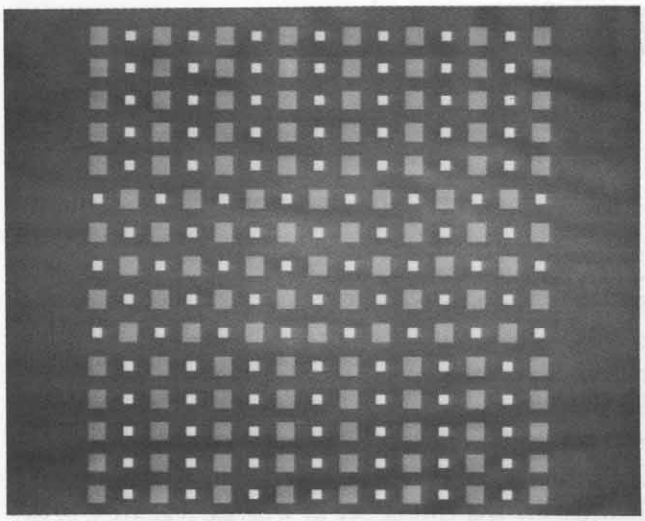
18

Complex Channels, Early Local Nonlinearities, and Normalization in Texture Segregation

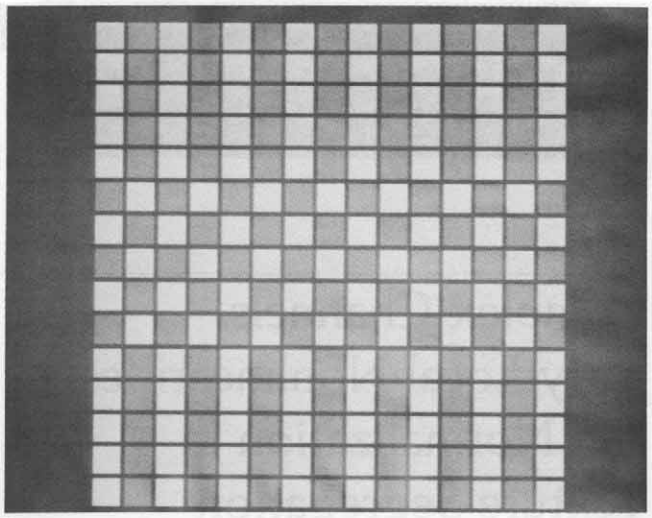
Norma Graham

Since the work of Jacob Beck and Bela Julesz and their colleagues in the 1960s and 1970s, it has been clear that two-dimensional spatial-frequency content (i.e., spatial-frequency and orientation) is a critical stimulus variable in how well differently patterned or textured regions of the visual field *segregate*, that is, how well they are perceived immediately and effortlessly as separate regions. (An excellent review of the perceived texture-segregation work can be found in Bergen, 1991.) This dependence of perceived segregation on spatial frequency and orientation —coupled with the mounting psychophysical and physiological evidence for low-level analyzers of spatial frequency and orientation in the visual system (reviewed, e.g., in Graham, 1985, 1989b)—has led a number of people to try to explain region (texture) segregation on the basis of such analyzers or sometimes, contrariwise, to try to prove that it cannot be so explained (e.g., chapter 17; Beck, Sutter & Ivry, 1987; Bovik, Clark & Geisler, 1987; Caelli, 1982, 1985, 1988; Chubb & Sperling 1988; Clark, Bovik & Geisler, 1987; Coggins & Jain, 1985; Daugman, 1987, 1988; Fogel & Sagi, 1989; Klein & Tyler, 1986; Landy & Bergen, 1988, 1989; Malik & Perona, 1989a,b; Nothdurft, 1985a,b; Turner, 1986; Victor, 1988; Victor & Conte, 1987, 1989a,b). Heavily influenced by Julesz's original conjecture about texture segregation—which was phrased in statistical language—discussions of texture segregation have often been mathematically sophisticated. (The relationship between such spatial-frequency analyzer models and Julesz's original statistical conjecture has been discussed in some detail by Klein and Tyler, 1986, and Victor, 1988. Related issues have been considered by Yellott and Iverson, 1990.) Only in the last years, however, has there been an emphasis on explicit and quantitative comparisons between psychophysical data on the one hand and predictions from computable models on the other.

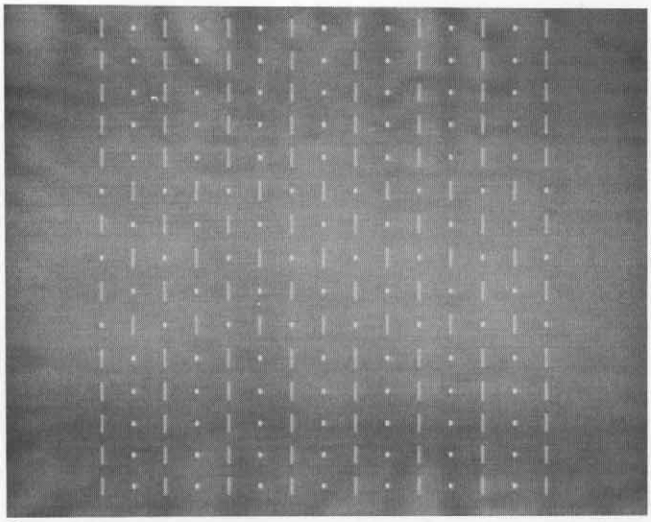
Jacob Beck, Anne Sutter, and I calculated quantitative predictions from a simple spatial-frequency analyzers



A



B



C

Fig. 18.1
Examples of element-arrangement textures. Identical elements are contained in all three regions but they are arranged in a checkerboard in the center region and in columns in the top and bottom regions.

model and compared these with quantitative measurements of region (texture) segregation collected for patterns like those in figure 18.1 (Graham, Beck & Sutter, 1989; Graham, Beck & Sutter, in preparation; Sutter, 1987; Sutter, Beck & Graham, 1989). These patterns are members of the class of *element-arrangement* texture patterns used originally by Beck, Prazdny, and Rosenfeld (1983) and explored further by Beck, Sutter, and Ivry (1987). They are composed of a uniform background on which are superimposed two types of elements arranged in stripes in the bottom and top regions and arranged in a checkerboard in the middle region. The observer's task is to indicate (on a scale from 0 to 4) to what degree the whole pattern seems effortlessly and immediately to contain two different kinds of region (striped vs. checkerboard).

Our simple spatial-frequency analyzers model is closely related to the class Chubb and Landy (see chapter 19) call *back pocket* models. We already knew on the basis of previous research that the simple model would work well but probably not perfectly. Our aim was to find out what such a simple model could do and, on the basis of systematic discrepancies between it and the data, to add further visual processes (either low- or higher-level) to the model and then again test the enhanced model. As it turned out, the comparisons between the predictions of the simple spatial-frequency analyzers model and the data from experiments using patterns like those in figure 18.1 were very revealing, showing both the simple spatial-frequency analyzers model's strengths and its weaknesses. The systematic discrepancies we found suggested adding two nonlinear processes to the simple model, where both nonlinearities are presumably the result of low-level (e.g., cortical V1 and V2 or below) visual processes. The second part of this chapter briefly describes these hypothesized nonlinear processes and presents some preliminary predictions from the enhanced model for one kind of experiment (a kind for which calculating predictions is easy if one is willing to make some reasonable simplifying

assumptions). These predictions are in very good agreement with the psychophysical results.

Simple Model

The simple spatial-frequency analyzers model we used consists both of assumptions about the analyzers themselves (described in Graham, 1989a, and Sutter, Beck & Graham, 1989) and assumptions relating the outputs of these analyzers to the responses produced by the observer in an experiment (described in Sutter, Beck & Graham, 1989). As in all such models, both sets of assumptions are crucial, although the second set is sometimes not explicit.

Characteristics of Simple Spatial-Frequency Channels (Linear Filters)

Each spatial-frequency channel in the model discussed first below is assumed to be a linear, translation-invariant filter. We described the spatial weighting functions characterizing each filter (the physiological substrate for which might be the receptive-field sensitivity profiles of the individual neurons) as a two-dimensional Gabor function (as used, e.g., by Bovik, Clark & Geisler, 1987; Clark, Bovik & Geisler, 1987; Daugman, 1987, 1988; Field, 1987; Fogel & Sagi, 1989; Rubenstein & Sagi, 1989; Turner, 1986; Watson, 1983), but the precise function makes little difference to these predictions (e.g., a pyramid computation like that used by Bergen and Adelson 1988 and Bergen and Landy, chapter 17, will produce similar enough filter outputs that it should make no difference to the conclusions here). The spatial-frequency and orientation half-amplitude full-bandwidths of each filter were 1 octave and 38° of rotation respectively (as in Watson, 1983) and in good agreement with results from near-threshold psychophysics (reviews in Graham, 1985, 1989b), but again, within rather a large range, the precise bandwidth makes little difference for anything said here. The filters' center spatial frequencies were separated by the squareroot of 2 over the visible range. Three orientations—vertical, horizontal, and half-way between (oblique)—were used. (These three are sufficient for the patterns under consideration here, and we used only three to save computing time; but for more general purposes one would want several more orientations.) The spatial weighting functions were all even-symmetric in the filter computations. (Given our full model, however, using multiple symmetries of weighting function at each position—that is, multiple phases—would make little or no difference to the predictions if the phases were combined as typically done by others. See discussion below.)

Figure 18.2 (middle and right panels) shows the outputs of two of these simple spatial-frequency channels (linear filters) to the pattern in the left panel. In this pattern, the two element types are squares, one had four times the area of the other, and the contrasts of the two element types were equal. The filter that is sensitive to the vertical orientation and to the fundamental spatial frequency of the striped regions (middle panel) has a spatial weighting function (i.e., receptive field) that will cover one column of squares with its center while covering the adjacent columns of squares with its inhibitory surround. This filter responds well in the top and bottom striped regions but responds little in the middle checkerboard region. A similar-sized obliquely oriented filter responds well in the checkerboard region and not at all well in the striped region. Thus these filters tuned to the fundamental frequency could easily contribute to an observer's ability to segregate the striped from the checkerboard regions.

Any filter tuned to spatial frequencies much higher than the fundamental (e.g., that producing the output in right panel of figure 18.2) has a weighting function much smaller than the period of the pattern. It will respond well (and, considering the whole of each region, about equally well) in both the checked and striped regions because: (1) its output is ample at the edges of all the individual elements (is proportional to edge contrast in fact); and (2) the total amount of element edge is the same in both

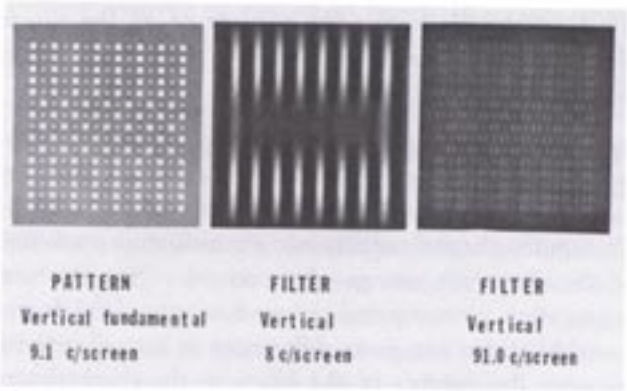


Fig. 18.2
A pattern (*left*) and the responses to it from two simple channels (*middle* and *right*).

the striped and checkerboard regions in these element-arrangement patterns. Thus simple linear channels sensitive to higher spatial frequencies can contribute little to segregation of the striped versus checked regions if any contribution depends on the difference between the total responses in the two regions (as it does in our simple model and in all models of the back pocket type).

Consider what happens, however, when the pattern in the left panel of figure 18.2 is replaced by one in which contrast of the small squares is raised to be a factor of 3 or 4 greater than the contrast of the large squares so that the product of area and contrast is approximately the same for the two sizes of squares. Now the channels sensitive to the fundamental do not respond at all well in either the checked or striped regions because excitation and inhibition balance out within each spatial weighting function (i.e, within each single neuron's receptive field). Thus these channels cannot contribute to perceived region segregation. Channels responsive to the higher spatial frequencies cannot contribute either, for the same reasons they could not for the pattern in figure 18.2. (They continue to respond well and equally well in both regions, since their response is to the edge of individual elements and the amount of and contrasts of element edges continues to be the same in both regions.) In short, neither the channels at the fundamental nor those at higher frequencies can contribute to segregation of the two regions. Thus our simple model (and all similar models) predicts that there should be a trade-off between contrast and area so that two regions differing in arrangement of elements (i.e., the striped versus the checkerboard regions) will not segregate well when the product of area and contrast is the same for the two types of elements. (Photographs illustrating the outputs of all the simple channels can be found in Graham, 1989a, and Sutter, Beck, and Graham 1989.)

From Channels' Outputs to an Observer's Responses

To turn the channel outputs into a quantitative prediction of the observer's ratings of perceived region (texture) segregation, we computed various measures of the degree to which there are gross differences in overall activity between the outputs of the filters to the checkerboard versus to the striped regions.

The first step is to compute a *spatially pooled response* from each channel in both the checked region and the striped region. Let the output at position (x, y) of the channel tuned to the ith frequency and jth orientation be called $O_{ij}(x, y)$. The spatially pooled response of the ijth channel to the checked region is taken to equal:

$$R_{ij}(ch) = \sqrt[k']{\frac{\sum_{(x,y)\ \text{in checked region}}^{N_x}\sum^{N_y} |O_{ij}(x,y) - A_{ij}|^{k'}}{N_x \cdot N_y}}, \quad (1)$$

where N_x and N_y are the numbers of spatial positions in the x and y directions in one period of the pattern and the summing is done over one period in the checked region. A_{ij} is the average value of $O_{ij}(x, y)$ over this one period. (The construction of the patterns assures that for each filter the values of A_{ij} are very similar in both the checked and striped region, and the fact that the filters are all bandpass means that these values are all very close to zero.) When the exponent k' is set equal to 2 the above measure is equal to the *standard deviation* of the outputs at different positions in one period of the given region. By crude analogy to other situations, this measure is also sometimes described as *energy*. We used exponents $k' =$ 1, 2, 3, 4, in the above formula as well as using the maximum output, the minimum output, and the maximum-minimum difference between the outputs at different positions. All conclusions given below held for all choices. The spatially pooled response in the striped region is exactly analogous to the definition in equation 1 for the checked region.

The difference between each filter's spatially pooled responses to the checked and to the striped regions is then computed yielding a *within-channel difference* for the ij^{th} filter of

$$Diff_{ij} = |R_{ij}(ch) - R_{ij}(st)|. \quad (2)$$

Finally, after weighting each within-channel difference by the observer's sensitivity to the corresponding spatial frequency and orientation, the within-channel differences are pooled across all channels to form R_{pool} given by the following definition:

$$R_{pool} = \sqrt[k]{\sum_{i}^{N_{freq}}\sum_{j}^{N_{orien}} \{Diff_{ij} \cdot S_{obs,ij}\}^k}, \quad (3)$$

where $S_{obs,ij}$ is the observer's contrast sensitivity to the ith frequency and jth orientation, N_{freq} is the number of

frequencies (usually 13—from 2 to 128 cycles/screen in steps of the square root of 2) and N_{orien} is the number of orientations (usually 3: horizontal, vertical, and oblique). With the exponent $k = 2$, R_{pool} is the root-mean-square difference between regions. We also use exponents of 1, 3, and 4 as well as infinity (taking the maximum of all the differences). All conclusions below hold for all choices. (Victor, 1988, and Chubb and Landy (chapter 19), also consider a whole family of rules, but typically calculations in the texture segregation literature have used pooling with an exponent equal to 2. An exponent of infinity corresponds to taking the channel that best discriminates the two regions which has sometimes been done in the psychophysical discrimination literature and was done in the multispatial-frequency filter model of Fogel and Sagi, 1989.)

The degree to which two regions (textures) segregate perceptually (as reflected by the observer's ratings of perceived segregation) is assumed to be *a monotonic function of* R_{pool}.

Before looking at some predictions of this model, let me briefly discuss alternatives to these assumptions relating the channels' outputs to the observer's responses.

Relationship of Our Pooling Measures to Other Approaches and Boundary Extraction

Instead of pooling across exactly one period of the pattern, which is like pooling with an abrupt-edged window exactly one period wide, one might use a gradual-edged window of about one period's width (about twice the width of the excitatory center of the corresponding filter's receptive field). With an exponent of 2, this is the local-energy measure which has been recently used in models of texture segregation and motion perception (e.g., chapter 17; Bergen, 1988; Bergen & Adelson, 1986, 1988; Landy & Bergen, 1988, 1989).

Instead of using only even-symmetric spatial weighting functions, one might use both odd-symmetric and even-symmetric ones (or any other pair having phase characteristics 90° apart) and then take the Euclidian (Pythagorean) sum of the outputs of the pair located at any internal position in the texture region (e.g., Adelson & Bergen, 1985; Bovik, Clark & Geisler, 1987; Clark, Bovik & Geisler, 1987; Fogel & Sagi, 1989, Turner, 1986). This will again produce essentially the same predictions as the pooling across a region using an exponent $k' = 2$ in the above. Either of these two alternative approaches can lead naturally into a model of how the boundaries between texture regions can actually be computed. In essence, one first computes, for each position in the pattern, the spatially pooled response or the sum of odd and even receptive fields. One then needs a process that ensures that all positions for which the set of these numbers from all channels is much the same belong to the same texture region (see, e.g., discussions in Bergen, 1991; Bovik, Clark & Geisler, 1987; Fogel & Sagi, 1989). This is an extremely interesting topic that has been discussed by a number of people (in addition to those just mentioned: Caelli, 1985, 1986, 1988; Grossberg, 1987; Grossberg & Mingolla 1985) although not explicitly studied here. Gradients at the boundary between two regions may well be important (chapter 17; Nothdurft, 1985b; Landy & Bergen, 1988, 1989). For simplicity, we currently ignore these gradients; this seems at least moderately safe for the results reviewed here since the spatial extent of the gradient in any one filter (the distance over which a local energy measure changes as one goes from the checked to the striped region, for example) is much the same for all the patterns used.

Contrast-Area Trade-off Experiments

In one series of experiments (Sutter, 1987; Sutter, Beck & Graham, 1989), we investigated the tradeoff between contrast and size of elements that is predicted by the simple spatial-frequency channels model (as discussed in connection with figure 18.2 above). The patterns all contained two element types that were the same shape but generally differed in size as in figure 18.1. The mean luminance of all the patterns in a given experiment was the same. Then the luminance (and hence contrast) of the larger element type was held constant while that of the smaller elements was varied and perceived texture segregation was measured. We did this for squares of different sizes, for different fundamental frequencies of patterns (different scalings of the overall patterns), for different duty cycles of patterns (different relationships between element size and interelement spacing), and for line-shaped elements as well as square-shaped elements. In each case, minimal segregation as rated by the observers occurred when the product of area times contrast was approximately equal for the two element types.

The predictions of the simple spatial-frequency channels model were computed for all the patterns used in all these experiments. As expected from the intuition discussed in connection with figure 18.2 above, the model always predicts minimal segregation when the product of area times contrast is approximately the same for the two element types (the exact contrast ratio producing the minimum depends on the duty-cycle of the pattern). In this respect the predictions of the model agreed with the experimental results.

However, the model did not predict all of the details of the experimental results correctly. While the contrast ratio at which the minimal segregation should occur was correctly predicted, the amount of segregation at this minimum was incorrectly predicted for many patterns. As is described briefly below and in much more detail in Sutter, Beck, and Graham (1989), we think we know how to modify the model by adding a "spatial nonlinearity" so that the enhanced model will make the correct predictions for the contrast-size tradeoff experiments.

Complex Channels: A Spatial (Rectification-type) Nonlinearity

The following modification of the simple model seems to resolve the discrepancies between our simple model and the contrast-area trade-off experimental results: Replace or supplement the totally linear spatial-frequency channels (simple channels as above) with "complex channels," as shown in figure 18.3. (The assumptions relating the channels' outputs to the observer's responses remain the same as in the simple model.) These complex channels have three stages: two stages of linear filtering with a pointwise nonlinearity (that is dramatic near zero) in between. For our purposes to date, the nonlinearity might be a full-wave rectification (that takes the absolute value of the first filter's output at each point in space), a half-wave rectification (that substitutes zero for the negative values and leaves the positive values untouched), a squarer (that squares the output at each point), or some other similar function. (See Heeger, chapter 9, for some discussion of the differences among these nonlinearities however.) These proposed complex channels are more complicated than the simple linear channels discussed earlier in much the same way that complex cells in area V1 of primate visual cortex are more complicated than

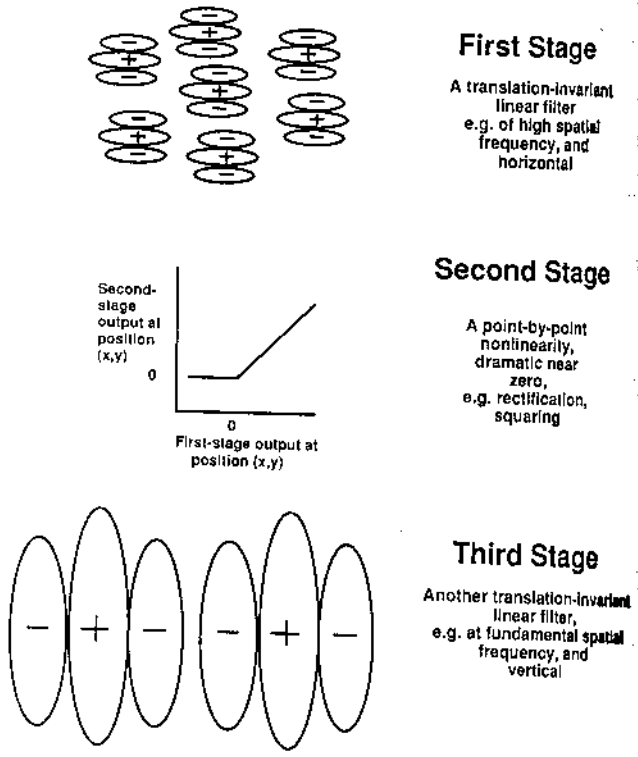
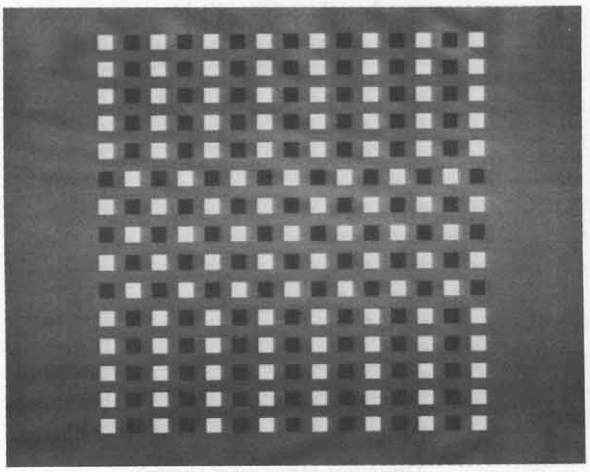


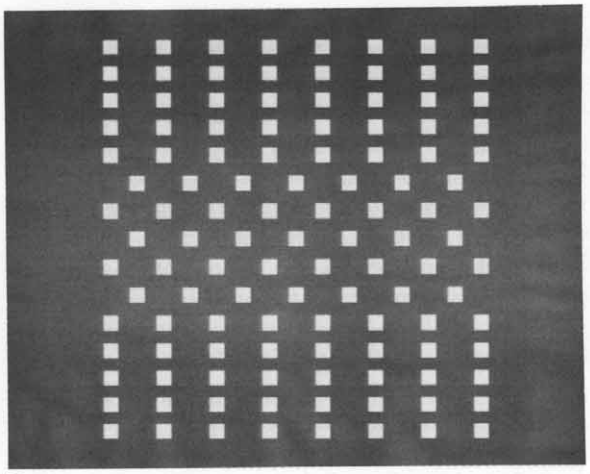
Fig. 18.3
Diagram of a complex channel.

simple cortical cells (e.g., Hochstein & Spitzer 1985). However, the bandpass nature of the third-stage filtering in the complex channels proposed here may not be confirmed for cortical complex cells; and, in any case, it is premature to take the possible physiological analogue of the complex channels too seriously. A number of other people have suggested the use of complex-cell-like computations in tasks like texture segregation (chapter 17; Bergen & Adelson, 1988; Chubb & Sperling, 1988; Fogel & Sagi, 1989; Grossberg & Mingolla, 1985; Robson, 1980; Sagi, 1989; Sperling, 1989; Sperling & Chubb, 1989). Some of these suggested complex-cell like computations, however, may be more like our simple model (simple linear channels followed by the nonlinear pooling operations contained in the assumptions relating channel outputs to observers' responses) than like our complex model (complex channels consisting of a linear-nonlinear-linear sandwich followed by the nonlinear pooling operations).

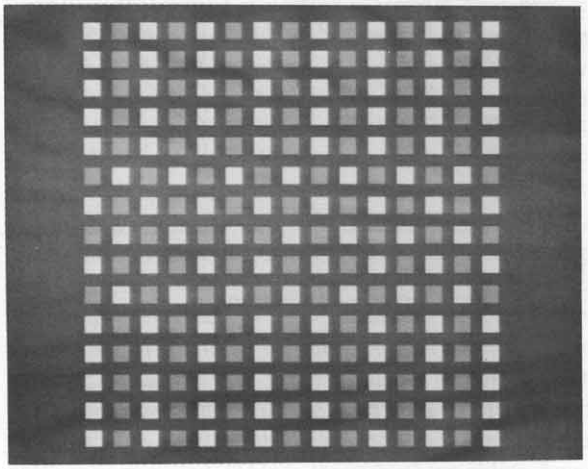
A qualitative description of how (i.e., insights into why) complex channels might explain all the discrepancies between the results of area-contrast trade-off experiments and the simple model's predictions can be found in



A



B



C

Fig. 18.4
Examples of element-arrangement textures used in *on-off* experiment. (A) An opposite-sign-of-contrast pattern. (B) A one-element-only pattern. (C) A same-sign-of-contrast pattern.

Sutter, Beck, and Graham (1989), but we have not yet done quantitative predictions. Briefly, the intuition is this: the kind of complex channel where the third-stage filter responds to much lower spatial frequencies than does the first-stage filter (as in the diagram of figure 18.3) will respond to *low-spatial-frequency patterns composed of high-spatial-frequency elements.* (To put it in terms of an auditory analogue that may be helpful to some people: this kind of complex channel is sensitive to the "missing fundamental".) Thus complex channels responding to higher harmonics can contribute to perceived segregation (although simple channels sensitive to higher harmonics do not, as was discussed in connection with figure 18.2). As it turns out, this response of the complex channels to higher harmonics can account for the discrepancies between the area-contrast trade-off experimental results and the predictions of the original simple model.

It is important to note that although a two-stage channel—just a pointwise nonlinearity followed by a linear filtering—works in principle to explain many previously-noted failures of simple linear channels (e.g., Peli, 1987), it will not work here.

On-Off Experiments

The rest of this chapter will describe another type of experiment—which we will call an *on-off* experiment—in somewhat more detail than were the area-contrast trade-off experiments. The results of the on-off experiments suggest the existence of (at least) two rather different nonlinearities: (l) the "spatial nonlinearity" embodied in the complex channels just described, and (2) an "intensity-dependent nonlinearity," for which we can propose at least two candidate processes known to occur at relatively low levels in the visual system.

For the patterns used in the on-off experiment (examples are shown in figure 18.4), the two types of elements were always squares of the same shape and size but differing in sign of contrast (i.e., lighter or darker than the

background—hence the "on" and "off" in the name of these experiments) and/or in amount of contrast. In the experiment described here, we used the set of stimuli diagrammed in figure 18.5, where each dot represents a stimulus; that is, we used the texture patterns defined by all possible pairs of a number of different contrasts. (Equivalently, since the background luminance was kept constant, all possible pairs of a number of different luminances were used.) There is only half a matrix of possibilities shown because the other half is presumably redundant. Notice that the contrast of any individual element relative to the background was never greater than 25 percent; this value is small for the literature on texture segregation (where black/white patterns, i.e., patterns of 100 percent contrast, predominate) but large for the literature on detection and identification of near-threshold patterns (from which much of the best evidence for the properties of spatial-frequency and orientation analyzers derives). (For the results shown here, the background luminance was 16 ft-L and the fundamental frequency of the striped texture pattern was 1 cycle/degree. The squares were slightly wider than the inter-square spaces.)

Constant-Difference Series and Simple-Model Predictions

Series of patterns like that illustrated in figure 18.6 are particularly interesting. In such a series the *difference* between the luminances of the two element types is held constant. The absolute luminances of the two element types vary together. In such a *constant-difference* series, which generally contains more patterns than the five shown here, there are patterns where both element types have the *same sign of contrast* (both darker or both lighter than the background), patterns where one element type has the same luminance as the background so there is *one element type only* apparent (which can be either dark or light), and patterns where the two element types are of *opposite sign of contrast*.

In the matrix-of-stimuli diagram (figure 18.5) any such series where $L_2 - L_1$ is constant occurs along lines parallel to the positive diagonal (bottom right). Notice also (bottom left) that all patterns on a line through the origin have the same incremental luminance ratio ($\Delta L_1 / \Delta L_2$) and also the same contrast ratio. For example, the patterns on the negative diagonal have element types with opposite (but equal) contrasts and thus a ratio of -1; all patterns

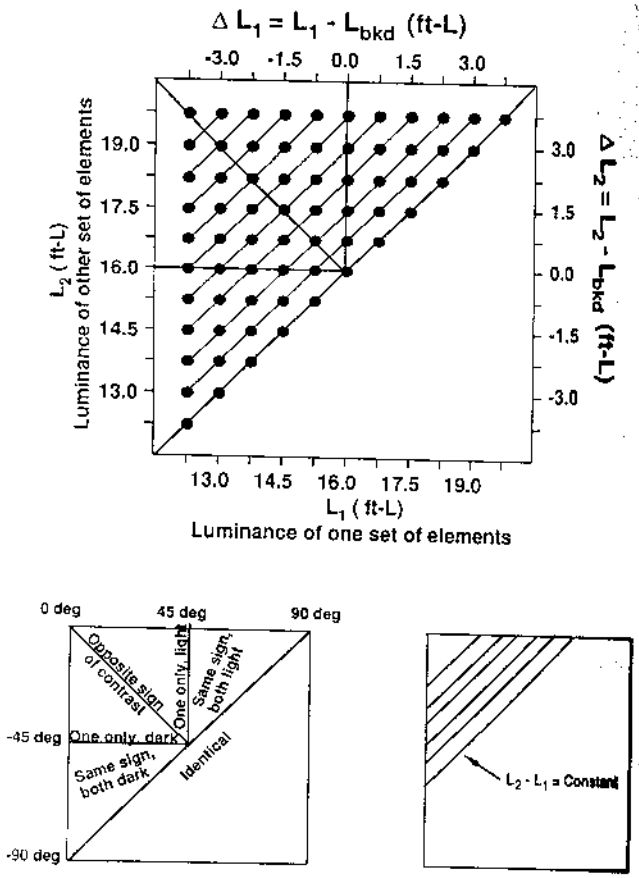


Fig. 18.5
Diagram of stimuli used in *on-off* experiment.

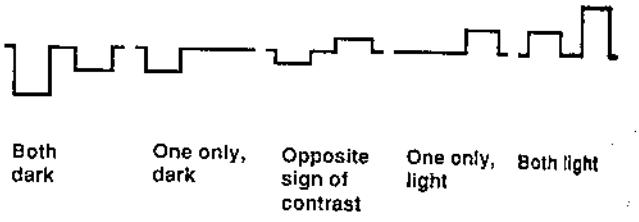


Fig. 18.6
Luminance profiles of the two element types from each of five patterns in a *constant-difference series* from an *on-off* experiment.

on the vertical ray upwards from the origin have one element type only, and that element type is bright. Rather than using L_2 and L_1 to describe a pattern, two other numbers have proved very useful: (1) the luminance difference $L_2 - L_1$, and (2) the angle, which we will call the *contrast-ratio angle*, between the negative diagonal in a plot such as figure 18.5 and the line going through the point representing a stimulus. (Some examples of contrast-

ratio angles are indicated in the bottom left of figure 18.5 on the left and top edges of the square.)

The constant-difference series of patterns are of particular interest because models of texture segregation involving multiple, linear, spatial-frequency and orientation channels (like our simple model) make a particularly simple prediction for such a series. These models predict that all members of a constant-difference series of patterns should show segregation to approximately the same degree (Graham, Beck & Sutter, 1989, and in preparation). Why do these simple linear models make this prediction? According to these models, segregation is based on the degree to which the linear spatial-frequency and orientation-selective channels show differences in amount of activity in the different texture regions. In other words, segregation depends approximately on the difference between the Fourier transforms' amplitudes in the checkerboard vs. the striped regions. For element-arrangement patterns like those used here, as mentioned earlier, the only substantial difference is at the fundamental frequency of the texture regions. And the fundamental frequency's amplitude is approximately constant throughout a constant-difference series of patterns; thus segregation is predicted to be approximately constant.

In terms of receptive fields (spatial weighting functions) the above explanation goes as follows. According to the simple models, segregation in element-arrangement texture patterns such as those of figures 18.1 and 18.4 is determined primarily by the receptive fields that are just big enough so that when a column of one element type stimulates the receptive field's excitatory center, then columns of the other element type stimulate the inhibitory surround. (This is the size of weighting function associated with the filter having the output displayed in the middle panel of figure 18.2.) Since for this size of receptive field the background is stimulating both the excitatory center and inhibitory surround to approximately the same extent, the effects of the background cancel out and it is only the two element types that contribute to the response. Further, for this size of receptive field in the appropriate position (that position producing the maximal response and thereby being the largest determinant of the whole channel's spatially pooled response), one element type stimulates the excitatory center and the other element type stimulates the surround; hence this receptive field essentially subtracts the effect of one element type from that of the other. Since in the on-off experiments the two element types have the same shape and size and differ only in luminance, this size of receptive field responds proportionally to the difference between the two element types' luminances.

In symbols, letting R_S be the pooled response of the simple channels (called R_{pool} earlier but distinguished here for reasons that will become clear later) and w_S the proportionality constant, then the following is true approximately (and the approximation is very good; Graham, Beck & Sutter, in preparation)

$$R_S = w_S \cdot |L_1 - L_2| = w_S \cdot |\Delta L_1 - \Delta L_2|. \quad (4)$$

The absolute value sign occurs because the quantity R_S is the value pooled across all such channels of the absolute value of the difference between each channel's spatially pooled response to one region (e.g., striped) and that to the other (e.g., checkerboard). Note that R_S is essentially entirely determined by the simple channels sensitive to the fundamental frequency of the texture regions. As mentioned earlier, the pooled response is assumed to be monotonic with (but not necessarily proportional to) the observer's segregation rating.

Empirical Results

Some empirical results are shown in figure 18.7. Each curve in the figure connects points representing patterns in a constant-difference series. The size of that constant difference increases from the bottom to the top curve. The horizontal axis shows the contrast-ratio angle of each pattern. The vertical axis gives the average observer rating of segregation.

As discussed above, the simple model predicts that all members of a constant-difference series segregate to the same extent approximately. Thus the predictions of the simple model would be approximately horizontal lines on this kind of plot. The results do not look at all like the predictions: First, each curve sinks dramatically at both its ends, with different curves actually converging for same-sign-of-contrast patterns when both element types' luminances are far from the background. (Thus, for the patterns at the ends of the curves, only the ratio of the contrasts in the two element types matters for segregation; the size of the difference between the contrasts or luminances does not.) Second, there are "ears" in the curves since there is maximal segregation for the one-element-type-only patterns with somewhat less segrega-

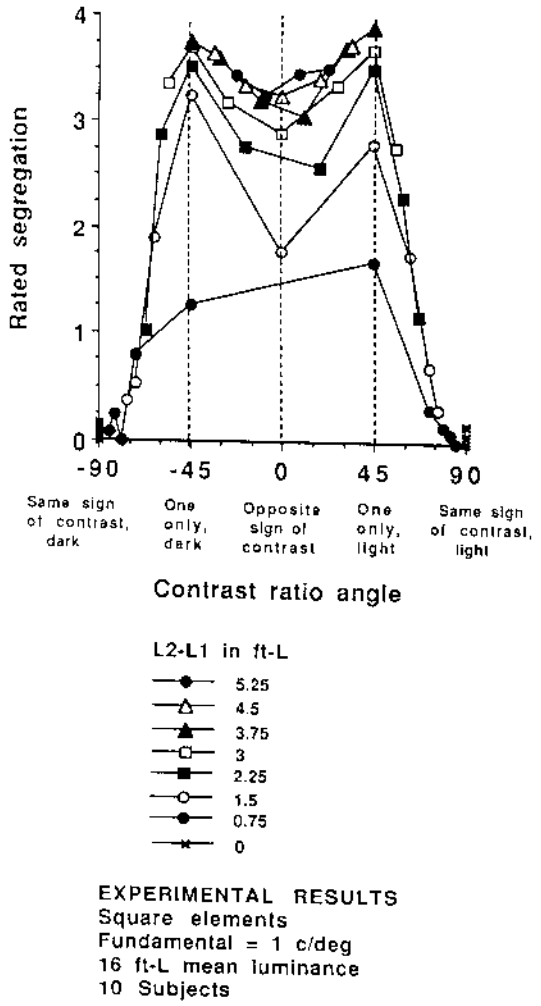


Fig. 18.7
Results from on-off experiment. The average rating from 10 observers is plotted vertically. The contrast-ratio angle is plotted horizontally. Each curve connects points representing patterns in a constant-difference series.

tion for the opposite-sign-of-contrast patterns in the middle of the curves (as well as much less segregation for the same-sign-of-contrast patterns at the ends of the curves).

Role of the Complex Channels in On-Off Experiments

Our explanation of these results rests on two different nonlinearities. The first we have already discussed. It is the *spatial nonlinearity* in the complex channels that was needed to explain the discrepancies between the simple model's predictions and the contrast-size trade-off experiments' results. These complex channels can also account for part of the discrepancies between the simple model's predictions and the on-off experiments' results, namely, for the fact that one-element-only patterns segregate better than opposite-sign-of-contrast patterns. The remainder of this paragraph attempts to describe (in the spatial rather than the frequency domain) the intuition behind this prediction. Consider a complex channel in which the first-stage filter is at relatively high spatial-frequencies (as in the diagram of figure 18.3) and the third-stage filter is at the fundamental frequency of the pattern. (To have a name for this type of channel in the following, we will call it a *high-low complex channel*.) No matter what the pattern, the high-frequency first-stage filter of such a high-low complex channel responds well at the edges of all elements, giving little Mach-bandlike responses, both positive and negative (cf. the right panel of figure 18.2). The second-stage rectification-type nonlinearity turns these positive and negative Mach-bandlike responses at the edges of every element into entirely positive responses. What happens then at the third-stage filtering, done at the fundamental frequency of the pattern? Think of a vertical receptive field superimposed on the output from the second-stage nonlinearity in the striped region. It is tuned to the fundamental frequency of the striped region and so can be placed with its center lying on a column of rectified responses to bright elements and its surround either on empty columns (in the case of *one-element-type-only* patterns) or on columns of rectified responses to dark elements (in the case of *opposite-sign-of-contrast* patterns). Thus this receptive field (and hence the overall complex channel) will respond strongly to the one-element-only patterns (since only its excitatory center is stimulated—it is getting no inhibition from its surround). But it will not respond at all to the opposite-sign-of-contrast patterns since both its center and surround are stimulated and stimulated equally (since the second-stage's rectified responses to both the bright and the dark elements are the same)!

In fact, one can write the following equation to (approximately) describe R_C, the pooled response of the high-low complex channels (the complex channels in which the first filtering is at higher harmonics of the pattern and the second filtering is at the fundamental frequency):

$$R_C = w_C \cdot \{|\Delta L_1| - |\Delta L_2|\}. \quad (5)$$

This is just like equation 4 for the simple channels (dominated by the simple channels at the fundamental frequency) except that $|\Delta L_i|$ has replaced ΔL_i. Why is this equation (approximately) correct? The first-stage filter of a high-low complex channel responds in proportion to the edge contrast ΔL_i, which the second-stage nonlinearity then rectifies to $|\Delta L_i|$ before sending it on as an input to the third-stage filtering done at the fundamental. Then, by an argument analogous to that given earlier for the simple channel at the fundamental frequency, the third-stage filtering in these high-low complex channels will respond in proportion to the difference between its inputs, that is to the difference between $|\Delta L_1|$ and $|\Delta L_2|$. It might be useful to consider what a plot of R_C in the format of figure 18.7 would look like. Each constant-difference curve would be flat for all same-sign-of-contrast patterns (for contrast-ratio angles greater than $+45$ or less than -45). But as the contrast-ratio angle moved in toward zero from either $+45$ or -45, the value of R_C drops steadily reaching zero at a contrast-ratio-angle of zero (zero segregation for opposite-but-equal patterns).

If the only channels that existed were these high-low complex channels (higher frequency at the first stage and fundamental frequency at the third stage), the opposite-sign-of-contrast patterns would not segregate at all. That they do segregate to some extent in figure 18.7, although not as well as the one-element-only patterns, can be explained by assuming that, in addition to high-low complex channels, there are either simple channels at the fundamental frequency or complex channels whose first filtering is at the fundamental frequency. (Such complex channels will act much like the simple channels at the fundamental, both kinds obeying equation 4.) These channels respond as much to the opposite-sign-of-contrast as to the one-element-only patterns and thus will contribute to the segregation of both.

Patterns Lacking Energy at the Fundamental

If one removes all the energy at the fundamental frequency from the pattern, however, neither the simple channels at the fundamental nor the complex channels having a first-stage filtering at the fundamental are stimulated. The high-low complex channels can segregate some patterns lacking energy at the fundamental (e.g., one-element-only patterns), but these high-low channels cannot segregate opposite-sign-of-contrast patterns at all. We used some patterns lacking energy at the fundamental; they were much like those in figures 18.1 and 18.4 but the individual elements had balanced sub-areas of positive and negative contrast and thus averaged out to the same luminance as the background. Indeed, the opposite-sign-of-contrast patterns made from such elements did not segregate at all although the one-element-only patterns segregated very well indeed; this result provides more and quite direct support for the notion of high-low complex channels (Graham, Beck, and Sutter, in preparation).

The Intensity-Dependent Nonlinearity: Two Candidate Processes

The second nonlinearity used to explain the on-off results (figure 18.7) will be called the intensity-dependent nonlinearity. This nonlinearity is needed to explain why the segregation decreases so sharply at the ends of the curves in figure 18.7. We will discuss two possible candidates for this second nonlinearity: (1) a local (pointwise) nonlinearity occurring early (before the channels), and (2) a normalization process, which might result from intracortical interaction, operating at the level of the channels themselves.

An Early Local Nonlinearity: One Possibility

Suppose, for example, that early local light adaptation processes readjust the operating range of the visual system to be centered on the recent mean luminance—the background luminance in this case—maximizing discriminability between luminances near that level, and, therefore, sacrificing discriminability for luminances far away. Figure 18.8 shows a hypothetical early local nonlinearity of this sort. ("Early" is meant only to imply that it occurs before the channels, that is, at the retina or lateral geniculate nucleur [LGN], presuming the channels are cortical; "local" is mean to imply that it is a process that is quite localized compared to the filtering done by the channels and can thus be assumed to act at each point on the stimulus.) Indicated on the horizontal axis of figure 18.8 are five pairs of luminances corresponding to five different patterns in a constant-difference series. Vertical lines

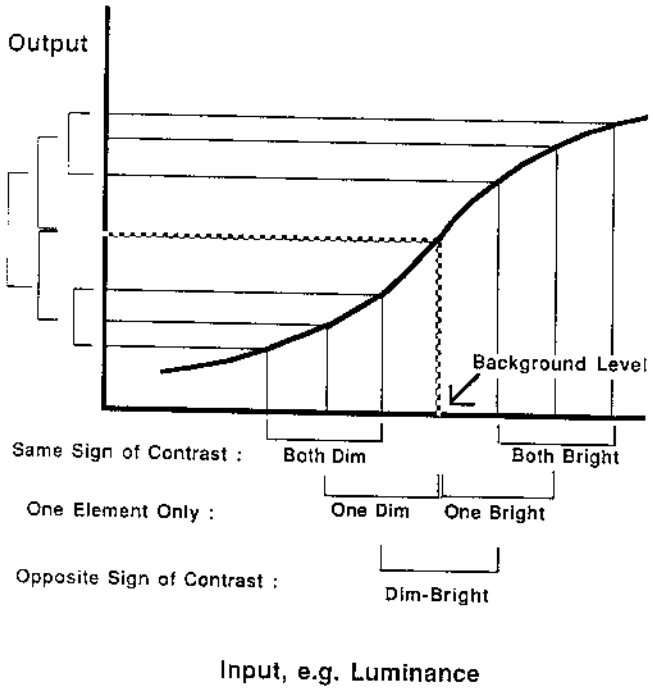


Fig. 18.8
Diagram of early-local-nonlinearity explanation of on-off experiment's results.

extend from these luminances up to the nonlinear curve and then horizontal lines extend over to the vertical axis to show the outputs for each pair of luminances. Note that the difference between the two outputs for the same-sign-of-contrast patterns is much smaller than that for the one-element-type-only patterns or for the opposite-sign-of-contrast patterns. In general, as the luminances get further from the background in either direction, this output difference gets smaller. (It is interesting to note that once you get fairly far from the background, the early local nonlinearity would need to be a logarithmic function of $|\Delta L|$ to make the curves for constant-difference series converge at their ends, that is, to make the amount of segregation depend only on contrast ratios and not on luminance differences, as is the tendency in the empirical results of figure 18.7.)

To calculate approximately the predictions from such an early local nonlinearity applied before the channels, one can start with equation 4 for the channels at the fundamental and equation 5 for the high-low complex channels but then substitute the outputs of the early local nonlinearity for the luminances in those equations and

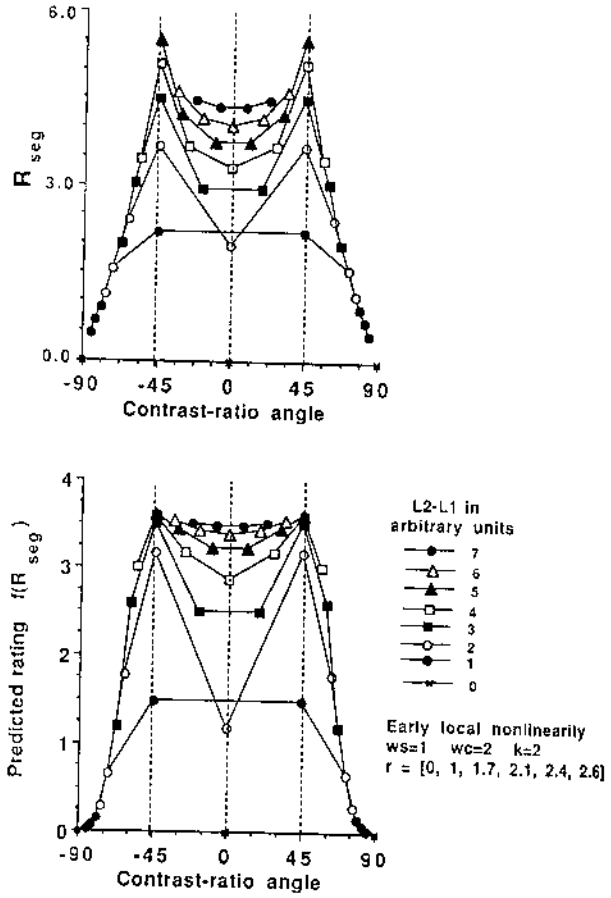


Fig. 18.9
Predictions of on-off experiment from a model incorporating complex channels and an early local nonlinearity.

combine the two pooled responses. In symbols, let R_{seg} be the pooled response over all the simple and high-low complex channels that contribute to segregation of the checkerboard versus the striped region; let k be the parameter that describes the manner in which outputs of the different channels are pooled as in equation 3. Then

$$R_{seg} = \{R_S^k + R_C^k\}^{1/k}, \tag{6}$$

where outputs from the early local nonlinearity have been substituted for luminances in equations 4 and 5. The texture-segregation ratings given by the observer are assumed to be a monotonic function of R_{seg}.

Figure 18.9 shows some predictions from equation 6 where the top panel shows the R_{seg} values themselves and the bottom panel shows R_{seg} transformed by a monotonic transformation f to produce a better fit of R_{seg} to the data

of figure 18.7. (Remember that the observer's ratings are only assumed to be a monotonic function of R_{seg}.) As intended, there are ears in the prediction (due to the presence of high-low complex channels), and the ends of the curves drop and different curves converge. Indeed, as comparison of figures 18.7 and 18.9 will show, these predictions fit the data very well. (The r^2 between $f(R_{seg})$ and the average observer rating across all 66 stimuli was 0.9873.) The parameter values are given in the figure inset but should not be overinterpreted as there are strong interactions among parameters and a wide range of any one parameter can work well given the correct value of other parameters.[1]

The good fit between the predictions of figure 18.9 and the data of figure 18.7 is pleasing on the one hand. On the other hand, one needs to think carefully about just what visual process this early local nonlinearity might correspond to. The heavy solid curve in figure 18.10, which shows the early local nonlinearity that was used for figure 18.9 (the dotted curves show alternatives that were clearly less good), is already strongly compressed at 18 or 14 ft-L, which—on a background of 16 ft-L—is a contrast of 13 percent. This seems too much compression for the light adaptation processes generally talked about in the retina (see reviews in Hood & Finkelstein, 1986; Shapley & Enroth-Cugell, 1985; Walraven, Enroth-Cugell, Hood, et al., 1989). It is also rather more compression than commonly reported for retinal-ganglion and LGN cells, although not extraordinarily more than that reported for some retinal and LGN M cells (e.g., Derrington & Lennie 1984; Sclar, Maunsell & Lennie, 1990; Shapley & Perry, 1986; Spekreijse, van Norren & van den Berg, 1971).

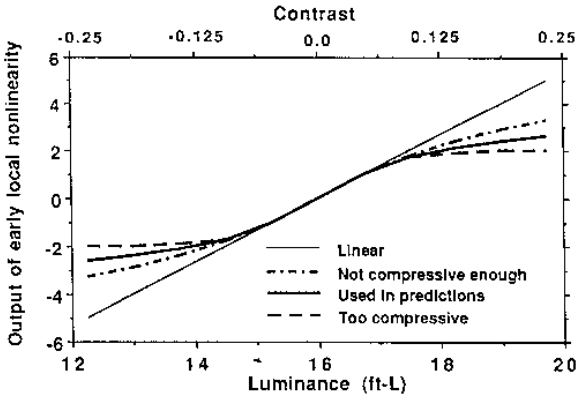


Fig. 18.10
Early-local nonlinearity used in predictions of figure 18.9

Normalization Across Channels—Another Possibility

Perhaps a better candidate for the intensity-dependent nonlinearity displayed in our texture segregation results lies in some known properties of cortical cells. Experiments on cortical cells produce results often described as involving cross-orientation or cross-frequency inhibition (e.g., Bonds, 1989; De Valois & Tootell, 1983; Morrone, Burr & Maffei, 1982). Further, the relationship between contrast and cortical cells' responses is known to be very compressive; some cortical cells do show compression at 10 or 20 percent contrast (e.g., Albrecht & Hamilton, 1982; Ohzawa, Sclar & Freeman, 1982; Sclar, Lennie & DePriest, 1989; Sclar, Maunsell & Lennie, 1990). As has been recently pointed out (see chapter 9; Heeger & Adelson, 1989; Robson, 1988a,b), both the intracortical inhibition and the response compression may result from

1. The early-local-nonlinearity model of equation has the following parameters: w_S, w_C, k, and the values of the early-local nonlinearity at the five different incremental luminances used in the experiment. (The nonlinearity was assumed to be odd-symmetric around the background luminance.) The normalization model of equation 8 and 9 has the following parameters: w_S, w_C, w_0, σ, k', and k.

For each of the models, a crude grid search was done over the above parameters, varying each parameter by a factor of about 2 over a reasonable range. (For the normalization model, the value of k' was arbitrarily set at 1 in the calculations done so far.) For each chosen set of values of these parameters, the values of R_{seg} for the 66 stimuli in the experiment were calculated.

The observer's rating of perceived segregation is assumed to equal $f(R_{seg})$ where f is a monotonic function otherwise unspecified. In the fits shown here, however, we assumed that f was a member of a particular four-parameter family of functions. (This family is a slight generalization of the Weibull distribution function, the Quick psychometric function, and the asymptotic regression function. It was picked merely because it contained the right variety of shapes to fit for the task at hand.)

$$f(x) = a[b + 1 - 2^{-(c \cdot x)^d}] \qquad (7)$$

The Nelder-Meade algorithm (which is nicely described in *Numerical Recipes* by Press, Flannery, Teykolsky, and Vetterling, 1986) as instantiated in MATLAB (available from The MathWorks Inc., 21 Eliot St., South Natick, MA 01760) was used to find the four parameters for f producing the smallest mean-square error (over the 66 stimuli in the experiment) between $f(R_{seg})$ and the average rating of the observers.

The fits shown in figures 18.9 and 18.11 are the best ones found using this procedure, but many others were just as good, since there are strong interactions between these parameters. Trivially, making all the weights w larger by some factor or making the outputs of the early-local nonlinearity larger by some factor will be exactly compensated by making a smaller by that factor. Of particular note for the normalization model, it is generally only the ratio of w_0 to σ that matters rather than the values of either one. The predictions are not, overall, very sensitive to the value of k for either model.

the same process, a normalization process which keeps the total response from some set of neurons at or below a ceiling. It accomplishes this by doing something like dividing (normalizing) each individual neuron's response by the total response from the set (once thresholds have been exceeded) as is discussed in some detail in chapter 9. Normalization-type processes have also been used in visual models by, among others, Grossberg (1987), Grossberg and Mingolla (1985), Lubin, (1989), and Lubin and Nachmias (1990).

How could normalization act to produce the aspects of our results on texture segregation that suggest intensity-dependent nonlinearity? One first needs to assume that the set over which normalization occurs consists of neurons responsive to rather a broad range of spatial frequencies in the same region. The general idea is then that the greater amount of higher harmonics present in the same-sign-of-contrast patterns (relative to both one-element-type-only and opposite-sign-of-contrast patterns) produces larger responses to the same-sign-of-contrast patterns (than to other patterns) from certain channels, henceforth called *other* channels. These *other* channels can be either simple channels at higher harmonics or complex channels having their first-stage and third-stage filterings at higher harmonics. Notice these *other* channels are not capable of segregating the texture regions because they do not respond differentially overall to the two texture regions. The larger responses from the *other* channels to the same-sign-of-contrast patterns (than to one-element-type-only or opposite-sign-of-contrast patterns) would enter into the denominator of the normalization process, however. This larger denominator then leads to a smaller postnormalization response to the same-sign-of-contrast patterns (than to the other patterns) from all channels over which normalization occurs, in particular, from the channels that do lead to segregation of the texture regions (the simple channels at the fundamental and high-low complex channels).

The simplistic approximating approach of equations 4 to 6 can be extended to model normalization instead of early local nonlinearity. To do this we need an expression for R_0, the responses of the *other* channels. Since these channels respond to the higher harmonics (and thus roughly in proportion to the edge contrast of the individual elements), the total response of all such channels in any not-too-small region of the pattern can be approximated by the following expression (and the approximation is very good; Graham, Beck, and Sutter, in preparation):

$$R_0 = w_0 \cdot \{ | |\Delta L_1|^{k'} + |\Delta L_2| |^{k'} \}^{1/k'}. \tag{8}$$

In words, R_0 is proportional to a power-sum of the absolute values of the incremental luminances at the edges of the two types of individual elements, where the constant of proportionality is called w_0 and the exponent k' is from the spatial-pooling asssumption of equation 1. It might be useful to think through what a plot of R_0 vs. contrast-ratio angle looks like for a constant-difference series of patterns: It is flat for angles between -45 and $+45$ and then increases precipitously as the angle gets smaller than -45 or larger than $+45$.

To finish the prediction of perceived segregation one puts the response of these *other* channels into a denominator that normalizes the responses of all the channels but does not put it into the numerator (representing the fact that the *other* channels do not respond differentially in the two texture regions):

$$R_{seg} = \frac{\{R_S^{k'} + R_C^{k'}\}^{1/k'}}{\{\sigma + R_S^{k'} + R_C^{k'} + R_0^{k'}\}^{1/k'}}. \tag{9}$$

The parameter σ is necessary to keep the equation from blowing up at zero. When σ is so large that it dominates the rest of the denominator, this equation is equivalent to equation 6 above (except that the weights have been divided by σ).

As an aside, notice that one might just take equation 9 to be a formulation of some kind of "masking" of the responses at the fundamental frequency by responses to the higher harmonics. It is not clear that this last sentence has much content but it does suggest that one might profitably search for analogies between these psychophysical results and others commonly called "masking."

Figure 18.11 shows some predictions from equation 9 where the top panel shows the R_{seg} values themselves and the bottom panel shows R_{seg} transformed by a monotonic transformation to produce a better fit of R_{seg} to the data of figure 18.7. These predictions fit the data of figure 18.7 very well, almost exactly as well as those of figure 18.9 from the early local nonlinearity. (The r^2 between $f(R_{seg})$ and the average observer rating across all 66 stimuli was 0.9874.) Again the parameter values are given in the figure inset but should not be overinterpreted. See note 1 for further details.

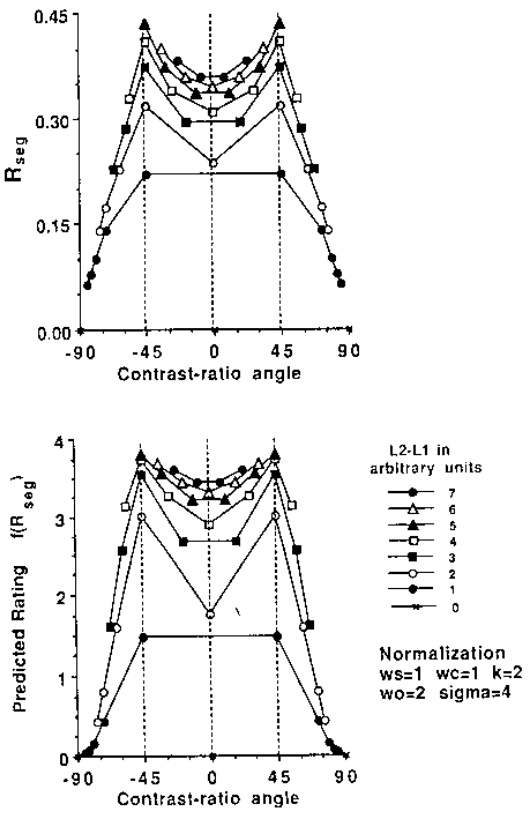


Fig. 18.11
Predictions for on-off experiment from a model incoporating complex channels and normalization (perhaps a result of intracortical inhibition).

To decide between early local nonlinearity and normalization on the basis of their predictions' fits to this data would be foolish. To decide between them totally on the basis of properties of their presumed physiological substrates (that the early local nonlinearity may not be plausible for the retina or the LGN) is at least premature. However, the processes differ in so many other properties (in particular, their spatial properties), not tested by this data, that further psychophysical experiments should be able to decide between them. The relationship of either of these processes in conjunction with complex channels to other nonlinear relationships that have been or are being suggested by others (e.g., the opponency in Bergen and Landy's model in this volume, the spatial inhibition in Malik and Perona 1989a,b, the thresholding in Victor and Conte 1989b) needs also to be made clearer conceptually and/or experimentally. Certain other experimental results (e.g., Victor & Conte 1987, 1989a,b) suggest that even more complicated nonlinearities may be necessary to explain texture discrimination and perhaps also texture segregation.

Perceived Lightness

Whatever the process underlying the intensity-dependent saturating nonlinearity in the region segregation judgments, the same kind of saturation may not appear in judgments of the perceived lightness of individual elements in the pattern. For the relationship between region segregation judgments and perceived lightness judgments is not simple for element-arrangement texture patterns such as those in figure 18.1 and 18.4 (Beck, Graham & Sutter, 1991; Beck, Sutter & Ivry, 1987).

Summary

Although much of what determines whether different regions in the visual field segregate immediately is accounted for by linear spatial-frequency channels models, there is clear nonlinear behavior. The nonlinear behavior exhibited in our experiments can probably all be accounted for by two different nonlinearities of types we know to exist at relatively low levels in the visual pathways: (1) A rectification-type (spatial) nonlinearity quite like that used to describe complex cortical cell behavior; (2) a very dramatic compressive nonlinearity, occurring at contrasts far less than 25 percent, which can be quantitatively predicted either by an early local nonlinearity occurring before the channels or by normalization among the channels (perhaps intracortical inhibition). In any case, ignoring this dramatic compressive nonlinearity in future attempts to explain region (texture) segregation would seem unwise.

Thus, as we and others are currently showing, models involving well-known low-level visual processes can explain a great deal about perceived region (texture) segregation and at a quantitative level. This is not to deny that still higher-level processes—perhaps more like that grouping processes of the Gestalt psychologists and even perhaps acting at a more categorical level—play a role in region segregation but only to suggest that such processes should not be invoked until they are needed. (For further discussion down this line, see Bergen, 1991.) Perhaps, after all, such higher-level processes may *not*

play a substantial role in region segregation if region segregation is a quick and easy computation done early in visual processing in order to ease the overload on higher processes by delimiting regions beyond which a given computation need not be done.

References

Adelson, E. H. & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A, 2,* 284–299.

Albrecht, D. G. & Hamilton, D. B. (1982). Striate cortex of monkey and cat: Contrast response function. *Journal of Neurophysiology, 48,* 217–237.

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

Beck, J., Graham, N. & Sutter, A. (1991). Perceived lightness and perceived segregation. *Perception and Psychophysics 49,* 257–269.

Beck, J., Sutter, A. & Ivry, R. (1987). Spatial frequency channels and perceptual grouping in texture segregation. *Computer Vision, Graphics, and Image Processing, 37,* 299–325.

Bergen, J. R. (1988). Visual texture segmentation based on oriented energy pyramids. Presented at Optical Society of America Annual Meeting, 30 Oct.-4 Nov. 1988. *Technical Digest Series, 11,* 162.

Bergen, J. R. (1991). Theories of visual texture perception. In D. Regan, (Ed.), *Vision and Visual Dysfunction, Vol. 10B: Spatial Vision* (pp. 114–134). New York: Macmillan.

Bergen, J. R. & Adelson, E. H. (1986). Visual texture segmentation based on energy measures. *Journal of the Optical Society of America A, 3,* 98.

Bergen, J. R. & Adelson, E. H. (1988). Early vision and texture perception. *Nature, 333,* 363–364.

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.

Bovik, A. C., Clark, M. & Geisler, W. S. (1987). Computational texture analysis using localized spatial filtering. *Proceedings of Workshop on Computer Vision.,* Miami Beach, FL, Nov. 30–Dec. 2, 1987 (pp. 201–206). The IEEE Computer Society Press.

Caelli, T. M. (1982). On discriminating visual textures and images. *Perception and Psychophysics, 31,* 149–159.

Caelli, T. M. (1985). Three processing characteristics of visual texture segregation. *Spatial Vision, 1,* 19–30.

Caelli, T. M. (1986). Three processing characteristics of texture discrimination. *Vision Interface, '86,* Proceedings, 343–348.

Caelli, T. M. (1988). An adaptive computational model for texture segmentation. *IEEE Transactions on Systems, Man, and Cybernetics, 18,* 1

Chubb, C. & Sperling, G. (1988). Processing stages in non-Fourier motion perception. *Supplement to Investigative Ophthalmology and Visual Science, 29,* 266.

Clark, M., Bovik, A. C. & Geisler, W. S. (1987). Texture segmentation using a class of narrowband filters. *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing.*

Coggins, J. M. & Jain, A. K. (1985) A spatial filtering approach to texture analysis. *Pattern Recognition Letters 3,* 195–203.

Daugman, J. G. (1987). Image analysis and compact coding by oriented 2D Gabor primitives. *SPIE Proceedings, 758,* 19–30.

Daugman, J. G. (1988). Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 36,* 1169–1179.

Derrington, A. M. & Lennie, P. (1984) Spatial and temporal constant sensitivities of neurones in lateral geniculate nucleus of macaque. *Journal of Physiology, 357,* 219–240.

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

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A, 4,* 2379–2394.

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

Graham, N. (1985). Detection and identification of near-threshold visual patterns. *Journal of the Optical Society of America A. 2,* 1468–1482.

Graham, N. (1989a). Low-level visual processes and texture segregation. *Physica Scripta. 39,* 147–152.

Graham, N. (1989b). *Visual pattern analyzers.* New York: Oxford University Press.

Graham N., Beck, J. & Sutter, A. K. (1989). Two nonlinearities in texture segregation. *Supplement to Investigative Ophthalmology and Visual Science, 30,* 161.

Graham, N., Beck, J. & Sutter, A. K. (in preparation). Effects of sign and amount of contrast on perceived segregation of element-arrangement textures: Evidence for two nonlinear processes.

Grossberg, S. (1987). Cortical dynamics of three-dimensional form, color, and brightness perception: Monocular theory. *Perception and Psychophysics, 41,* 87–116.

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

Heeger, D. J. & Adelson, E. H. (1989). Mechanisms for extracting local orientation. *Supplement to Investigative Ophthalmology and Visual Science, 30,* 110.

Hochstein, S. & Spitzer, H. (1985). One, few, infinity: Linear and nonlinear processing in the visual cortex. In D. Rose and V. G. Dobson (Eds.), *Models of the Visual Cortex* (pp. 341–350). New York: Wiley.

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

Klein, S. A. & Tyler, C. W. (1986). Phase discrimination of compound gratings: Generalized autocorrelation analysis. *Journal of the Optical Society of America A, 3*, 868–879.

Landy, M. S. & Bergen, J. R. (1988). Texture segregation by multiresolution energy or by structure gradient? Presented at the Annual Meeting of the Optical Society of America, 30 Oct.–4 Nov. *Technical Digest Series, 11*, 162.

Landy, M. S. & Bergen, J. R. (1989). Texture segregation for filtered noise patterns. *Supplement to Investigative Ophthalmology and Visual Science, 30*, 160.

Lubin, J. (1989). Discrimination contours in an opponent motion stimulus space. *Supplement to Investigative Ophthalmology and Visual Science, 30*, 426.

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

Malik, J. & Perona, P. (1989a). *A computational model of texture perception*. (Report No. UCB/CSD 89/491). Computer Science Division (EECS), University of California, Berkeley, California.

Malik, J. & Perona, P. (1989b). Computational model of human texture perception. *Supplement to Investigative Ophthalmology and Visual Science, 30*, 161.

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 B, 216*, 335–354.

Nothdurft, H. C. (1985a). Orientation sensitivity and texture segmentation in patterns with different line orientation. *Vision Research, 25*, 551–560.

Nothdurft, H. C. (1985b). Sensitivity for structure gradient in texture discrimination tasks. *Vision Research, 25*, 1957–1968.

Ohzawa, I., Sclar, G. & Freeman, R. D. (1982). Contrast gain control in the cat visual cortex. *Nature, 298*, 266–268.

Peli, E. (1987). Seeing the forest for the trees: The role of nonlinearity. *Supplement to Investigative Ophthalmology and Visual Science, 28*, 365.

Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1986). *Numerical recipes*. Cambridge: Cambridge University Press.

Robson, J. G. (1980). Neural images: The physiological basis of spatial vision. In C. S. Harris (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. *Supplement to Investigative Ophthalmology and Visual Science, 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 Sixteenth Symposium of the Center for Visual Science, University of Rochester, Rochester, New York, June, 1988. Abstract p. 5.

Rubenstein, R. S. & Sagi, D. (1989). *Spatial variability as a limiting factor in texture discrimination tasks: Implications for performance asymetries*. (Technical Report CS89-12). The Weizmann Institute of Science, Department of Applied Mathematics and Computer Science, Rehovot, Israel.

Sagi, D. (1989). Hierarchy of spatial filtering in early vision. *Supplement to Investigative Ophthalmology and Visual Science, 30*, 161.

Sclar, G., Lennie, P. & DePriest, D. D. (1989). Contrast adaptation in striate cortex of macaque. *Vision Research, 29*, 747–755.

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. & Enroth-Cugell, C. (1985). Visual adaptation and retinal gain controls. In N. N. Osborne and G. J. Chader (Eds.), *Progress in retinal research, Vol. 3* (pp. 263–343). Oxford: Pergamon Press.

Shapley, R. & Perry, V. H. (1986). Cat and monkey retinal ganglion cells and their visual functional roles. *Trends in Neurosciences*, May, 229–235.

Spekreijse, H., van Norren, D. & van den Berg, T. J. T. P. (1971). Flicker responses in monkey lateral geniculate nucleus and human perception of flicker. *Proceedings of the National Academy of Science, USA, 68*, 2802–2805.

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

Sperling, G. & Chubb, C. (1989). Apparent motion derived from spatial texture. *Supplement to Investigative Ophthalmology and Visual Science, 30*, 161.

Sutter, A. (1987). The interaction of size and contrast in perceived texture segregation: A spatial frequency analysis. Doctoral Dissertation, University of Oregon, Eugene.

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.

Turner, M. R. (1986). Texture discrimination by Gabor functions. *Biological Cybernetics, 55*, 71–82.

Victor, J. D. (1988). Models for preattentive texture discrimination: Fourier analysis and local feature processing in a unified framework. *Spatial Vision, 3,* 263–280.

Victor, J. D. & Conte, M. (1987) Local and long-range interactions in pattern processing. *Supplement to Investigative Ophthalmology and Visual Science, 28,* 362.

Victor, J. D. & Conte, M. (1989a) Cortical interactions in texture processing: Scale and dynamics. *Visual Neuroscience, 2,* 297–313.

Victor, J. D. & Conte, M. (1989b) What kinds of high-order correlation structure are readily visible? *Supplement to Investigative Ophthalmology and Visual Science, 30, 3,* 254.

Walraven, J., Enroth-Cugell, C., Hood, D. C., MacLeod, D. I. A. & Schnapf, J. (1989). Control of visual sensitivity. In L. Spillman and J. Werner (Eds.), *Physiological foundations of perception.* New York: Springer-Verlag.

Watson, A. B. (1983). Detection and recognition of simple spatial forms. In O. J. Braddick and A. C. Sleigh, (Eds.), *Physiological and biological preprocessing of images* (pp. 110–114). New York: Springer-Verlag.

Yellott, J. & Iverson, G. (1990). Triple correlation and texture discrimination. *Supplement to Investigative Ophthalmology and Visual Science, 31,* 561.