* (1992a) for further discussion.

An example of predictions from Eqn (1) is shown in the bottom left of Fig. 6. The horizontal axis shows the contrast-ratio angle. The vertical axis gives predictions of the model. Each curve connects points representing the predictions for a constant-difference series, and the size of that constant difference increases from the bottom to the top curve.

Complex Channels

Experimental results, however, were very unlike the simple-channel models' predictions (Graham, 1991; Graham *et al.*, 1992a). For one thing, the opposite-sign-of-contrast patterns, like that in Fig. 3(a), were somewhat less segregatable than the one-element-only members [like that in Fig. 3(b)] in the same constant-difference series. (This effect will be shown here in the data for individual observers; it is the dip in the middle of many of the curves in Fig. 8, for example.)

This dip in sensitivity for opposite-sign-of-contrast patterns is easily explained by complex channels due to the rectification-type-nonlinearity that occurs between the two filtering stages (Graham, 1991; Graham *et al.*, 1992a). The following equation gives approximately the contribution of the complex channels to perceived segregation:

$$D_C = w_C \cdot ||\Delta L_1| - |\Delta L_2||. \quad (2)$$

The parameter w_C reflects the number and sensitivity of the observer's complex channels and, like w_S, depends

on the pattern (i.e., the spatial scale and the background luminance) but not on the luminances of the elements.

Equation (2) assumes that the complex channels are completely insensitive to opposite-sign-of-contrast patterns as if the nonlinearity between the two filtering stages were full-wave rectification (or some other even-symmetric function). The characteristics of the intermediate nonlinearity in complex channels are still unknown, so this is a tentative assumption. But weakening it to be, for example, half-wave rectification would have no effect on our conclusions below.

To predict the observer's segregation, the contributions of both simple and complex channels are pooled to yield the following:

$$D_{S+C} = \{D_S^k + D_C^k\}^{1/k}, \quad (3)$$

where the exponent k is a parameter determining the characteristics of the pooling across channels. The observer's rating is assumed to be a monotonic function of this quantity.

Model with Early–Local Nonlinearity

There was a second discrepancy between the predictions of the simple models and the results. Same-sign-of-contrast patterns [e.g. Fig. 3(c)] are in general much less segregatable than the others in the same constant-difference series, and the further that the elements' luminances get from the background luminance, the worse segregation gets. (See, for example, the downturn of the ends of most curves in Fig. 8.) This poor segregation between regions in these same-sign-of-contrast element-arrangement patterns occurs even though the difference between the two types of elements is easily perceivable (Beck *et al.*, 1991).

An early local nonlinearity acting pointwise before the channels (Fig. 1) predicts this result by adjusting the limited operating range of the visual system to be centered at the current background luminance in order to insure discriminability among luminances near the background luminance; this has the consequence that luminances far from the background luminance are compressed. To calculate approximate predictions from such an early local nonlinearity, one can use Eqns (1) and (2) above, but substitute the outputs of the early local pointwise nonlinearity, which will be called $r(\Delta L_i)$, for the luminances ΔL_i in those equations. The resulting quantities will be called D_{S*} and D_{C*} to distinguish them from the corresponding quantities in Eqns (1) and (2):

$$D_{S*} = w_S \cdot |r(\Delta L_1) - r(\Delta L_2)| \quad (4)$$

and

$$D_{C*} = w_C \cdot ||r(\Delta L_1)| - |r(\Delta L_2)||. \quad (5)$$

Next combine the two overall differences, as in Eqn (3), but call the result D_{ELN}:

$$D_{ELN} = \{D_{S*}^k + D_{C*}^k\}^{1/k}. \quad (6)$$

This observer's rating is assumed to be a monotonic function of this predicted value D_{ELN}.

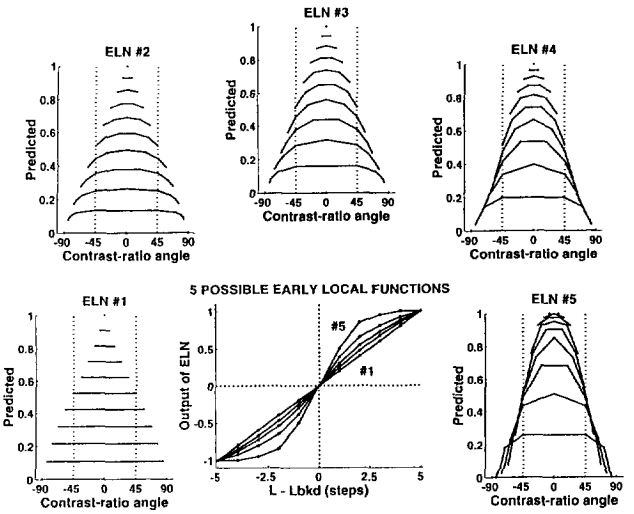

FIGURE 6. Predictions from an early–local nonlinearity model. The central lower panel shows five different early, local transformations that vary from linear (#1) to very compressive (#5). The value $r(\Delta L)$ [see Eqns (4) and (5)] is plotted on the vertical axis and the contrast or ΔL (in steps) is plotted on the horizontal axis. The other five panels show the predictions from these five transformations; the predicted value of D_{ELN} is plotted as a function of contrast-ratio angle for a number of constant-difference series (different curves). For the bottom curve (coincident with the horizontal axis), the constant difference is 0; for the second curve from the bottom, the constant difference equals one step, and so on. The predicted value shown on the vertical axis is D_{ELN} computed from Eqn (6) with the values of ΔL_1 and ΔL_2 expressed in steps. The weights on the channels set at $w_S = 1$ and $w_C = 0$. (The absence of complex channels means there is no dip in the middle of these curves.) The value of the exponent for pooling across channels was set at $k = 2$. The predictions are plotted relative to the maximum for that set of parameters so the maximum = 1.0 in each panel.

Figure 6 shows predictions from Eqn (6) for several versions of a model incorporating simple channels and an early–local nonlinearity. The five different early–local functions are shown in the middle bottom panel and the predictions in the surrounding panels. Function #1 is linear and thus the resulting model is the same as the simple-channels model; function #5 is the most compressive. The exact values of the other parameters of the model are given in the figure legend. Notice that, as the early local function becomes more compressive (moving from the panel on the lower left in an arc around to the panel on the lower right), the predicted curves become more curved downward at the ends; i.e., the same-sign-of-contrast patterns become less and less segregatable relative to the others in their constant-difference series.

Model with Interchannel Interaction in a Normalization Network

There is an alternative candidate for the intensive nonlinearity, namely inhibition among the channels (the physiological substrate for which might be inhibition among V1 cortical cells). Notice that the same-sign-of-contrast stimuli contain the same amount of energy at the fundamental frequency as do other stimuli in the constant-difference series, but they contain a great deal more energy at some other spatial frequencies, in particular at those higher frequencies that define the

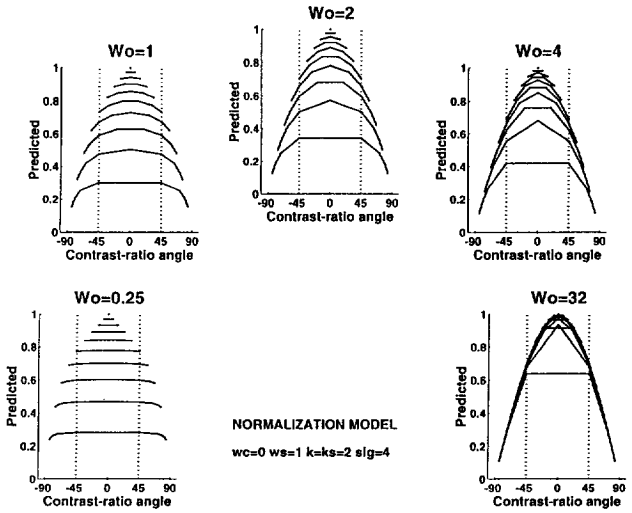

FIGURE 7. Predictions from the normalization model embodying inhibitory interaction among channels. The axes and conventions are like Fig. 6. The predicted value shown on the vertical axis is D_{NORM} computed from Eqn (8) with the values of ΔL_1 and ΔL_2 expressed in number of steps. The weights on the channels were set at $w_S = 1$ and $w_C = 0$. The values of the exponents for pooling across channels and for pooling across space were set at $k = 2$ and $k_S = 2$. To control the amount of intensive nonlinearity in these example predictions, the value of σ was fixed at 4 and the value of w_O was varied. The five panels show the predictions from five values of w_O. The predictions are plotted relative to the maximum for that set of parameters so the maximum = 1.0 in each panel. In general, as the ratio w_O / σ becomes larger, the degree to which the predicted curves bend down at the ends becomes more pronounced, with the curves converging to a common envelope at the highest values of w_O / σ.

individual elements. If there is inhibition among channels, the responses to these "other" frequencies will suppress the responses to the frequencies (in particular, the fundamental frequency) that distinguish the checkerboard from the striped arrangement. (Nothing in the results here can answer the questions of how large a range of orientations or spatial frequencies or spatial positions contributes to the inhibition, so we will continue to talk in general about "other" channels.) The spatially pooled regional responses of channels sensitive to these "other" frequencies is approximately equal to

$$R_O = w_O \cdot \{|\Delta L_1|^{k_S} + |\Delta L_2|^{k_S}\}^{1/k_S}, \quad (7)$$

where k_S is an exponent describing pooling across the spatial position of the output within any single channel, and w_O is a parameter describing the sensitivity of the "other" channels. To finish the prediction, we use the following expression (Graham *et al.*, 1992a):

$$D_{NORM} = \frac{\{D_S^k + D_C^k\}^{1/k}}{\{\sigma + D_S^k + D_C^k + R_O^k\}^{1/k}}, \quad (8)$$

where σ is a constant which, among other things, serves to keep the expression from going to infinity as the contrast of the stimuli is reduced to zero.

Figure 7 shows some predictions from Eqn (8) for five versions of a model incorporating simple channels and a normalization network. As the amount of inhibition is increased (in this case by increasing the parameter w_O for fixed values of the other parameter values—see legend to Fig. 7), the curves bend down more and more until they converge at the highest value (lower right panel of Fig. 7).

Fitting the Models to the Results

The half-matrix of data from a single pattern for a single observer consists of 66 data points; each is the average of the 4 segregation ratings given by the observer to one of the 66 contrast combinations at which the pattern was presented (as in Fig. 5). There were 51 such half-matrices of data: 15 from Expt 1 (5 observers × 3 spatial scales) and 36 from Expt 2 (3 observers × 3 spatial scales × 4 background luminances). Each such half-matrix of data was fit separately to each of the two kinds of intensive nonlinearity [to Eqn (6) and to Eqn (8)]. Details of the methods for fitting predictions from these equations to the experimental results are described more fully in the Appendix. Briefly, the methods are like those of Graham *et al.* (1992a) with several modifications. The modifications were necessary because there was substantially more fitting to be done here [51 half-matrices of results as compared to 3 in Graham *et al.* (1992a)] and here we needed parameter estimations, not just good fits. In overview, a crude grid search was done over most of the parameters of both models, but the Nelder–Meade algorithm was used to search for the best values of the function describing the early local nonlinearity at the ± 5 steps of contrast used in the experiment, and for the best-fitting final monotonic transformation between the predicted values D and the observer's segregation ratings. The goodness-of-fit was assessed by a Pearson product–moment correlation between the experimental results and the transformed predictions. While this is an imperfect measure of goodness of fit (since the variances at different points on the response scale are not the same), it has the advantage of speed and seems quite adequate for our purposes here.

RESULTS

Preview of the Comparisons Between Model and Experiment

As we will show in the first and longest section below, the results of the two experiments reported here are *not* consistent with the original version of the early, local, nonlinearity hypothesis (Fig. 1); they *are* consistent with a modified version of the early, local model which will be presented (Fig. 10). They are also consistent with the normalization model if the inhibitory interaction occurs among the channel outputs (as in Fig. 2) although not if the interaction occurs before the filtering action of the channels. In short, at least within the framework of these kinds of models, sensitivity to different spatial frequencies and different backgrounds is set *prior* to the intensive nonlinearity that acts in texture segregation.

Or to describe these results from another perspective, the effects of spatial scale and background luminance on the intensive nonlinearity in texture segregation can be

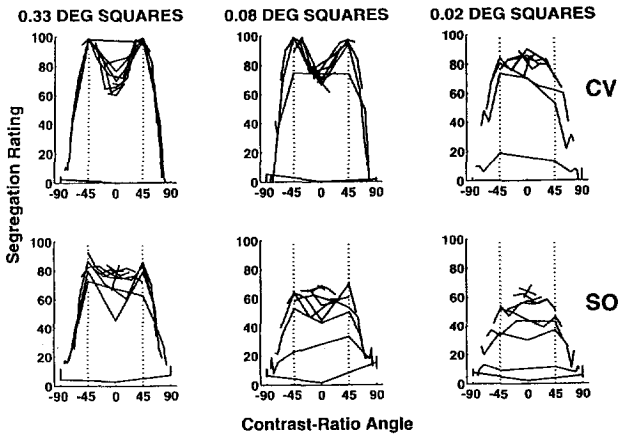

FIGURE 8. Some results from Expt 1 plotted in the same way as the predictions in Figs 6 and 7. Results from two observers are shown (rows) and three spatial scales (columns). In each panel, segregation rating is plotted as a function of contrast-ratio angle. Each curve is for a particular constant-difference series (with the difference being 0 steps for the bottom curves very near to the horizontal axes and 11 steps for the topmost curves). The step-size was 4% contrast for the left and middle columns and 8% for the right column. Each point is the average of four ratings.

accounted for by variation in a single parameter—the segregation threshold. More particularly, although the amount of compression at a given *absolute contrast* varies with spatial scale and background luminance, the amount of compression at a given *relative contrast* (where relative contrast equals absolute contrast divided by the segregation threshold) stays constant.

The second section below describes several quantitative features of this intensive nonlinearity—whether it is modeled as a relatively early local one or as inhibitory interaction—and describes its variation among individual observers.

The third section below discusses the spatial nonlinearity. The contribution of complex channels increases at larger spatial scales, as expected from their proposed structure, but the contribution does not change with background luminance in the range studied here.

I. Results About Models of the Intensive Nonlinearity

Figure 8 shows representative results from Expt. 1 plotted in the same form as the predictions shown in Figs 6 and 7 for two observers (two rows) at all three spatial scales (three columns). Notice that the ends of the curves in Fig. 8 bend down more and therefore overlap more at the larger spatial scales (left columns) than at the smaller scales. To put it another way, the results at larger spatial scales are more compressive (act like the higher-numbered members of the early–local nonlinearity family shown in Fig. 6 or like the higher values of w_O / σ shown in Fig. 7) than those at small scales.

When the models incorporating the two kinds of intensive nonlinearity are fit to the results for any particular observer and any particular pattern (e.g. to any of the six panels of Fig. 8) they both produce equally good fits. These fits are generally very good, explaining between 95 and 99% of the variance in the average segregation ratings. (See the Appendix for more details.) One cannot choose between the two kinds of intensive nonlinearity, therefore, on the basis of the goodness of these fits. The next sections explore instead the effect of spatial scale and background luminance on the parameter values estimated from the best-fitting versions of both models. The systematic variation in these parameter values is informative.

Effect of Spatial Scale on the Deduced Early, Local Nonlinearity. The curves in Fig. 9 show the early, local functions—$r(\Delta L)$ in Eqns (4) and (5)—from the best fits of the early, local model [Eqn (6)] to the results of Expt 1. Each panel shows the functions for the three different spatial scales for an individual observer. The horizontal axis gives element contrast (or ΔL) expressed in steps. The vertical axis gives Δr (the difference between the response to the element luminance and the response to the background luminance). The symbols just above the horizontal axis show the segregation thresholds and will be discussed later. Quite clearly the shapes of the functions at different spatial scales are not identical. They are more compressive (less linear) for larger spatial

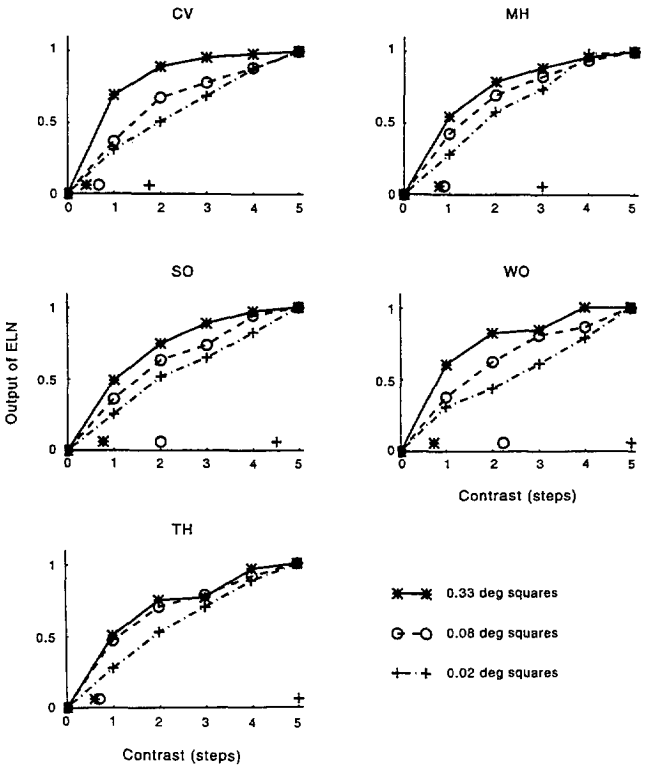

FIGURE 9. The deduced early local function Δr as a function of element contrast (or ΔL) expressed in steps. (The step-size for contrast was 4% for the two larger scales and 8% for the smallest scale.) The different curves in each panel shows the results at the three spatial scales for a single observer. These functions come from the best fits of the early, local model to the results of Expt 1 under the constraint that $k = 2.0$, but these functions are essentially identical even if k is allowed to vary. The functions were scaled to equal 1.0 at the largest contrast used to clearly show changes in shape. (The multiplicative constant is not determined by the model fit.) The symbols just above the horizontal axis show the segregation thresholds. The panels for observers CV and SO are the functions fit to the data in Fig. 8. For observers WO and TH, the threshold for the smallest scale was actually equal to six steps but is plotted at the maximum shown on the horizontal axis.

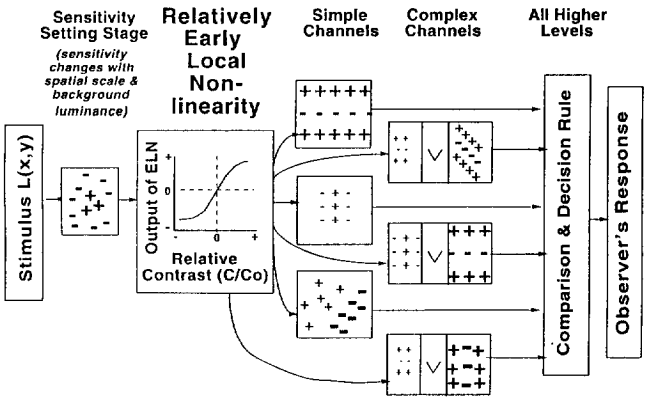

FIGURE 10. Modified version of the early–local nonlinearity hypothesis. A sensitivity-setting stage has been inserted before the early–local nonlinearity in the original model (Fig. 1). The sensitivity-setting stage may be thought of as a linear filter that attenuates some spatial frequencies and orientations more than others, but which has parameters that depend on background luminance. The modified model in this figure is consistent with the texture-segregation results while the original early–local nonlinearity model (Fig. 1) is not.

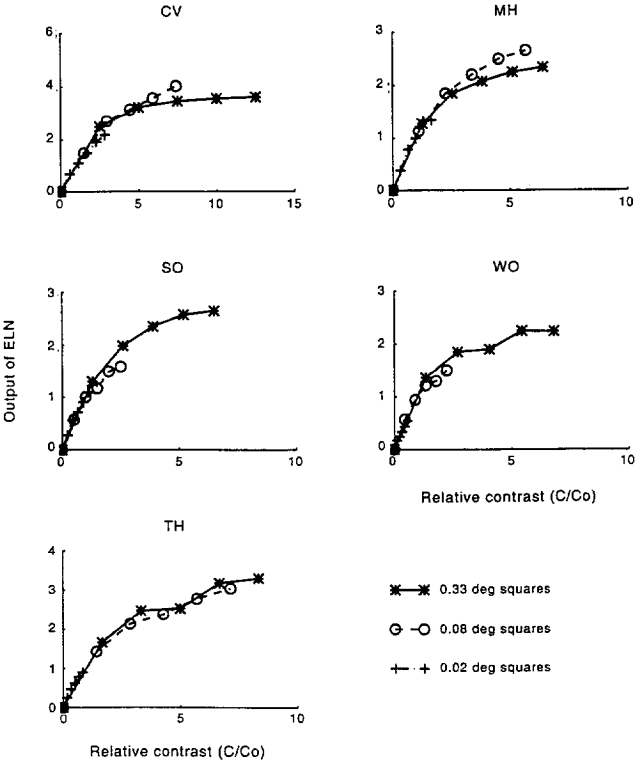

FIGURE 11. Deduced early–local functions from Fig. 9 plotted against relative contrast (contrast c divided by segregation threshold c_0). Each function was vertically scaled so it is equal to 1.0 at the segregation threshold (when $c/c_0 = 1$). The functions in this figure superimpose as they should if the modified early–local model in Fig. 10 is correct.

scales than for smaller. This difference is even more dramatic than it looks in Fig. 9, since in that figure the contrast step-size was twice as great for the smallest spatial scale as for the other scales. This same effect of spatial scale occurred in the results of Expt 2 although less dramatically since the spatial scale varied less.

The model in Fig. 1 assumes there is a single early–local, nonlinearity occurring before the channels and acting directly point-by-point on the stimulus luminance. Thus, according to this model, the deduced early–local functions should be identical for all patterns. It clearly is not. This is evidence against the early–local nonlinearity hypothesis of Fig. 1.

However, this version of the early nonlinear hypothesis is very strong—it says that the compressive nonlinearity acts directly, point-by-point, on the luminance values. To dramatize its strength, note that it says that the compression acts before optical blurring as well as before all other spatial filtering. This is clearly too strong. Whenever investigators speak of textures having only two intensity values (black and white), they are making the same assumption, however; it might be a good enough approximation in some circumstances, but it certainly is not here.

A Relatively Early–Local Nonlinearity. A modified version of the early–local, nonlinearity hypothesis is shown in Fig. 10. Here the local nonlinearity still occurs relatively early (i.e., before the channels) and still occurs locally (point-by-point), but, unlike the original early–local model (Fig. 1), it does not act directly on the luminance. Rather it acts point by point on the output of a sensitivity-setting stage. This sensitivity-setting stage incorporates both the effects of optical blurring and of all neural processes (e.g. retinal-ganglion cells) that adjust sensitivities to different orientations and spatial frequencies prior to the simple and/or complex channels. Further, as will be discussed in connection with Expt 2, the effects of light adaptation are incorporated into this early sensitivity-setting stage by having its sensitivity to different spatial frequencies and orientations depend on background luminance.

Rather than computing the predictions from this relatively early–local model in full detail, we continued the approximate-equation approach we have been using. While inexact, this approach seems quite adequate here and has the advantage of transparency and simplicity as well as the savings in computation. On this approximate approach, a single parameter needs to be estimated for each spatial scale of pattern; this parameter represents the contrast threshold of the sensitivity-setting stage. As it turned out, taking the measured segregation threshold c_0 (plotted near the horizontal axes of Fig. 9) to be this parameter worked very well so we did not try any other estimation schemes. The ratio c/c_0 is taken to be an estimate of the output of the sensitivity-setting stage in Fig. 10 or, equivalently, of the input to the early–local stage. Hence, to test the relatively early–local model of Fig. 10, the early–local functions deduced from the best fits of Eqn (6) should be replotted against the quantity (c/c_0) with the output at $c/c_0 = 1$ set equal to a constant. If the modified model is correct, these functions should superimpose. (There is no guarantee that this procedure will work to produce superimposed functions; it will definitely not work if the deduced early–local functions for different spatial scales have different shapes on log–log axes.) This procedure is carried out on the functions of Fig. 9 and the results illustrated in Fig. 11. It is quite

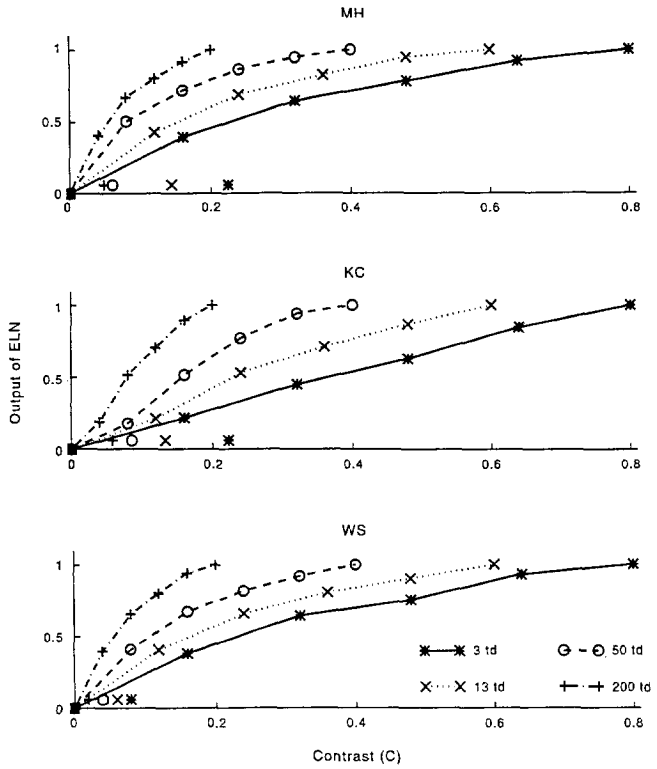

FIGURE 12. Deduced early–local functions from Expt 2 at different background luminances and the medium spatial scale. The horizontal axis shows contrast. The vertical axis shows the early local function scaled to be 1.0 at the maximum contrast used at each background luminance. Each panel shows the results for one observer. Only one spatial scale (0.17 deg squares) is shown for visual clarity, but the results of varying background luminance were similar at all scales. The symbols near the horizontal axis indicate the segregation thresholds.

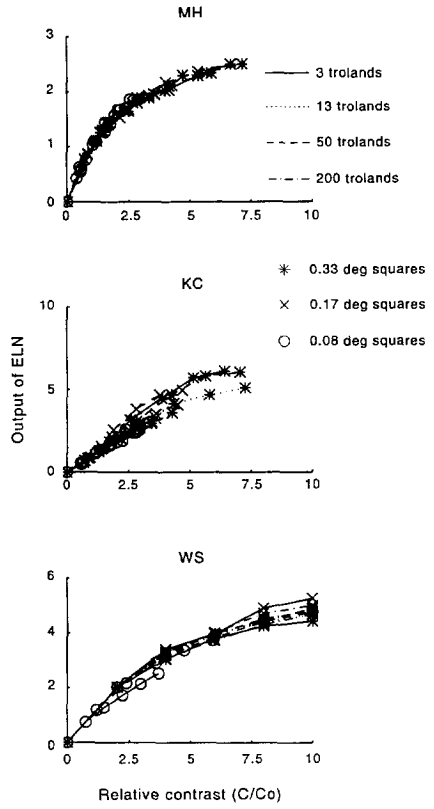

FIGURE 13. Deduced early–local functions for all conditions of Expt 2 plotted against relative contrast (c/c_0). Each function was vertically scaled so it is equal to 1.0 at the segregation threshold (when $c/c_0 = 1$). The functions in this figure superimpose as they should if the modified early–local model in Fig. 10 is correct.

clear that the procedure did produce superimposed functions in this case. That is, the relatively early–local model (Fig. 10) works well for Expt 1; further, the superimposed curves in Fig. 11 provide an estimate of the relatively early–local transformation in that model.

Effect of Background Luminance Interpreted Within the Relatively Early–Local Model. Since effects of background luminance are often thought to occur at a very early sensitivity-setting stage, it is natural to wonder whether the modified early–local model of Fig. 10 can also account for the effects of varying background luminance. The four curves in any panel of Fig. 12 show the deduced early–local functions plotted against contrast for the four background luminances of Expt 2. (Only the medium-spatial-scale results are shown for the sake of visual clarity. Again the symbols just above the horizontal axis represent segregation thresholds.) Note that the compressiveness at any given contrast decreases with decreasing background luminance. For example, at a contrast of 0.20, the function for the lowest background luminance (*) is still quite linear (absolutely linear for observer KC), while those for the highest two background luminances (○ and ×) are quite compressive.

Figure 13 shows all the deduced early–local functions from Expt 2 (all spatial scales and mean luminances) plotted against relative contrast c/c_0. These functions superimpose very well. (Observer KC's results in the middle panel show the most variability. The estimated standard deviation of KC's responses to any single stimulus is about twice that of any of the other observers in Expts 1 or 2.) Thus, the relatively early–local model (Fig. 10) can account well both for the effects of varying background luminance and spatial scale.

Interchannel Inhibitory Interaction (Normalization Model). The left and right panels of Fig. 14 are, for the normalization hypothesis, the analogs of Figs 9 and 11, respectively, for the early–local hypothesis. The quantities plotted on the vertical axes of the two panels summarize the best fits of the normalization Eqn (8) to results of Expt 1. The two panels show the appropriate measures of compressiveness if the input to the normalization network were proportional to contrast c (left panel) or to relative contrast c/c_0 (right panel). The horizontal axis gives spatial scale, and each curve is for a different observer. As with Fig. 9, the left panel of Fig. 14 shows greater compressiveness at larger spatial scales than at smaller scales. As with Fig. 11, the right panel of Fig. 14 shows approximately constant compressiveness. In short, the parameter determining compressiveness is only invariant when relative contrast—not contrast—is taken to be the input to the normalization network. The same result holds for Expt 2.

In the normalization model as shown in Fig. 2, the

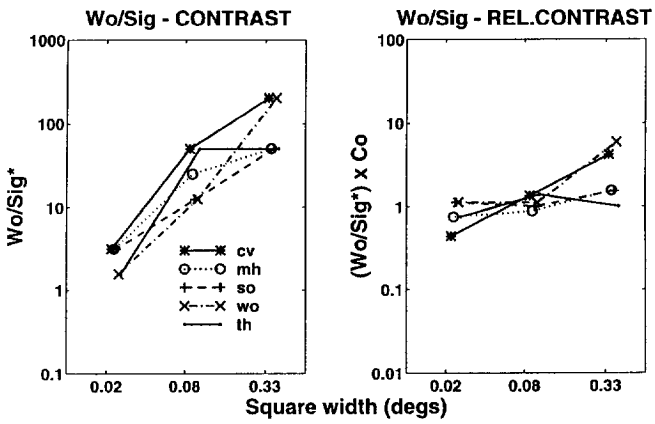

FIGURE 14. The fit of the normalization model to Expt 1. In each panel, the five curves come from the five different subjects and the horizontal axis gives the spatial scale represented as the width of the square elements. (The symbols were offset slightly to better reveal individual points.) The quantities plotted on the vertical axes come from the best fit of the normalization model. The left panel's vertical axis shows the quantity $w_O/\sigma^* = w_O/(\sigma \cdot c_{step})$. This quantity is a measure of the degree of compressiveness at a given physical contrast. The right panel's vertical axis shows the product $(w_O/\sigma^*) \cdot c_0$. This product represents the degree of compressiveness at a given relative contrast.

input to the normalization network is relative contrast (not contrast) since the channels' spatial-filtering action is shown as coming before the inhibitory interaction of the network. The constancy of compressiveness considered as a function of relative contrast (the approximate flatness of the curves in the right panel of Fig. 14) implies that this model is consistent with the experimental results.

On the other hand, the inconstancy of compressiveness as a function of contrast (the lack of constancy in the left panel of Fig. 14) implies that the results are *not* consistent with a model in which the input to the normalization network is directly proportional to contrast (as it would be if the normalization network were to operate on the inputs to the channels in Fig. 2).

Is all Sensitivity-Setting Prior to the Texture Intensive Nonlinearity? Thus, within the framework of either early–local nonlinearities or of normalization networks, the observer's sensitivity to different spatial frequencies and background luminances is set prior to the process producing the texture intensive nonlinearity.

Could, however, some of the sensitivity change still occur after the intensive nonlinearity although much occurs before? The fact that the curves are so well juxtaposed in Fig. 11 suggests the answer is "no", that all of the sensitivity change must be occurring before the early–local nonlinear function rather than after. The curves in the analogous plot for the normalization model (the right panel of Fig. 14) are not quite so convincingly flat, but they are flat enough to leave very little room for sensitivity changes occurring after the normalization network. Thus, we will conclude that, for the ranges of spatial frequency and background luminance studied here, most if not all of the sensitivity changes occur before the intensive nonlinearity.

II. Quantitative Characteristics of the Intensive Nonlinearity

Linear at Low Relative Contrasts, Logarithmic at High. Figure 15 shows the relatively early–local functions from Figs 11 and 13 replotted on log–log axes (○). At lower relative contrasts, the data points are fit well by a line of slope 1—i.e., they exhibit linear behavior. At higher relative contrasts, the points break away from the line of slope 1 and tend to a much shallower slope of about 1/2, showing that, at higher contrasts, these functions are approximately logarithmic. (Semilogarithmic plots illustrate this last point directly.) In the range of high relative contrasts where these functions are logarithmic, it is only the ratio of contrasts in the two elements—not their actual values—that determines segregatability. On plots like those in Figs 6–8 which are plots as a function of contrast-ratio angle, the constant-difference curves in that range of high relative contrasts (the ends of the curves) completely superimpose.

By assumption the normalization model necessarily embodies this quality of linear behavior at low contrast and logarithmic-type behavior at high contrasts. At relatively low contrasts, the predicted segregation depends linearly on the contrast (σ dominates the other terms in the denominator) and at higher relative contrasts, it depends only on the ratio of ΔL_1 to ΔL_2 (as algebraic derivation shows when σ is negligible).

Individual Observers. Also apparent in Fig. 15 is a

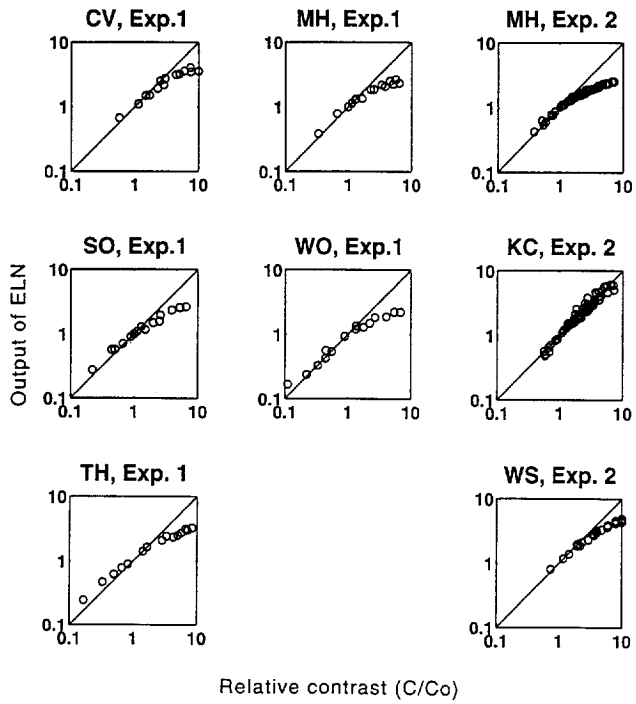

FIGURE 15. The scaled early–local nonlinearities for all observers from Expts 1 and 2 on log–lot plots. The solid line has slope 1 and is plotted through the point (1, 1), where all the functions were pinned. At low relative contrasts, the deduced early–local data points all lie near the line representing the linear function. At higher contrasts they deviate and tend toward a line of slope about 0.5, with the point of transition varying somewhat from subject to subject.

difference among individual observers. For six of the seven observers, the transition where the function ceases to be linear occurs when $c/c_0 = 2$ or 3. This corresponds to a contrast of < 10% for the highest background luminance and largest spatial scale. However, observer KC is substantially more linear. In fact, in Fig. 15 there is very little sign of a nonlinear range for observer KC at all. (Auxiliary experiments including even higher contrasts confirm that observer KC's results do show compression at high enough contrasts.) Fits of the normalization model also differentiate KC from the other observers: the estimated parameter measuring compressiveness is approximately 0.1 for KC rather than 1.0, as for the other observers.

Observer KC's segregation and detection thresholds are very similar to those of some other observers (e.g. KC vs MH in Fig. 12), so a difference in sensitivity does not seem to be connected to the difference between KC and the other observers. One does wonder if the greater linearity in KC's behavior is related to his greater variability. As mentioned earlier, the standard deviation of his responses to repeated presentations of the same stimulus was about twice that of the other observers. Unfortunately, KC was no longer working as a subject by the time his difference from other observers was noticed so no further experiments [e.g. forced-choice experiments like those of Sutter and Graham (1995)] could be done.

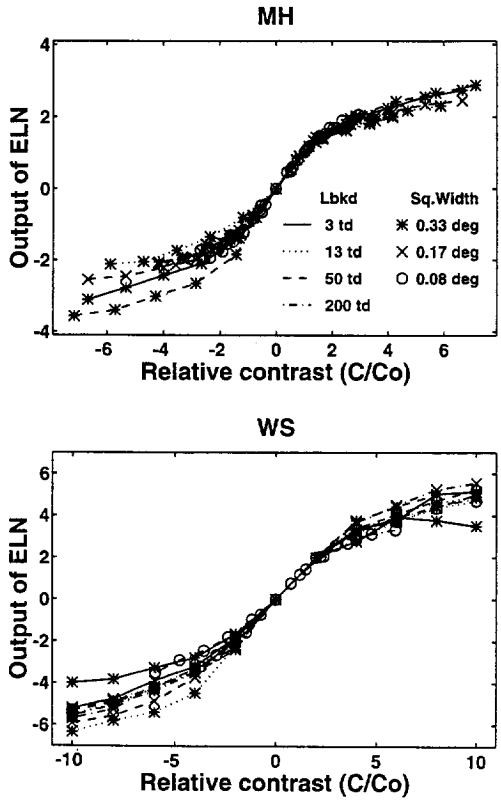

FIGURE 16. Deduced early–local functions when values at negative and positive contrasts were allowed to vary independently. Each panel shows the results from one observer of Expt 2 for all 12 conditions (4 background luminances × 3 spatial scales) plotted as a function of relative contrast. The function is approximately odd-symmetric although, particularly for MH, the left part of the function (for decrements) is slightly less compressive than the right part. [The 12 functions plotted in a given panel here are constrained to intersect at the origin and at (1, 1).]

Increments vs Decrements. For the fits shown in Figs 9, 11–13 the early–local function was assumed to be odd-symmetric, i.e.,

$$r(-\Delta L_1) = -r(\Delta L_1),$$

and thus only the function at positive values needed to be plotted. That this assumption is at least approximately correct can be seen in results plotted as in Fig. 8, where the left and right ends of the curves dip down by about the same amount.

One might well suspect, however, that decremental same-sign-of-contrast stimuli should be less compressed and therefore more segregatable than the corresponding incremental stimuli. (The effective adapting luminance is probably closer to the mean luminance than to the background luminance, and thus is lower for the decremental than for the incremental case.) It is easy to test this suspicion by considering matched pairs of same-sign-of-contrast patterns, where one member of the pair (incremental) has luminances above the background and the other (decremental) has luminances below background and the magnitudes of increments and decrements were matched. The suspicion is confirmed. There is definitely a statistically significant difference between increments and decrements. (For observer MH in Expt 2, the decrement pattern segregated more than the increment pattern in 82 of 120 matched pairs; the reverse was true in 31; and the other 7 cases were tied. For observer KC, these three numbers were 71, 47, and 2; and for observer WS, they were 86, 34, and 0.) But how substantial is this difference between increments and decrements?

Within the relatively early–local model, the difference between increments and decrements can be measured straightforwardly by allowing the early local function to be asymmetric when fitting the model to the data (see the Appendix). The functions estimated this way from two observers of Expt 2 are shown in Fig. 16. They are still, however, almost completely odd-symmetric. Careful inspection of the endpoints shows that there is some asymmetry, particularly for MH; the left side of the function (for decrements) is not quite as compressive as the right side (for increments). In general, there does not seem to be any regularity in Fig. 16 between the two observers as to which condition is most asymmetric and/or most compressed and which least. No other subject (in Expts 1 or 2) showed any greater asymmetry than MH, although several showed as much.

Thus, while there is asymmetry, it is a minor feature of these results. It is easily incorporated in the early–local model, as we did in fact in order to compute the curves shown in Fig. 16. Nor does it pose a serious problem for the normalization model as some earlier, more local nonlinearity (e.g. conventional retinal light adaptation) might well precede the normalization network and cause some asymmetry between the incremental and decre-

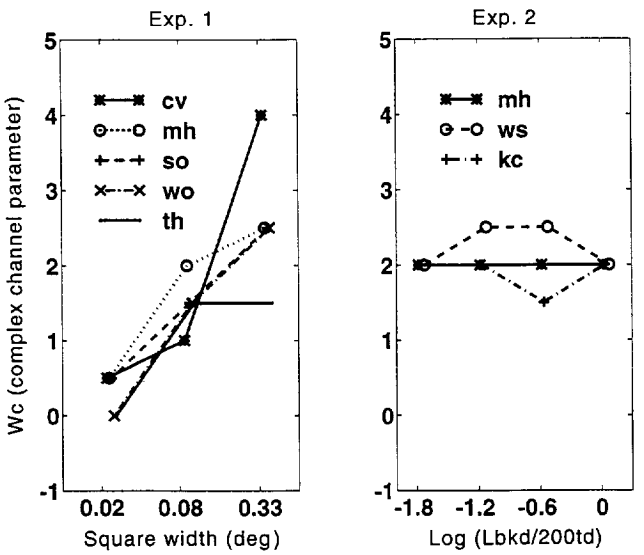

FIGURE 17. The effect of spatial scale on the estimated complex-channel weight (w_C) plotted as a function of square width (left panel: from Expt 1) or background luminance (right panel: from Expt 2 for 0.17 deg squares). Each curve is from a separate observer. (The symbols were offset slightly to better reveal individual points.) These estimates are from fitting an early–local nonlinear model with $w_S = 1$ and $k = 2$. The actual values of w_C should be treated with caution as there are differences due both to model and to the value of k (see the Appendix). The relative values of w_C as spatial scale or background luminance is changed can be considered seriously.

mental stimuli [as in, for example, the model of masking suggested by Bowen and Wilson (1994)].

III. The Spatial Nonlinearity (Complex Channels)

Effect of Spatial Scale and Background Luminance on the Complex Channels. As the spatial scale of the patterns becomes smaller, the whole spectrum moves to higher spatial frequencies where it is affected by the observer's high spatial frequency cutoff. Thus, the visibility of the higher harmonics relative to the fundamental frequency should decrease, and, since the complex-channel response may well depend on the higher harmonics, the complex-channel contribution should decrease. Visual inspection of the results in Fig. 8 suggests there is a less pronounced dip at the middle of the curves (a less pronounced signature of complex channels) for smaller than for larger scales. Figure 17 (left panel) shows the estimated complex-channel weights (w_C) from Expt 1. The complex-channel weight increases from something very close to zero for the smallest spatial scale to being two times the simple-channel weight for the largest scale. The effects of spatial scale in Expt 2 were similar although less dramatic since spatial scale was varied over a smaller range. (See legend to Fig. 8 and the Appendix for more details.)

If the relative visibility of the higher harmonics and fundamental frequency changed with background luminance, one would expect the complex-channel contribution to change as well. For the ranges of background luminances and spatial scales used here, however, the shape of the relevant part of the contrast sensitivity function did not change dramatically with background luminance, and, as Fig. 17 (right panel) illustrates, the estimated value of w_C was much the same at all background luminances.

DISCUSSION

Summary

I. About Models of the Intensive Nonlinearity. The effects of background luminance and spatial scale on the intensive nonlinearity in texture segregation can be accounted for by variation in a single parameter—the segregation threshold. (The segregation threshold is that contrast which, in the one-element-only version of a pattern, produces a criterion amount of segregation. Spatial scales from 0.75 to 12 c/deg in fundamental frequency and background luminances from 3 to 200 td were studied.) Specifically,

(a) there is greater compression at *a given absolute contrast* for larger than for smaller spatial scales and for higher than for lower background luminances; and (b) there is approximately the same amount of compression *at a given relative contrast* (where relative contrast is absolute contrast divided by the segregation threshold) for all spatial scales and for all background luminances studied here.

As a consequence, these experimental results *are not* consistent with the original version of the early–local, nonlinearity hypothesis (Fig. 1). All the results *are* consistent, however, with a modified version—the relatively early–local model—in which a sensitivity-setting stage precedes a local nonlinearity which precedes the channels (Fig. 10). They are also consistent with the normalization model in which inhibitory interaction occurs among the channel outputs (Fig. 2) although not if the interaction occurs among unfiltered channel inputs.

Thus, at least within the framework of these models and within the ranges of spatial frequency and background luminance studied here, the relative sensitivities to different spatial frequencies and background luminances are set prior to the compressive action of the intensive nonlinearity acting in texture segregation.

II. Quantitative Aspects of the Intensive Nonlinearity. The texture intensive nonlinearity acts linearly for low relative contrasts and logarithmically—depending only on contrast ratios—for high relative contrasts. (Very low contrasts were not used in this study.) For the highest background luminances (200–300 td) and largest spatial scale (0.75 c/deg fundamental frequency) studied, the texture intensive nonlinearity is very compressive, showing clear deviation from linearity by 10% contrast for six of the seven observers studied. The seventh observer is substantially more linear.

There is a small asymmetry between incremental and decremental stimuli with the decremental stimuli being more segregatable than incremental ones of the same magnitude. This is the expected effect if any light-adaptation process that came before the channels was

responding to something more like the space-average luminance than the background luminance.

III. The Spatial Nonlinearity. The contribution of complex channels (relative to that of simple channels) decreases as the spatial scale becomes smaller which is expected if, as postulated, the response of complex channels depends on the higher harmonics. The contribution of complex channels is not affected by background luminance in the range studied.

The Modified Early–Local vs the Normalization Hypotheses

The modified early–local model (Fig. 10) is quite different in a number of ways from the inhibitory interaction among channel outputs in a normalization network (Fig. 2). Yet, from the results here, we cannot distinguish them. In an earlier study (Beck *et al.*, 1991), we compared region-segregation judgments like those studied here with population-segregation judgments. (In the population-segregation task, observers were asked how well the elements of one type stood out from among elements of the other type when the elements were randomly intermixed rather than arranged in stripes and checkerboards.) If the intensive nonlinearity that acts in region segregation were local enough to act on single elements, it might have been expected to show up in population segregation also. But judgments in these two kinds of tasks were very different. This discrepancy suggested that the intensive nonlinearity acting in region segregation is *not* local enough to act on single elements and thus led us to favor the normalization hypothesis over any hypothesis of local nonlinearity; in fact, we still incline that way. On the other hand, the results of the earlier study could also be explained if the underlying compressive nonlinearity does act on individual elements and is the same for both population segregation and region segregation judgments, but population segregation is simply easier. To rule out this alternative explanation one will have to actually estimate the early–local nonlinearity in the population-segregation task. Meanwhile, one ought to keep an open mind on this issue.

This open-mindedness is, however, unfortunate in at least one way. Deductions about other aspects of region segregation—e.g., about properties of the complex channels—frequently depend on the nature of the intensive nonlinearity (and on other unknown aspects of the model such as pooling exponent). The extent of the contribution of the complex channels here, for example, was somewhat larger when estimated with the early–local model than with the normalization model. Similarly, the deduction of the actual bandwidth of the first-stage filters depends on the intensive nonlinearity (Graham *et al.*, 1993). It is quite likely that analogous confounding is present in many conclusions about different mechanisms of motion perception.

Comparisons of Compressiveness in Texture Segregation and Other Visual Processes

The question arises as to how the texture intensive nonlinearity compares to that measured for a number of other perceptual processes. While this comparison could be done within either the framework of the normalization model or of the early–local nonlinearity model, the latter is more convenient for these comparisons and is used here.

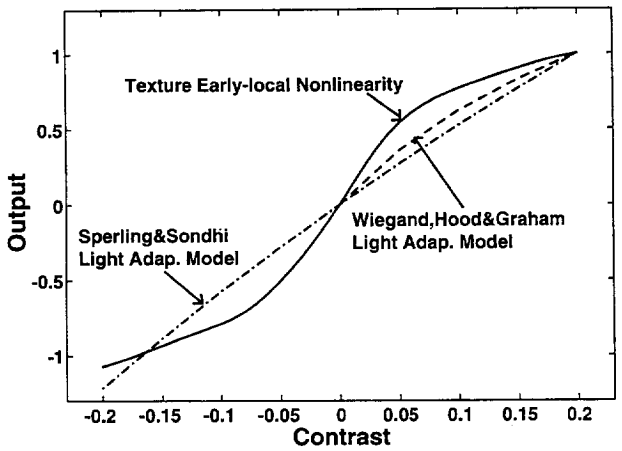

FIGURE 18. The intensive nonlinearity for texture segregation compared to predictions of the compressive nonlinearity from models of light adaptation. To make the comparison of shapes easier, the functions are all scaled to equal 1.0 at the highest contrast shown (20%). The solid line shows the early–local nonlinearity deduced from texture segregation. It is the splined version of the median curve from the seven observers at the largest spatial scale and highest background luminance. (The curve is a good representation of all observers except KC.) The – · – and – – – lines show predictions from models of psychophysical light-adaptation processes: the model of Sperling and Sondhi (1968) and that of Wiegand *et al.* (1995). What is plotted here for both light-adaptation models are the peaks in the responses to step increments (both models) and the trough in response to decrements (Sperling and Sondhi's model) from a steady background luminance of 600 td.

Light Adaptation. Figure 18 compares the texture-segregation intensive nonlinearity with the compression predicted by two models of light adaptation. The solid line shows the texture intensive nonlinearity at the highest background luminance and largest spatial scale. The - · - · - line shows predictions from the classic model of Sperling and Sondhi (1968). This model was developed to deal with contrast thresholds for various spatiotemporal stimuli on steady homogeneous backgrounds. The - - - - - line shows predictions from the more recent model of Wiegand, Hood and Graham (1995); also see Graham and Hood (1992) and Wiegand (1993). This recent model merges a model similar to that of Sperling and Sondhi (1968) with a model designed to explain the dramatic compressive nonlinearities that occur at transients as demonstrated in the probe-flash paradigm. Predictions from the more recent model are shown only for increments because the model's response to decrements has not been defined. [Initial tests suggest it should be much less compressive for decrements than for increments (Chase *et al.*, 1993).]

The predictions from these light-adaptation models are much less compressive than the texture intensive nonlinearity. This is particularly dramatic for decre-

ments, since the light-adaptation model of Sperling and Sondhi (1968) is actually slightly expansive for decrements: the larger the decrement, the greater the magnitude of the response to it.

This difference between the light-adaptation and texture nonlinearities would be difficult to handle in a framework where the texture intensive nonlinearity occurs as early or earlier than light adaptation (e.g. the original early–local model shown in Fig. 1). It is handled quite comfortably, however, within either the normalization model of Fig. 2 or the relatively early–local model of Fig. 10.

Suprathreshold Pattern Discriminations. Responses in some psychophysical experiments involving discriminations among suprathreshold patterns, however, are consistent with a very compressive intensive nonlinearity. For example, the minimum spatial frequency or orientation at which two suprathreshold gratings are just discriminable does not change once the contrasts are greater than several times detection threshold. It might, therefore, be modeled using a very compressive function, although other interpretations are possible (e.g. Campbell *et al.*, 1970; Smith & Thomas, 1989; Bowne, 1990). Similarly, the extent of binocular summation in orientation discrimination ceases to change for contrasts of 20% and above [with 5 c/deg stimuli (Bearse & Freeman, 1994)]. This kind of very compressive effect is also seen with temporally varying stimuli. Discrimination of the direction of motion reaches its maximum at contrasts of 2–5% (Derrington & Goddard, 1989), and the effect of uniform-field flicker masking increases no further once contrast is raised above 10% (Badcock & Smith, 1989). Pantle and Sekuler (1969) measured contrast-response functions for both the direction-selective and the non-direction-selective components of the threshold-elevating effect of adaptation to gratings (square wave with spatial frequency about 0.5 c/deg, velocity 6 deg/sec, and mean luminance somewhat less than 1 ft-L); the direction-selective component increased in response only up to about 16% contrast.

A psychophysical task of particular interest here is masking. The reduced segregatability of the same-sign-of-contrast patterns can be described as a kind of masking: the differences between the two regions are "masked" by the higher harmonics in the same-sign-of-contrast stimuli. (Remember that the same-sign-of-contrast effect in the texture segregation judgments is NOT due to an inability to tell the two kinds of elements apart; the two kinds of elements are frequently clearly discriminable from one another even in cases where there is no perceived segregation between the checkerboard and striped regions.) Underscoring the possible relationship between texture segregation and masking, a number of people have recently suggested that masking and related suprathreshold discrimination results might be accounted for by intracortical inhibition modeled by a normalization network like that used here for texture segregation (e.g. Bowen & Wilson, 1994; Foley, 1994; Heeger, 1994; Lubin, 1992; Lubin & Nachmias, 1990; Thomas *et al.*, 1993). Before these normalization explanations, masking experiments had typically been interpreted as the result of a nonlinear transducer in the channel detecting the test stimulus (e.g. Legge & Foley, 1980; Swanson *et al.*, 1984). While the nonlinear-transducer models may well turn out not to be the correct model of masking, the calculated nonlinear transducer provides a means of comparing the compressiveness in masking with that in texture segregation. The most compressive of the eight functions presented in one study of masking (Swanson *et al.*, 1984) is extremely similar to the texture nonlinearity; the nonlinear function used in another (Legge & Foley, 1980, Fig. 8) is, if anything, even more compressive. To compare masking and texture results more definitively, however, would require further work; one would need to compute the masking expected by some model(s) for the texture-segregation stimuli—perhaps, for example, assuming that the test stimulus was a grating at the texture fundamental frequency whereas the mask stimulus was a compound of higher frequencies like the higher harmonics in the textures. Whether that is worth doing at this time, given the possible inadequacies of all these models of masking (e.g. Nachmias, 1993), is not clear.

Neurophysiological Results. The response-contrast

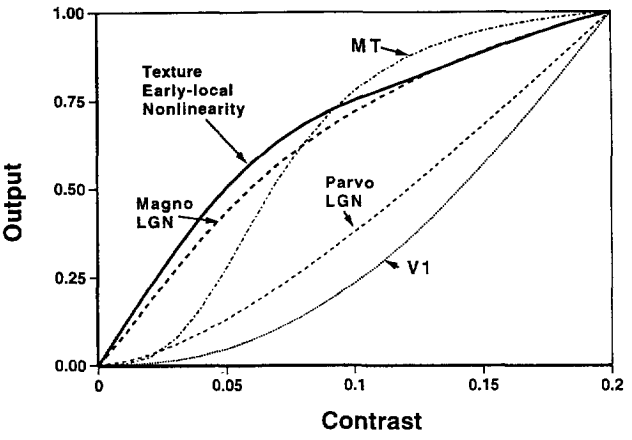

FIGURE 19. The intensive nonlinearity for texture segregation is compared to contrast-response functions for single neurons from four visual areas in the macaque monkey: the parvocellular and magnocellular layers of the lateral geniculate nucleus (LGN), the visual striate cortex (V1) and the middle temporal area (MT). These physiological functions use the median values of the parameters given in Table 1 of Sclar *et al.* (1990) for functions fitted to single neurons' firing rates plotted against the contrast of sinusoidal gratings drifting at each neuron's preferred values. The background luminance for the physiological experiments was 120–200 cd/m^2 with a 2.5 mm pupil. All functions are scaled to equal 1.0 at the highest contrast shown (20%). The value of the texture nonlinearity plotted here is the average of the absolute values at corresponding positive and negative contrasts in Fig. 18; in algebraic form this function is satisfactorily described by the following function which is linear at low contrasts and logarithmic at high:

$$f(c) = \alpha \cdot [\log_2(1 + c/\beta)],$$

where $\beta = 0.02$ and the value of α is set to make $f(0.20) = 1.0$. This function is plotted against the contrast of the elements in our pattern, not the contrast at the fundamental frequency. If replotted against contrast in the fundamental frequency, it would compress at even lower contrasts.

functions of neurons in a number of places of visual cortex have now been measured. Figure 19 shows the neural functions from one study of four visual areas (Sclar *et al.*, 1990). Each neuron's response was measured for various contrasts of a drifting sinusoidal grating (at the orientation, spatial frequency and velocity preferred by the neuron). The proper physiological quantity to compare to a psychophysical quantity is always a matter of some question, of course; some might argue that a signal/noise ratio would be more appropriate. However, the observer's rating of perceived segregation seems more like a magnitude of response (as plotted here for the neurons) than like a discrimination between two values (for which a signal/noise ratio might well be more appropriate) and, on a more practical note, response magnitude was the physiological measure easily available to use in this comparison. Figure 19 also shows the early–local nonlinearity inferred from the texture segregation at the highest background luminance and largest spatial scale. The functions from parvocellular LGN and from V1 cells are not nearly as compressive as the function inferred from texture-segregation results; indeed the parvo-LGN and V1 cells are expansive in this range of contrasts. The contrast-response functions from magnocellular LGN (see also Kaplan & Shapley, 1986) and from MT, on the other hand, are as compressive as that from texture segregation.

What do Comparisons of Compressiveness Imply? What should be made of these comparisons even if one were completely happy with the choices for the horizontal and vertical axes? Can compressiveness be taken as a signature of "pathway" or "stream"? Should we conclude from Fig. 19 that perceived texture segregation is done by something called the magno-pathway? That seems, at best, premature. The number of assumptions going into such a conclusion would be immense. The functions from a higher area in another pathway (e.g. the parvo-pathway) might be as compressive as those recorded from the magno LGN and MT here. On the other hand, it is thought-provoking.

Rather than viewing compressiveness as a signature of pathway, it might seem more appealing (although only slightly less dubious) to view it as a signature of level in a hierarchy of visual processing (cf. Sclar *et al.*, 1990). A function that *is not* particularly compressive could represent a continuum of output magnitudes (gray-level images so to speak). However, a function that *is* very compressive is almost like a step function; and a step function basically just says yes or no to the stimulus, that is, it categorizes. Lower-level processes might typically produce continuously graded output, and higher-level perceptual processes might categorize the stimulus. If so, texture segregation judgments in these experiments seem to be affected by a process somewhere in the middle, not retinal and not at the highest levels of perception either. It would certainly be premature to believe this as fact (and remember that magno-LGN cells are already very compressive). But, again, the idea of compressiveness as a signature of level is thought-provoking.

REFERENCES

Badcock, D. R. & Smith, D. (1989). Uniform field flicker: Masking and facilitation. *Vision Research, 29,* 803–808.

Bearse, M. A. Jr & Freeman, R. D. (1994). Binocular summation in orientation discrimination depends on stimulus contrast and duration. *Vision Research, 34,* 19–29.

Beck, J., Graham, N. & Sutter, A. (1991). Lightness differences and the perceived segregation of regions and populations. *Perception and Psychophysics, 49,* 257–269.

Beck, J., Prazdny, K. & Rosenfeld, A. (1983). A theory of textural segmentation. In Beck, J., Hope, B. & Rosenfeld, A. (Eds), *Human and machine vision* (pp.1–38). New York: Academic Press.

Bergen, J. R. (1991). Theories of visual texture perception. In Regan, D. (Ed.), *Vision and visual dysfunction, Vol. 10B: Spatial vision.* New York: Macmillan.

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

Bonds, A. B. (1989). Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. *Visual Neuroscience, 2,* 41–55.

Bowen, R. W. & Wilson, H. R. (1994). A two-process analysis of pattern masking. *Vision Research, 34,* 645–657.

Bowne, S. F. (1990). Contrast discrimination cannot explain spatial frequency, orientation or temporal frequency discrimination. *Vision Research, 30,* 449–461.

Campbell, F. W., Nachmias, J. & Jukes, J. (1970). Spatial-frequency discrimination in human vision. *Journal of the Optical Society of America, 60,* 555–559.

Cannon, M. W. & Fullenkamp, S. C. (1993). Spatial interactions in apparent contrast: Individual differences in enhancement and suppression effects. *Vision Research, 33,* 1685–1695.

Chase V. M., Wiegand T., Hood D. & Graham N. (1993). Exploring the dynamics of light adaptation using a sinusoidally modulated background and a probe. *Investigative Ophthalmology and Visual Science (Suppl.), 34,* 1036.

Chubb, C. & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A, 5,* 1986–2007.

DeValois, K. K. & Tootell, R. B. (1983). Spatial-frequency-specific inhibition in cat striate cortex cells. *Journal of Physiology, 336,* 339–376.

Derrington, A. M. & Goddard, P. A. (1989). Failure of motion discrimination at high contrasts: Evidence for saturation. *Vision Research, 29,* 1767–1776.

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

Foley, J. M. (1994). Human luminance pattern–vision mechanisms: Masking experiments require a new model. *Journal of the Optical Society of America A, 11,* 1710–1719.

Gorea, A. & Papathomas, T. V. (1993). Double opponency as a generalized concept in texture segregation illustrated with stimuli defined by color, luminance, and orientation. *Journal of the Optical Society of America A, 10,* 1450–1462.

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

Graham, N. (1994a). Low-level visual processes in perceived region segregation. *Investigative Ophthalmology and Visual Science, 35,* Abstr. #1034.

Graham, N. (1994b). Nonlinearities in texture segregation. In *Higher-order processing in the visual system (Ciba Foundation Symposium No. 184)* (pp. 309–329). Chichester: Wiley.

Graham, N. & Hood, D. (1992). Modeling the dynamics of light adaptation: The merging of two traditions. *Vision Research, 32,* 1373–1393.

Graham, N., Beck, J. & Sutter, A. (1992a). Nonlinear processes in spatial-frequency channel models of perceived texture segregation:

Effects of sign and amount of contrast. *Vision Research, 32,* 719–743.

Graham, N. & Sutter, A. (1992). Intensive nonlinearities in perceived texture segregation. Presented at the *Annual Meeting of the Optical Society of America*, Albuquerque, N.M., Sept. 1992.

Graham, N., Sutter, A. & Venkatesan, C. (1992c). Characteristics of complex channels (second-order processes) in texture segregation. Presented at the *Annual Meeting of the Optical Society of America*, Albuquerque, N.M., Sept. 1992.

Graham, N., Sutter, A. & Venkatesan, C. (1993). Spatial-frequency- and orientation-selectivity of simple and complex channels in region segregation. *Vision Research, 33,* 1893–1911.

Graham, N., Sutter, A., Venkatesan, C. & Humaran, M. (1992b). Nonlinear processes in perceived region segregation: Orientation selectivity of complex channels. *Ophthalmic and Physiological Optics, 12,* 142–146.

Grossberg, S. & Mingolla, E. (1985). Neural dynamics of perceptual grouping: textures, boundaries, and emergent features. *Perception and Psychophysics, 38,* 141–171.

Heeger, D. J. (1991). Computation model of cat striate physiology. In Landy, M. S. & Movshon, J. A. (Eds), *Computational models of visual processing*. Cambridge, MA: MIT Press.

Heeger, D. J. (1994). The representation of visual stimuli in primary visual cortex. *Current Directions in Psychology, 3,* 159–163.

Hood, D. C. & Finkelstein, M. A. (1986). Sensitivity to light. In Boff, K. R., Kaufman, L. & Thomas, J. P. (Eds), *Handbook of perception and performance, Vol. I: Sensory processes and perception* (Chapter 5). New York: Wiley.

Kaplan, G. & Shapley, R. M. (1986). The primate retina contains two types of ganglion cells, with high and low contrast sensitivity. *Proceedings of the National Academy of Sciences U.S.A., 83,* 2755–2757.

Landy, M. S. & Bergen, J. R. (1991). Texture segregation and orientation gradient. *Vision Research, 31,* 679–691.

Legge, G. E. & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America, 70,* 1458–1471.

Lubin, J. (1992). Interactions among motion-sensitive mechanisms in human vision. Ph.D. Dissertation, University of Pennsylvania, Pittsburgh, PA.

Lubin, J. & Nachmias, J. (1990). Discrimination contours in an F / 3F stimulus space. *Investigative Ophthalmology and Visual Science (Suppl.), 31,* 409.

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

Morrone, M. C., Burr, D. C. & Maffei, L. (1982). Functional implications of cross-orientation inhibition of cortical visual cells. I. Neurophysiological evidence. *Proceedings of the Royal Society of London B, 216,* 335–354.

Nachmias, J. (1993). Masked detection of gratings: The standard model revisited. *Vision Research, 33,* 1359–1365.

Pantle, A. & Sekuler, R. (1969). Contrast response of human visual mechanisms sensitive to orientation and direction of motion. *Vision Research, 9,* 397–406.

Robson, J. G. (1980). Neural images: The physiological basis of spatial vision. In Harris, C. S. (Ed.), *Visual coding and adaptability* (pp. 177–214). Hillsdale, N.J.: Erlbaum.

Robson, J. G. (1988a). Linear and non-linear operations in the visual system. *Investigative Ophthalmology and Visual Science (Suppl.), 29,* 117.

Robson, J. G. (1988b). Linear and non-linear behavior of neurones in the visual cortex of the cat. Presented at *New Insights on Visual Cortex, the 16th Symposium of the Center for Visual Science*, University of Rochester, Rochester, N.Y., June 1988 (Abstr., p. 5).

Sclar, G., Maunsell, J. H. R. & Lennie, P. (1990). Coding of image contrast in central visual pathways of the macaque monkey. *Vision Research, 30,* 1–10.

Shapley, R. & Gordon, J. (1985). Nonlinearity in the perception of form. *Perception and Psychophysics, 37,* 84–88.

Smith, B. G. & Thomas, J. P. (1989). Why are some spatial discriminations independent of contrast? *Journal of the Optical Society of America A, 6,* 713–724.

Sperling, G. (1989). Three stages and two systems of visual processing. *Spatial Vision, 4,* 183–207.

Sperling, G. & Sondhi, M. M. (1968). Model for visual luminance discrimination and flicker detection. *Journal of the Optical Society of America, 58,* 1133–1145.

Sutter, A., Beck, J. & Graham, N. (1989). Contrast and spatial variables in texture segregation: Testing a simple spatial-frequency channels model. *Perception and Psychophysics, 46,* 312–332.

Sutter, A. & Graham, N. (1995). Investigating simple and complex mechanisms in texture segregation using the speed–accuracy tradeoff method. *Vision Research, 35,* 2825–2843.

Swanson, W. H., Wilson, H. R. & Giese, S. C. (1984). Contrast matching data predicted from contrast increment thresholds. *Vision Research, 24,* 63–75.

Thomas, J. P., Olzak, L. A. & Shimozaki, S. S. (1993). The role of Fourier components in discrimination between two types of plaid patterns. *Vision Research, 33,* 1573–1579.

Turano, K. & Pantle, A. (1989). On the mechanism that encodes the movement of contrast variations: Velocity discrimination. *Vision Research, 29,* 207–221.

Victor, J. D. & Conte, M. M. (1991). Spatial organization of nonlinear interactions in form perception. *Vision Research, 31,* 1457–1488.

Wiegand, T. (1993). A psychophysically-based computational model of the dynamics of light adaptation. Ph.D. Dissertation, Columbia University, N.Y.

Wiegand T., Hood D. & Graham N. (1995). Testing a computational model of light-adaptation dynamics. *Vision Research, 35,* 3037–3051.

Wilson, H. R. (1990). Psychophysics of contrast gain control. *Investigative Ophthalmology and Visual Science (Suppl.) 31,* 430.

Wilson, H. R., Ferrara, V. P. & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion pereception. *Visual Neuroscience, 9,* 79–97.

Acknowledgements—This research was supported by National Eye Institute Grant 1 RO1 EY08459. Preliminary accounts of some of this work appeared in Graham and Sutter (1992) and Graham (1994a,b).

APPENDIX

FITTING THE MODELS

As stated in the text, Eqn (6) for the early–local model and Eqn (8) for the normalization model were fitted to each half-matrix of data coming from a single observer and a single pattern (where each of the 66 data points in the half-matrix was the average of the observer's 4 responses to a given stimulus). This appendix provides details of the fitting beyond those given in the Methods section of the main text. The fitting procedure was very similar to that used by Graham *et al.* (1992a) and the reader is referred there for general motivation and explanation.

The Final Transformation from Model Predictions to Observer Ratings and Assessment of Goodness-of-Fit

Our fitting procedures here differ from those in our earlier paper in one major way. This modification was important in speeding up the process. This modification was in the final transformation F from the predictions of the model D_{ELN} [Eqn (6)] or D_{NORM} [Eqn (8)] to the observer's ratings. In our previous study we had required that this transformation F be monotonic. For practicality we had restricted it to belong to a particular family of monotonic function—the Quick–Weibull family [see Fig. 16 of Graham *et al.* (1992a)]. Using the fmins procedure in Matlab (which is an instantiation of the Nelder–Meade algorithm), we found the best-fitting member of this family where the least-squares r^2 (the square of the Pearson product–moment correlation coefficient between the transformed predictions and the data) was used as the measure of fit. This old procedure was effective and sensible but (on the available computers) rather slow. For the current study,

therefore, we represented the final transformation F by a polynomial of fourth degree. Therefore, we could use standard methods of fitting the polynomial which were much faster (polyfit in Matlab—again the least-squares r^2 is the measure of fit).

The fact, however, that the fourth-degree polynomial can be nonmonotonic produced spurious best fits occasionally. (Sometimes it bounced up and down to capture stray values or, worse, it sometimes became a distinctly U-shaped function, thus allowing clear and systematic violations of the model to be counted as good fits.) By doing some fits both the old and the new ways, and by visual examination of the fitted final transformation for a number of the new fits to the current experiments, we satisfied ourselves that only a small proportion of the new fits were contaminated this way and that they could be easily found by the following procedure. The fits for different values of the pooling parameter k were almost always extremely close to each other but, using the new fitting procedure, there were occasional cases (about 3%) where the best-fitting function r or w_O depended dramatically on which value of k was used. We examined these cases and found there was a pronounced nonmonotonicity in the final polynomial. Then we redid the fits in a less automatic fashion by throwing out all the fits based on extreme nonmonotonic polynomials and taking the next best fit. (This procedure has never involved any subtleties of judgment. The nonmonotonicities have occurred for parameter values in entirely the wrong parts of parameter space, and the next highest locally maximal r^2 has served to identify a perfectly reasonable fit.)

Parameter Choices Common to Both Models of Intensive Nonlinearity

Choice of k*—the exponent for pooling across channels.* For both the early–local nonlinearity and the normalization models, we used four different values of the pooling parameter k, namely 1, 2, 4, and 8. As it turned out, this parameter did not matter for the conclusions reported here. The deduced early–local nonlinearity was generally completely unaffected (and on the few occasions when it was affected, it was due to an artifactually good fit—as just discussed—which was then corrected). The best-fitting value of w_O in the normalization model was affected minimally. When affected, the value of w_O increased with k more often than it decreased. The value of w_C—the weight on the complex channels—was slightly affected (without affecting our conclusions) as will be discussed below. The results for $k = 2$ are shown in the figures in the text. (We choose 2 because many other investigators use that value routinely in the analogous places in their models, e.g. "energy" models.)

Choice of w_S *and* w_C*—the Weights on the Simple and Complex Channels.* For convenience in setting up the bookkeeping to fit the models, we always specified the luminances ΔL_1 and ΔL_2 in units of steps (i.e., integers from – 5 to + 5) rather than changing them to reflect the actual contrasts. Also, for both models we held the weight on the simple-channels w_S constant at 1.0 while using a number of different possible weights on the complex channels; specifically, w_C was set at values ranging from 0 to 5.0 in steps of 0.5.

Choices Specific to Fitting the Interchannel Inhibition (Normalization) Model.

For the normalization model, there are three further parameters: k_S (the spatial pooling parameter), w_O (the weight on "other" channels), and σ [the parameter in the denominator of Eqn (8)]. It is not reasonable to use the Nelder–Meade algorithm to find the best-fitting values for this set of three, however, as there are large ranges over which one or the other does not matter (e.g. due to tradeoffs between them) and in which therefore the algorithm will not converge. Instead we did crude grid searches with some further restrictions. For one thing, we held the value of the spatial-pooling parameter k_S constant at 2, since preliminary results suggested this parameter made less difference than the other two. Secondly, earlier calculations [e.g. those in Graham *et al.* (1992b) and some other unpublished ones] had shown that it is the ratio of w_O to σ (the weight on "other" channels divided by the parameter σ) rather than the absolute value of either that plays the largest role in determining the predictions. Thus, for this study, we used three disparate values of σ (0.001, 4, and 64), while varying w_O in a series of finer steps (0 and all numbers from 0.125 to 4096 by a factor of 2).

It again turned out to be the ratio w_O/σ that primarily matters. In Fig. 14 we report those parameter estimates when $\sigma = 4$. (The ones for $\sigma = 64$ would have been essentially the same except for occasional random events. Similarly those for $\sigma = 0.001$ would have been similar except in some cases where the fits were actually systematically worse.) Further, when examining the best-fitting values of the other-channel weight w_O we noticed the following: all low values of w_O ($w_O \leqslant 0.25$ for $\sigma = 4$) produced indistinguishable fits as did all high values of w_O ($w_O \geqslant 32$ for $\sigma = 4$). See the plots in Fig. 7. Thus, to prevent apparently large changes that were actually meaningless, we substituted the lower bound (i.e., 0.25 for $\sigma = 4$) for any lower best-fitting w_O and substituted the higher bound (i.e., 32 for $\sigma = 4$) for any higher best-fitting w_O when giving best-fitting values of w_O in Fig. 14.

To summarize, in fitting the normalization model to the results of this study, we searched for the best fit over the following values of the parameters:

- three disparate values of σ (0.001, 4, and 64) which could have been restricted to only one value ($\sigma = 4$) without loss,
- a series of values for the other-channel weight w_O, which (for $\sigma = 4$) could have been restricted without loss to the values from 0.25 to 32 by a factor of 2,
- the spatial-pooling exponent $k_S = 2$;

and also over the following which were used for both models:

- the channel-pooling exponent $k = 1$, 2, 4, and 8,
- the simple-channel weight $w_S = 1$ (while expressing the element contrasts in number of steps),
- values for the complex-channel weight w_C from 0.0 to 5.0 in steps of 0.5,
- a fourth-degree polynomial to represent the final transformation from predicted value D_{NORM} to the observer's response, which was allowed to vary under control of the Nelder–Meade algorithm.

The Normalization Model Parameters Plotted in Fig. 14.

To understand the quantity plotted on the vertical axes of Fig. 14, consider the following: as just discussed, it is the ratio w_O/σ rather than w_O alone that primarily determines compressiveness in a normalization model. Also the step size in Expt 1 varied from scale to scale (with the step-size being larger at the smallest scale so that the upper left panel underestimates the actual effect of scale). To take into consideration these two facts, the quantity reported on the vertical axis in Fig. 14's left panel is the ratio w_O/σ corrected for the contrast step-size c_{step} used in each condition, i.e.,

$$w_O/\sigma^* = w_O/(\sigma \cdot c_{step}). \quad \text{(A1)}$$

This quantity w_O/σ^* is the value of the ratio w_O/σ that would have led to the best fit if all the ΔL_i had been expressed in actual contrast values rather than being expressed, as they were in calculating the fits, in number of steps.

The quantity plotted on the vertical axis in the right panel is the product of w_O/σ^* times the segregation threshold c_0. This quantity is the value of w_O/σ that would have led to the best fit if the ΔL_i in Eqn (7) had been expressed relative to segregation threshold when doing the fits.

Interactions Among k, w_O *and* w_C *in Fits of the Normalization Model.* Although the numerical values of the goodness of fits of the normalization model did not depend on k, there were some interactions among k and w_O and w_C. In particular, as k increased (while holding σ constant), both the values of w_C and of w_O that produced the best fit increased.

In particular, as k increased from 1 to 8, the estimate of w_C typically increased by one unit. For example, the estimate of w_C might increase from 0.5 to 1.5. (Remember that w_S was held constant at 1.0; and w_C values spaced by 0.5 units were used.) More specifically, it increased by one unit about 70% of the time, by 0.5 unit about 10% of the time, and various other things happened 10% of the time. This increase occurred both for the early–local model and the normalization models to the same extent.

Quantitatively, the value of w_O (and thus w_O/σ) typically increased

by a factor of 2 as k increased from 1 to 8. As we make nothing of the absolute values of w_O in this paper, we will not discuss this interaction further except to say one more thing. Since the deduced early–local transformation in the early–local model did *not* depend on k at all (see below), the fact that the analogous parameter of the normalization model (w_O) did depend on k needs some explanation. It is probably connected to the increase in w_C with increases in k; this increase in w_C probably requires an increase in w_O to keep the amount of signal coming from the other channels high enough to produce normalization.

Choices Specific to Fitting the Early–Local Nonlinearity Model

In fitting the early–local-nonlinearity model, there was one feature quite different from the earlier procedure (Graham *et al.*, 1992a). There we had considered only a small family of possible functions (those shown in Fig. 6) as candidates for r in Eqns (4) and (5).

Here we wanted to explore the intensive nonlinearity in greater detail. Therefore, rather than assuming a certain form of function, we allowed the value of r in Eqns (4) and (5) to vary at each value of $|\Delta L_i|$ independently. In most of the fits reported here we assumed that the function was odd-symmetric, i.e., equally compressive for increments and decrements:

$$r(-\Delta L_1) = -r(\Delta L_1) \text{ and } r(0) = 0. \quad \text{(A2)}$$

Since there were then only five different nonzero magnitudes of ΔL_i used, and the final monotonic transformation between predictions and data trades off with the multiplicative unit of this early–local function, there were only four independent values of the early local function r. We chose to fix the value of the highest point—in particular, to fix $r(5 \text{ steps}) = 5$ and let the others vary—because fixing the value of the highest one produced the most stable behavior of the fitting routine. We used the Nelder–Meade algorithm as embodied in Matlab to find the best-fitting function $r(\Delta L_i)$. The starting values for the early local function $r(\Delta L_i)$ specified a completely linear function, and the goodness-of-fit between the values predicted by Eqn (6) and the data were assessed as described above.

To summarize, in fitting the early–local model of Eqn (6) to the results of this study, we searched for the best fit across the following parameter values:

- the four values determining an odd-symmetric early–local function r for 5 steps of contrast which were allowed to vary under control of the Nelder–Meade algorithm;

and also over the following which were used for both models:

- the channel pooling exponent $k = 1$, 2, 4, and 8,
- the simple-channel weight $w_S = 1$ (while expressing the element contrasts in number of steps),
- values for the complex-channel weight w_C from 0.0 to 5.0 in steps of 0.5,
- a fourth-degree polynomial to represent the final transformation from predicted value D_{ELN} to the observer's response, which was allowed to vary under control of the Nelder–Meade algorithm.

Interactions Among Parameters. In the Results, the early–local nonlinearities that are reported (Fig. 9 and Figs 11–13) are, strictly speaking, those from the best fits holding $k = 2$. The values of the early–local function depended minimally on k, however, and these figures would look essentially identical if plotted for any other k.

On the other hand, the estimate of w_C typically increased with k in very much the same manner as it did for the normalization model. The estimates shown in Fig. 17 are in the middle of the range.

Further, the complex-channel weights for a given value of k were slightly greater when estimated by the early–local model than by the normalization model. While the difference is small, it is very systematic. For example, over the various conditions of Expts 1 and 2, the average difference between the estimates was 0.63 units with a standard error of 0.01 (for $k = 2$). Or to look at it another way, in 41 of the 51 cases, the weight estimated from the early–local model was greater than that estimated by the normalization model; 6 were ties; and 4 were reversed.

Subsidiary Set of Fits to Investigate Asymmetry in the Early–Local Function. In a subsidiary set of fits (producing Fig. 16), we allowed increments and decrements to act differently—i.e., we allowed r in Eqns (4) and (5) to be asymmetric. There are then 11 values of $r(\Delta L_i)$—5 decrements and 5 increments and zero. We fixed $r(0) = 0$ and $r(5 \text{ steps}) = 5$, leaving 9 free parameters which varied under control of the Nelder–Meade algorithm.

For this subsidiary set of fits, we used only the data from same-sign-of-contrast and one-element-only patterns in order to speed up the process; thus the value of the weight on the complex channels was irrelevant.

Obtained Goodness-of-Fit for the Two Models

Each model was fitted separately to each of the 51 half-matrices of data (15 from Expt 1 for 5 observers at each of 3 spatial scales plus 36 from Expt 2 for 3 observers at each of 36 combinations of background luminance and spatial scale). For concreteness, the reader can assume the following summary is about the values of r^2 from the very best fits for each model, that is when all parameters were allowed to vary (including k, and, for the normalization model, σ). However, the summary applies equally well to any of various ways of looking at the fits (in particular, looking just at the fits holding $k = 2$ and $\sigma = 4$).

The fits of the two models behave extremely similarly. Both models produced r^2 values between 0.95 and 0.99 almost all of the time. Further, their fits were well correlated across data sets (i.e., for any observer–pattern combination where one model led to a particularly high or low correlation, so did the other model). The only exceptions to these two statements are the following:

(1) For subject KC (Expt 2) the 12 r^2 values for each of the models varied from 0.80 to 0.94. These low values are probably a consequence of KC's greater variability compared to that of other subjects: the estimated standard deviation of KC's responses to any stimulus was about twice that of any of the other six observers in these two experiments.
(2) There is one other r^2 value that is below 0.95; it is for observer TH in Expt 1 and is about 0.85 (for both models). The cause of this low value is unclear.
(3) There may be some subtle but systematic individual differences that involve differences between the models: in Expt 2, the normalization model fits better than the early–local model for 11 of the 12 conditions for subject WS but for only 2 of the 12 for subject MH. The difference in fits is, however, extremely small. The source of this effect is also not clear.

In short, there is no way to choose between these two models when restricting your attention to the goodness-of-fits done for single spatial scales and background luminances. Both models do a good job.