Vision Res. Vol. 33, No. 14, pp. 1893–1911, 1993
Printed in Great Britain.

0042-6989/93 $6.00 + 0.00

Spatial-frequency- and Orientation-Selectivity of Simple and Complex Channels in Region Segregation

NORMA GRAHAM,*† ANNE SUTTER,* CHARU VENKATESAN*

Received 24 June 1992; in revised form 10 February 1993

Models incorporating spatial-frequency- and orientation-selective channels explain many texture-segregation results, particularly when known nonlinearities are included. One such nonlinearity is *complex channels*. A complex channel consists of two stages of linear filtering separated by a rectification-type nonlinearity. Here we investigate the spatial-frequency- and orientation-selectivity of simple (linear) channels and of the complex channels' first stage. Observers rated the degree of segregation between two "textures" both composed of elements which were Gabor patches. When the textures differed in *type of element* (e.g. one composed of vertical and the other of horizontal Gabor patches), the segregation results yield bandwidth estimates for *simple channels* of approx. 0.5–1.0 octave on the spatial-frequency dimension and 5–20 deg of rotation on the orientation dimension. When the textures differ in *the arrangement of elements* (e.g. striped vs checkerboard arrangements, both of horizontal and vertical patches), the segregation results yield bandwidth estimates for the *first stage of complex channels*. These estimates, while differing substantially from one observer to another, were always substantially wider than those for simple channels (by at least a factor of two) but narrower than bandwidths of LGN cells (particularly on the orientation dimension where LGN cells show little selectivity at all).

Spatial frequency Orientation Texture segregation Complex channels

INTRODUCTION

While linear channels selective for spatial frequency and orientation have been useful in explaining many aspects of human vision, nonlinear processes are clearly necessary as well. They are necessary even when the space- and time-average luminance of the stimuli are held constant to minimize the effects of nonlinear processes usually thought to control light and dark adaptation. Several investigators have considered nonlinear channels like those diagrammed in Fig. 1, that is, two stages of linear filtering separated by a pointwise rectification or similar nonlinearity (e.g. squaring).

In our studies of region segregation, we have referred to these nonlinear channels as *complex channels* (Sutter, Beck & Graham, 1989; Graham, 1991; Graham, Beck & Sutter, 1992a). We were inspired by John Robson's suggestion (Robson, 1980) that complex cortical cells might perform such perceptual tasks, but one should treat this possible physiological analog cautiously. Other investigators have invoked similar processes in the study of texture and motion perception, often calling them *non-Fourier* or *second-order* processes (e.g. Chubb & Sperling, 1988; Fogel & Sagi, 1989; Grossberg & Mingolla, 1985; Pantle, 1992; Sperling & Chubb, 1989; Sperling, 1989; Turano & Pantle, 1989; Wilson, 1991; Wilson & Kim, 1992; Wilson & Richards, 1992).

Examples of complex channels' responses to particular patterns will be shown later. One way to describe the action of complex channels is to say that they can respond to low-frequency arrangements of high-frequency elements [as in AM (amplitude modulation) radio; see further description in Sperling (1989)]. Subsequent to the channels themselves (which may include simple linear channels as well as complex channels) there must be some calculation on the outputs at different spatial positions in different channels in order to determine the observer's response. We have considered a variety of rules for pooling across spatial positions and across different channels [briefly presented in the Appendix here and discussed further in Sutter *et al.* (1989) and Graham *et al.* (1992a)].

To completely specify a model containing complex channels, one would need to know many other pieces of information. Very little is known about the complex channels acting in the perceived segregation of regions in nominally-stationary patterns. One study does give some information about the mapping between the preferred values of spatial frequency of the two filters, suggesting that the preferred spatial frequency of the second filter of a particular channel (the one after the rectification-type

*Department of Psychology, Columbia University, New York City, NY 10027, U.S.A.
†To whom all correspondence should be addressed.

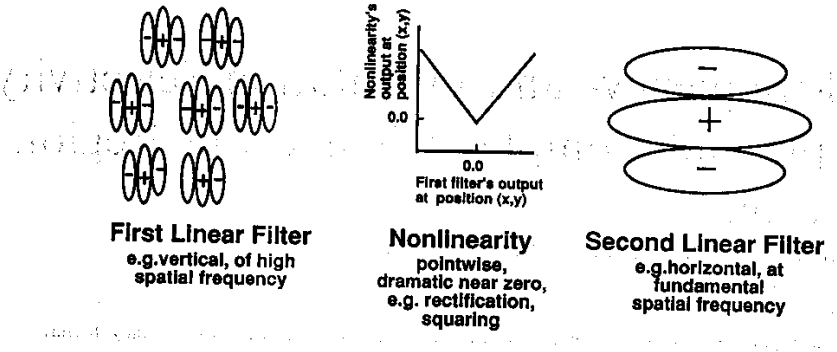

FIGURE 1. Diagram of complex channel with an orientation-selective first filter.

nonlinearity) is 3–4 octaves lower than that of the filter before the rectification-type nonlinearity (Sutter, Sperling & Chubb, 1991). Our goal here is to further specify the first-stage filters of the complex channels by investigating their orientation and spatial-frequency bandwidth.

If, for example, the physiological substrate for the first-stage filters of the complex channels were lateral geniculate nucleus (LGN) cells, the first stage would show little orientation selectivity (since LGN receptive fields are approximately concentric as in Fig. 2). If, however, the physiological substrate for the first stage were simple cortical cells, the first stage should show considerable orientation selectivity (as implied by the diagram in Fig. 1). In a study with moving "texture-quilt" stimuli. Werkhoven, Sperling and Chubb (1992) reported that there was very little orientation selectivity in the second-order processes serving motion-from-texture perception, that is, the correct diagram would be more like Fig. 2 than Fig. 1. Here we will use nominally-stationary patterns.

The patterns used in the first part of this study are like those shown in Fig. 3. In general, they are composed of two types of Gabor-patch elements, arranged in stripes in the left and right regions of the pattern and in a checkerboard arrangement in the middle. (In one panel of Fig. 3, only one type of element is visible because the contrast of the other type has been set to zero.) The Gabor-patch elements used in the study generally had more cycles than those shown (twice as many—except when spatial frequency was varied), but we reduced the number of cycles for this figure to insure visibility after reduction. In the experiment, the observers were asked to rate (on a scale from 0 to 4) how well the patterns perceptually segregated into three regions. They were asked to give their global impressions, not to attempt to scrutinize the patterns.

Responses of simple channels to element-arrangement patterns

Since the space-average luminance of the Gabor-patch elements was the same as the background luminance, there was very little energy in these patterns at the fundamental frequencies of the striped or of the checkerboard region. (A fundamental frequency is a reciprocal of a period of repetition.) The energy at the higher spatial frequencies, resulting from the individual elements, is essentially identical in both striped and checkerboard regions. Hence, as the next paragraph discusses in more detail, the regions of these patterns cannot be segregated by models containing only simple (linear) channels.

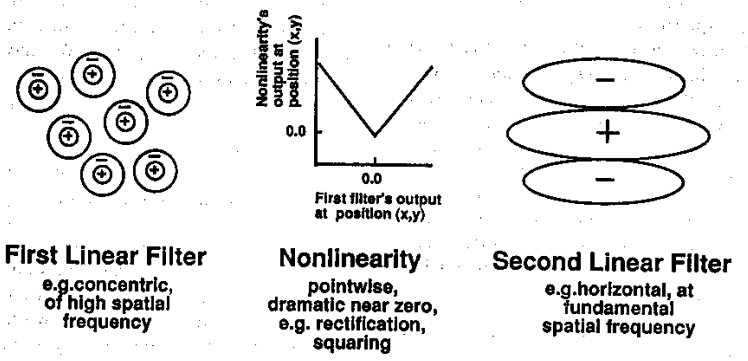

FIGURE 2. Diagram of complex channel with a non-orientation-selective first filter.

A complete description of simple-channel model predictions for patterns similar to those used here is given in Graham *et al.* (1992a). In that study, however, the elements were not Gabor patches but were either center-surround elements (where the space-average luminance equaled that of the background) or solid-square elements (where the space-average luminance differed from that of the background). The two element types in any one pattern had identical spatial configuration, and their contrasts were systematically varied in amount and sign. The computed outputs of simple-channels models to the center-surround-element patterns (which is primarily determined by the energy at the fundamental frequency in the center-surround-element patterns relative to the energy at that frequency in the solid-square-element patterns) was shown to be much less than necessary to predict the perceived segregation reported by observers for center-surround-element patterns. We have since done analogous experiments and computations using grating-patch elements instead of center-surround elements and obtained analogous results. Since the experiments, computations, and results are so similar to those reported in Graham *et al.* (1992a), we do not report them in detail here.

Responses of complex channels to element-arrangement patterns

Figure 4 illustrates the logic of the experiments reported here. It pictures only a small portion of the striped region of the patterns in Fig. 3 and only those complex channels sensitive to horizontal arrangements of elements (as occurs in the striped regions). Analogous reasoning would apply to the checkerboard region and complex channels sensitive to diagonal arrangements.

The righthand columns in Fig. 4 illustrate the responses of two different complex channels to the pattern which is composed of two element types that are perpendicular to one another (the pattern in Fig. 3 where both elements are visible). The top pair of panels in those columns show part of the striped region from that pattern with superimposed receptive fields. On the left the superimposed receptive fields are *concentric* (i.e. non-orientation-selective) and on the right they are *vertically-oriented* (i.e. orientation-selective).

The second row of the righthand columns in Fig. 4 shows the outputs of the first stages. Mid-gray represents

(a)

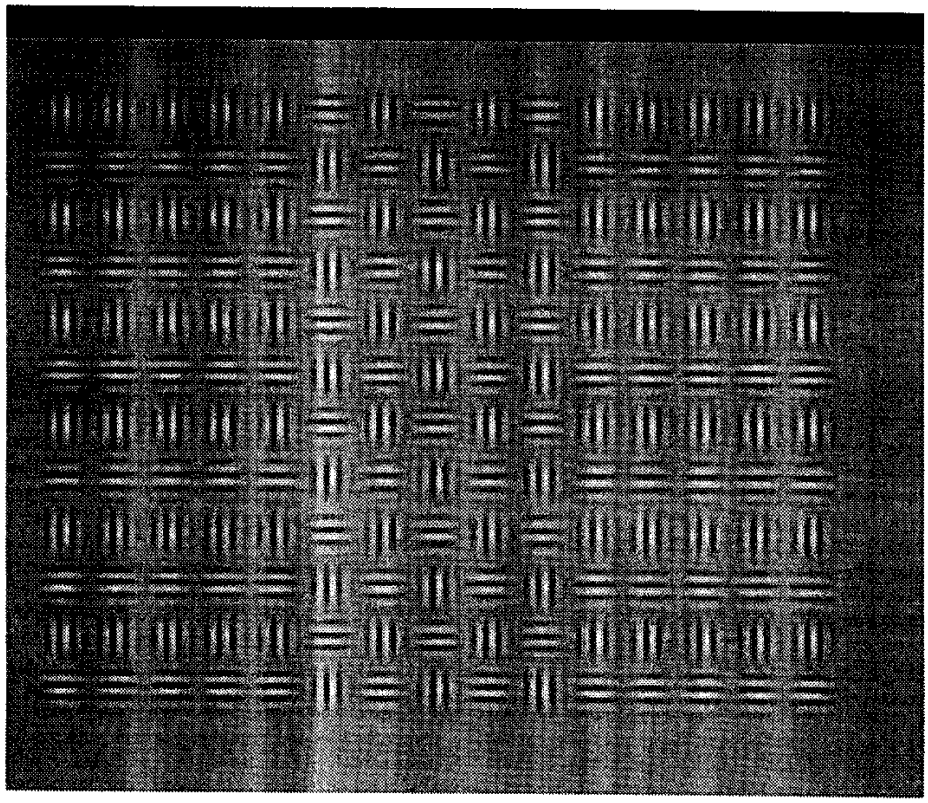

FIGURE 3(a). *Caption overleaf.*

(b)

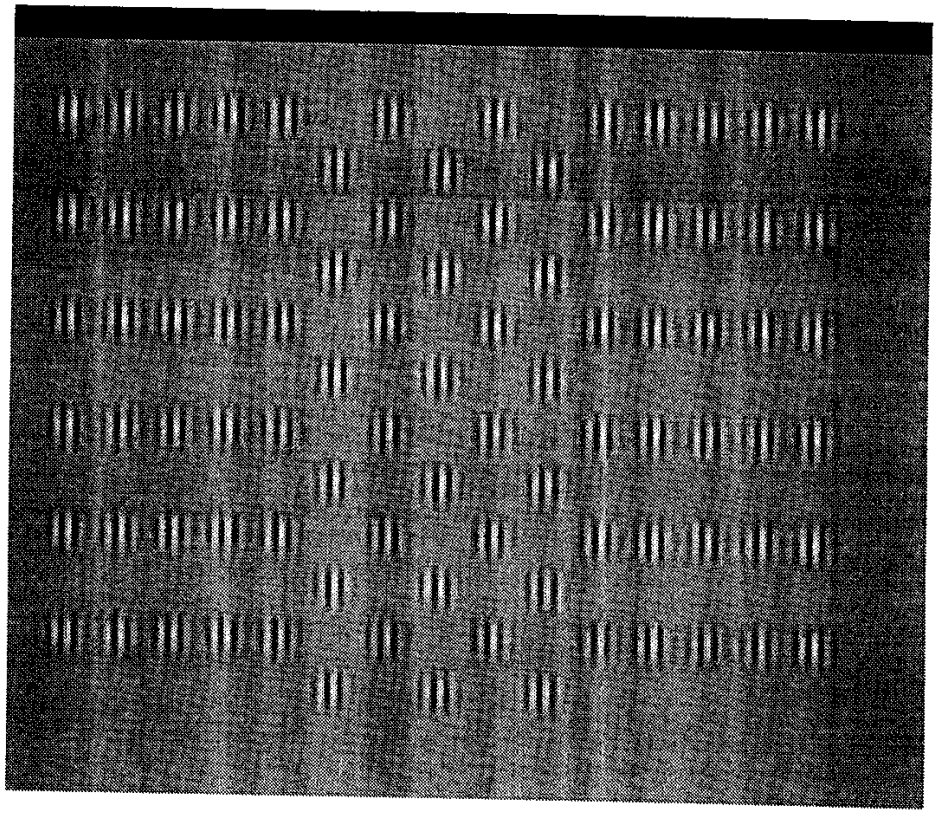

FIGURE 3. Two examples of element-arrangement patterns where the elements are Gabor patches. In (a) the two types of elements are of the same contrast and differ in orientation by 90 deg of rotation. In (b) only one type of element (a Gabor patch of vertical orientation) is visible (i.e. elements of the other type have zero contrast). In the stimuli used in the experiments, the luminance averaged across each grating was the same as the background luminance. The Gabor patches generally had twice as many cycles as those shown here. Reproduction will have distorted the stimuli in these figures somewhat. The black bar visible above the area of mean luminance was not visible on the screen.

a zero response, brighter areas represent positive responses, and darker areas represent negative responses. Note that the non-orientation-selective first-stage filter (left member of the pair) responds to both element types, while the orientation-selective first-stage filter responds only to the appropriate orientation.

These outputs from this first stage of filtering go into the rectification-type nonlinearity and come out looking like those in the third row of Fig. 4. Half-wave rectification was assumed here, so there are still positive responses (bright areas) but the negative responses have been rectified to zero (mid-gray).

Superimposed on the outputs in the third row is a sample receptive field from the second filter of each complex channel. These second filters are identical for the two complex channels under consideration because the second filters respond to the arrangement of elements into horizontal stripes. Notice that this receptive field is large enough to average over the responses to several elements. These second filters' responses are shown in the bottom right panels of Fig. 4. Notice that, if the first-stage is *not* orientation-selective (left member of the pair), the complex channel ends up not being able to detect the horizontal stripes in this pattern because its first stage cannot distinguish the two types of elements. If, however, the first-stage is orientation-selective (right panel) the complex channel does detect the stripes.

In the checkerboard region, the outputs of either of these two complex channels will be almost uniformly zero (because both types of elements fall within the excitatory center and also both fall within the inhibitory surrounds of their second-filter receptive fields).

Consequently, the complex channel with *concentric* first-stage receptive fields responds uniformly (and identically) both in the checkerboard and in the striped regions. Thus it cannot lead to perceptual segregation between the regions.

On the other hand, the complex channel with *an orientation-selective* first stage responds in a strongly modulated fashion to the striped arrangement in the striped region but responds uniformly in the checkerboard region. Hence this complex channel should lead an observer to perceptually segregate these patterns. [Further description of the relationship between the channels' outputs and the observer's responses is given briefly in the Appendix here and discussed more completely in Sutter *et al.* (1989) and Graham *et al.* (1992a).]

More generally, patterns composed of two element types should not segregate well if first-stage receptive fields respond well to both element types but should segregate well if first-stage receptive fields are selective enough to respond well to only one element type.

A one-element-only pattern, on the other hand, should segregate whether or not the complex channel is orientation-selective, as illustrated in the leftmost column of Fig. 4. The top panel in that column shows a portion of the one-element-only pattern with both concentric and oriented receptive fields superimposed. Either kind

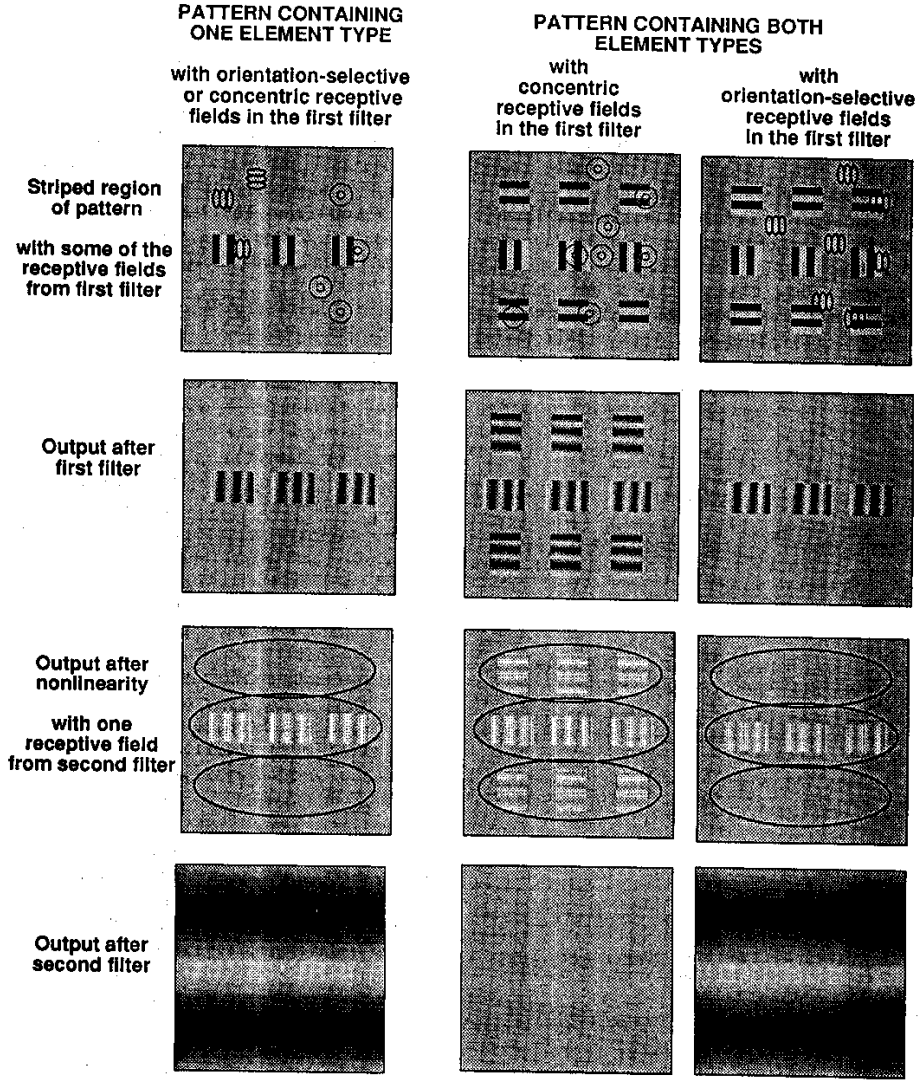

FIGURE 4. Diagram of responses of two complex channels—one with an orientation-selective first stage and one with a non-orientation-selective first stage—to the patterns shown in Fig. 3. The responses of both complex channels are identical to the pattern containing one element type (lefthand column); the responses differ dramatically to the pattern containing both element types (righthand pair of columns).

of first-stage filter responds to the one element type that is present (second row) which is then rectified (third row) and sensed by the second filter (fourth row).

More generally, a one-element-only pattern should segregate whether or not the first stage of the complex is selective for element type. There is only one element type present, after all.

The experiments

In the experiments reported here, we used patterns like those of Fig. 3 with various combinations of orientations and of spatial frequencies. One type of element was a vertical Gabor patch having a spatial frequency of 8 c/deg (each such patch had twice as many cycles as those in Fig. 1), and the second type of element differed from the first in spatial frequency or orientation (but not both). For each pattern, we systematically varied the contrast in the two element types to investigate the degree to which they interfered with one another (or did not) in producing perceptual segregation.

Figure 5 shows hypothetical results. Each curve gives the results from a set of stimuli in which the contrast of element 1 is constant while the contrast in the other element varies widely. (The sketches at the bottom of Fig. 5 illustrate three stimuli from such a set; the first element type is a vertical grating patch and the second element type a horizontal patch.) From such curves, one can tell whether or not the presence of the second element type ever depresses perceived segregation relative to the case where only the first element type is present, that is, one can tell the extent of interference between the two element types. The curve at the top of Fig. 5 illustrates no interference; that at the bottom illustrates a great deal of interference.

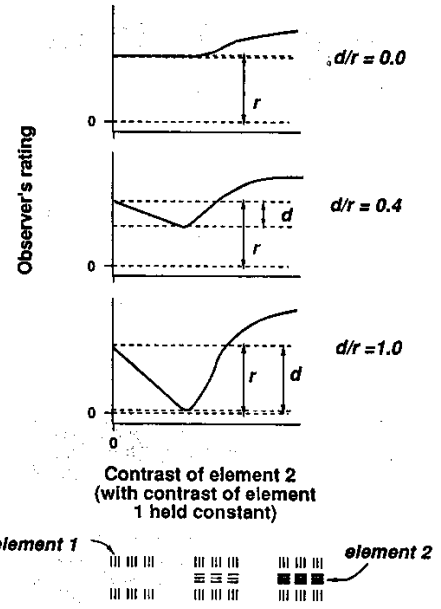

FIGURE 5. Hypothetical results from experiment where contrast in first element type is held constant (e.g. horizontal grating patches as in the sketches at the bottom of the figure) while the contrast in the second element type (e.g. the vertical grating patches in the sketches at the bottom of the figure) is varied.

The d/r ratio as a measure of first-stage sensitivity

The quantity d/r (for dip/range) diagrammed in Fig. 5 is a measure of how much interference two element types cause each other. The precise relationship of the d/r ratio to the sensitivity function of a complex channel's first stage necessarily depends upon many assumptions of the full model. We can derive it analytically for one very simple model (see Appendix). This model contains exactly two complex channels, one having a first-stage filter that is maximally sensitive to the first element type, and the other having a first-stage filter maximally sensitive to the second element type. In both channels, the second filter is maximally sensitive to the overall periodicity of the pattern and the nonlinearity between the two filters is a full-wave rectification. The observer's response is assumed to be equal to the larger of the outputs from the two channels. For this model, the following relationship holds:

$$\frac{d}{r} \approx \frac{S_A(2)}{S_A(1)} \times \frac{S_{obs}(1)}{S_{obs}(2)} \quad (1)$$

where channel A is a channel in which the first stage is more sensitive to element 1 than to element 2, S_A is the sensitivity of channel A, and S_{obs} is the sensitivity of the observer. In words, the d/r ratio is equal to the relative sensitivity of a channel to the two elements weighted by the reciprocal of the observer's relative sensitivity. When the observer happens to be equally sensitive to the two elements—as often happens in the case of different orientations—then the d/r ratio directly gives the relative sensitivity of the complex channels.

This naive, two-channel model can easily be shown to be inadequate. It completely ignores all the other simple and complex channels, which are certainly less important than the two under consideration but not entirely negligible. Further, it does not address the fact that the observer's ratings are constrained to be in a finite range but the output of the channels continues to infinity. At the least, the observer's ratings must be assumed to be a nonlinear function of the pooled outputs from the channels. More interestingly, an intensive nonlinearity like those we have invoked in previous work (Graham, 1991; Graham *et al.*, 1992a) is probably necessary. This might be an early, local nonlinearity occuring before the channels or something more like inhibition among the channels (as captured in a normalization network). More work will be needed to decide on the form of the intensive nonlinearity. However, preliminary calculations suggest that equation (1) will hold rather well for a number of more realistic models in a number of situations (see Appendix). In any case, this ratio seems to stay quite constant over different values of contrast (as long as the contrast is high enough for the elements to be clearly visible). Thus, we are going to use the d/r ratio here to summarize our experimental results.

METHODS AND PROCEDURES

Both element types were Gabor patches with a concentric Gaussian window having a half-width half-height of 8 pixels (0.25 deg at a viewing distance of 0.91 m) truncated at ± 16 pixels so as not to overlap with the neighboring elements. The harmonic oscillation was in sine phase with respect to the window so that the space-average luminance across each element was the same as the background luminance. The center-to-center spacing between neighboring elements was 32 pixels (1 deg); thus the fundamental frequency of the striped region (one period of which consists of two rows of elements and two rows of inter-element spaces) was 0.5 c/deg. The numbers, spacing, and arrangments of the elements can be seen in Fig. 3.

The spatial frequency of the first element type was always 8 c/deg (a period of 4 pixels). (In order to insure their visibility, the period used in Fig. 3 was 8 pixels; thus each patch in Fig. 3 contains half as many cycles as for element 1 in the experiments reported here.)

In all element-arrangement patterns used in this study, the overall orientation of the stripes was horizontal as shown in Fig. 3. The orientation of the first element type was usually vertical as also shown in that figure. (To serve as a control, the orientation of element 1 was horizontal for several experiments described below.)

Values of the base contrast for each orientation or spatial frequency were roughly equated for visibility on the basis of pilot experiments with each observer. For the first element type (8 c/deg and vertical) the base contrast was typically 5%.

In each session, the spatial frequency *and* orientation of element 1 were held constant. Either the spatial frequency or orientation of element 2 was held constant (and identical to element 1) while the other was varied randomly from trial to trial over seven different values (sampling without replacement).

From one presentation of a pattern to the next the contrasts of both elements varied randomly (sampling without replacement). The contrast of element 1 could be one of several equally-spaced levels; generally three different levels were used, and they were equal to 0, 3, and 6 times the base contrast. The contrast of element 2 could be one of a number of equally-spaced levels; generally ten levels were used, and they were equal to $0, 1, 2, \ldots, 9$ times the base contrast. Thus there were generally 30 different contrast combinations for each of the seven different values of element 2's frequency or orientation, making a total of 210 different stimuli. Also, about 10% of the trials were of stimuli not described here but used to check that the observer's behavior remained stable over time.

For each subject there were four replications of each stimulus: two replications in each of two sessions.

To compute the d/r ratios shown below, observer's responses were first averaged across the four replications of each stimulus. (Some resulting curves are shown in Fig. 6.) The d/r ratio was then computed separatedly for each level of contrast in element 1. There were no systematic differences between the d/r ratios computed from different levels of first-element contrast, and thus they were averaged for presentation here.

Each trial started when the subject pressed the appropriate part of a response pad. (The response was a commercially-available pad, an "Unmouse", approx. $4 \times 2''$ in size, which can be pressed at any position to substitute for a mouse press at the corresponding position on the Macintosh display.) A fixation pattern then appeared for 1 sec. It was a cross located in the middle of the screen, at 10% contrast, extending 8 pixels horizontally and 16 pixels vertically. 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 before a beep signaled to the observer that a response would now be accepted.

The observer's response was to indicate the degree of perceived segregation by pressing the appropriate position within a 4" wide and 1" high rectangle on the "Unmouse" response device. Although the "Unmouse" recorded responses on an effectively-continuous scale, five equally-spaced numerals (0, 1, 2, 3, and 4) were written on the face of the "Unmouse", and observers were instructed about the meaning of the numerals (as described in the next subsection). The numerals on the "Unmouse" were visible during the experiment due to illumination from the screen. Observers knew, however, that their responses were registered on a continuous scale. This scale was scored as going from 0 to 100 for the results shown in Fig. 7. After the observer's response, there was a double-beep.

The 1 sec delay after the offset of the stimulus before allowing a response on the "Unmouse" is worth some comment. We used it on the basis of experiments we ran to replicate and extend the results of Graham *et al.* (1992a). These experiments had shown that, when responding using an "Unmouse" but without the delay, observers go extremely quickly from one stimulus to the next; further, their responses differ to some extent from those they give when writing down their ratings [as was done in Graham *et al.* (1992a)]. Three factors may contribute. When responding using an "Unmouse" without the delay, observers may go so fast that (i) they do not consistently follow the instructions to respond on the basis of perceived global segregation and/or (ii) they may become distracted or otherwise affected by the patterned aftereffects that appear with too short an inter-stimulus interval and/or (iii) they may respond as soon as the response of some one channel is big enough and, if the responses of different types of channels rise to asymptote at different speeds, this strategy will produce different results than when waiting for all channels' responses to asymptote. Thus we decided to use a delay regularly in this study, although we intend to pursue this difference further.

The patterns were generated and experiments run by a Macintosh IIci on a standard Apple Monitor with Pascal programs that were built upon programs kindly supplied by Hugh Wilson.

The space-average luminance of the patterns presented on a CRT screen was about 18 ft-L. Observers viewed the CRT binocularly while sitting in a chair in front of the CRT screen with unrestrained head. Light from a partially obscured lamp behind the subject pro vided some general background illumination of the room.

Observers and instructions

One of the observers (CV) is an author. The rest were naive as to the purpose of the experiment. Observers CV, MH, RI, and AK had participated in previous experiments which were replications and extensions of the experiments described in Graham *et al.* (1992a). These prior experiments had used the same general procedure but with patterns in which the elements were squares or center-surround elements rather than Gabor patches. For observers JW and KC, the data reported here was the first data we collected.

Before participating in their first segregation-rating experiments, the observers all received at least 15–30 min of instructions which included going through a series of practice patterns. In addition to being instructed as to the meaning of the response scale (see below), the observers were instructed to maintain a focus of attention that was global and to give their first and overall impressions rather than scrutinizing the pattern. They were asked to rate each pattern according to how distinct the three different regions appeared to them. They were asked to ignore factors such as the overall size and the degree of brightness of the pattern. They were instructed *not* to focus on the individual squares or any other form of local information when determining the rating. The instructions were repeated in several ways for emphasis. They were told to use the following guide to the ratings:

0—no perceptible difference between the regions
1—barely perceptible difference between regions
2—perceptibly distinct regions
3—moderately distinct regions
4—highly distinct regions.

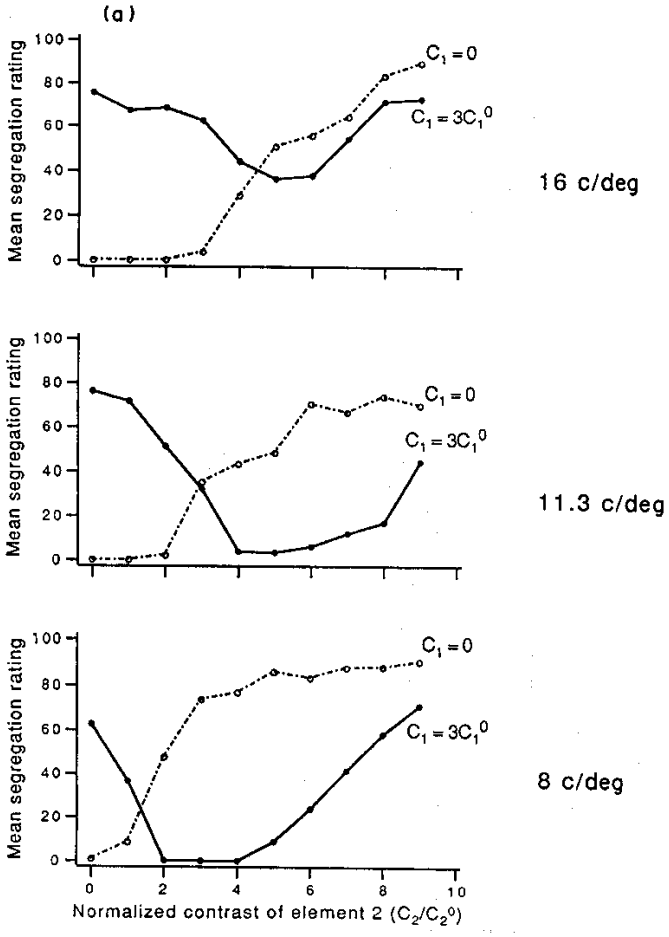

FIGURE 6(a). *Caption on facing page.*

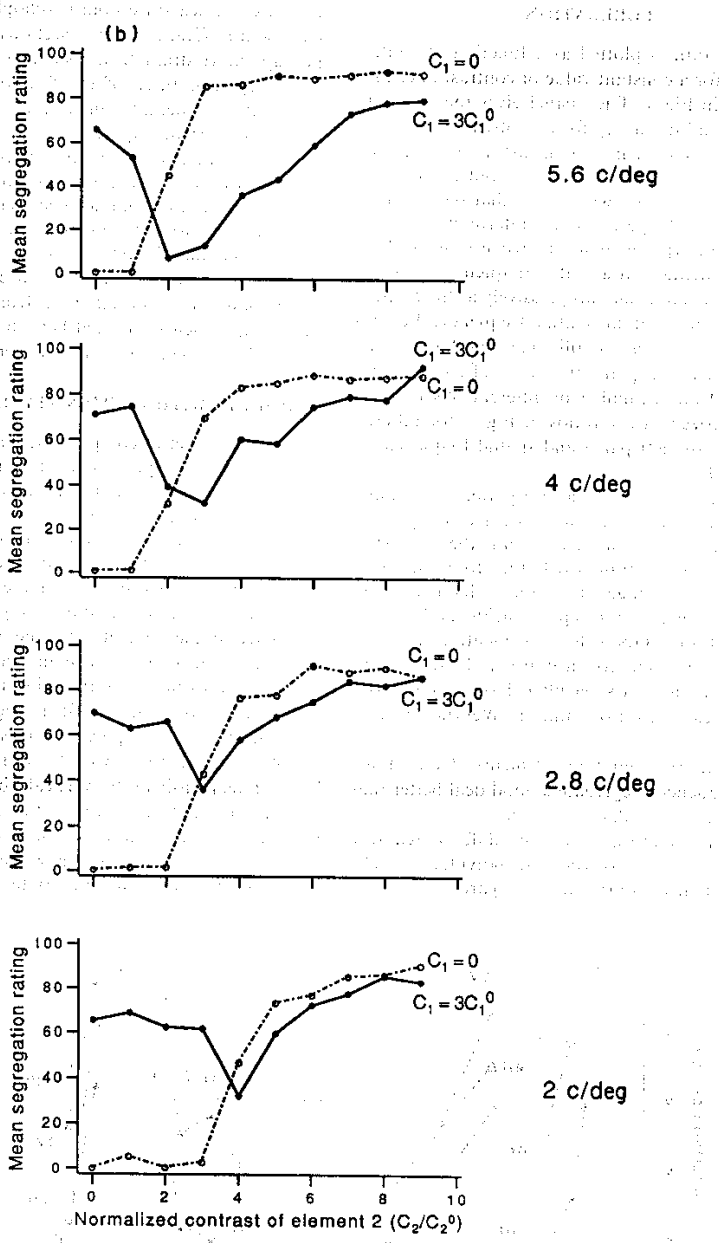

FIGURE 6. Results for element-arrangement textures of several different spatial-frequency combinations for observer CV. Average segregation rating as a function of contrast in element 2 when the contrast in element 1 was 0% (dotted curves) or 16% (3 times the base contrast—solid curves). The results when the contrast of element 1 was six times base contrast has been omitted for visual clarity. Element 1 was 8 c/deg. The spatial-frequency of element 2 is indicated next to each panel. The contrast in element 2 is plotted relative to C_2^0, the base contrast for the spatial frequency of that element. For this observer, corresponding to the set of spatial frequencies {16, 11.3, 8, 5.6, 4, 2.8, 2} c/deg, the set of base contrasts was {0.08, 0.08, 0.053, 0.027, 0.013, 0.009, 0.007}. Each point on the figure gives the average of four trials.

RESULTS WITH ELEMENT-ARRANGEMENT SEGREGATION

Segregation ratings plotted as a function of contrast of element 2 (for a constant value of contrast in element 1) are shown in Fig. 6. Each panel gives the results for a different spatial frequency for one observer (CV). The functions measured when the contrast of element 1 was zero (dashed lines) or three times base contrast (or about 16%—solid lines) are shown. Notice that the dip in the solid-line curves is biggest—i.e. the interference between the two element types is greatest—when element 2 and element 1 are identical at a spatial frequency of 8 c/deg. The interference diminishes progressively as the elements become more different in spatial frequency. For this observer, however, there is still some interference even when element 2 has a spatial frequency of 2 or 16 c/deg.

The results from a number of different observers are shown summarized as d/r ratios in Fig. 7 for differing orientations in the left panel and spatial frequencies in the right panel.

There is much less interference for patterns containing elements of very different orientations than for patterns containing similar orientations. Within the context of this general theoretical framework, therefore, the results imply that the first stage of complex channels is *not* composed of concentric receptive fields and thus is not composed of LGN cells. As mentioned in the Introduction, a different conclusion may hold for moving stimuli where the first-stage filters have been reported to be quite unselective for orientation (Werkhoven *et al.*, 1992).

Similarly, patterns containing elements of very different spatial frequencies segregate a good deal better than patterns containing similar frequencies.

Second, as shown in Fig. 7, subjects differ a great deal from each other. It is also true that individuals remain constant over many months in this regard.

Third, the bandwidths tend to be wider than those measured in near-threshold psychophysical experiments (reviewed in Graham, 1989, Sections 6.4, 6.5, and 12.2), perhaps more dramatically so for orientation than for frequency. For three of the four observers in Fig. 7, for example, definite interference occured between orientations 45 deg apart.

The near-threshold experiments presumably measure the bandwidth of something like simple linear channels rather than that of complex channel's first stages. Thus it is tempting to conclude that the bandwidths of the complex channels' first stages are somewhat broader than those of simple channels. Since the near-threshold experiments were so different from the segregation-rating experiments reported here, however, it seemed a little risky to make comparisons across them.

RESULTS WITH ELEMENT-TYPE SEGREGATION

In order to better compare the bandwidths of simple channels and complex channels' first stages, we estimated simple-channel bandwidths in segregation-rating experiments. For this purpose we used patterns like those in Fig. 8, where each region contains only a single type of element. Thus the difference between regions does not depend on differences in the *arrangement* of two types of elements (as it did in Fig. 3) but rather on differences in the element *type*. The two types of element could differ in orientation or in spatial frequency.

The patterns in Fig. 8 are quite similar to those used by Caelli and Moraglia (1985). To segregate patterns like those in Fig. 8, complex channels are not required. Simple linear channels will do. For example, a simple linear filter sensitive to vertical (but not horizontal) orientations and sensitive to the appropriate spatial frequency will respond well throughout the left and right regions of the patterns in Fig. 8 (i.e. will give a modulated response reflecting the bright and dark stripes in

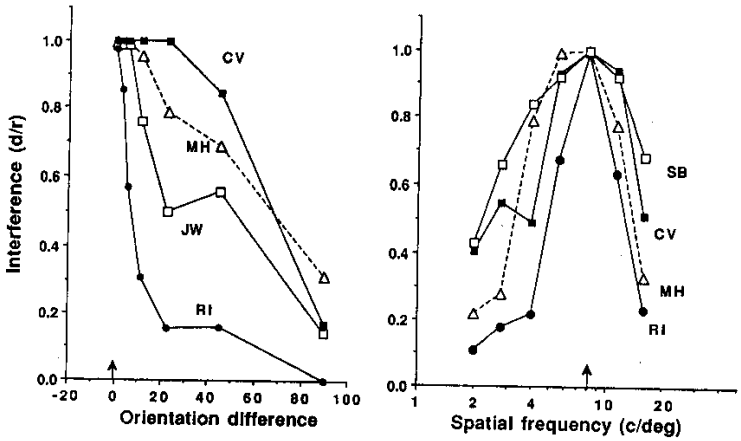

FIGURE 7. The relative sensitivity of the first stage of complex channels as estimated by the d/r ratio for element-arrangement pattern segregation ratings. The d/r ratio is plotted as a function of orientation difference (left panel) or spatial frequency of element 2 (right panel—the spatial frequency of element 1 was always 8 c/deg). Each curve gives results from a different observer.

each vertical Gabor patch) while responding poorly in the middle regions. Conversely, a simple linear filter sensitive to horizontal (but not vertical) orientations will respond well in the middle regions but not in the left and right regions. Thus, within this theoretical framework, the bandwidth derived from d/r ratios for these element-type patterns is an estimate of simple-channels' bandwidths.

These element-type patterns were used in the same kind of experiment as those reported above. In some cases, there were several slightly-different variations of the experiment. None of these variations turned out to matter, but the details of the experiments are described in the next paragraph.

Details of element-type experiments

The individual elements used in these experiments were identical to those used in the correspondent element-arrangement experiments. (It was only the arrangement of the elements in the patterns that differed.) In some experiments only element-type patterns were studied. Then the procedure was exactly like that described earlier. In other experiments, trials of all the element-arrangement and element-type textures were randomly intermixed. when both types of patterns were intermixed, there were about twice as many stimuli, so each subject took four sessions rather than two to complete the four replications of each stimulus.

In the case of observers RI and CV, we ran the intermixed experiments with two slightly different sets of patterns. In one, the orientation of element 1 was vertical (as was standard) and in the other set, the orientation of element 1 was horizontal. (The global orientation of the stripes in the striped region of element-arrangement patterns was always horizontal, as in Fig. 3.) This difference in stimuli did not seem to affect the result.

Both CV and RI replicated experiments of the type reported here over a period of ten months, and MH over a period of 3 months. RI and MH showed no changes whatsoever beyond those conservatively expected from sampling variability. CV's bandwidth for element-arrangement discriminations on the orientation dimension seemed to broaden slightly between the first experiments done (which are not shown here) and those shown here, but CV's bandwidth was never narrower than that for MH.

(a)

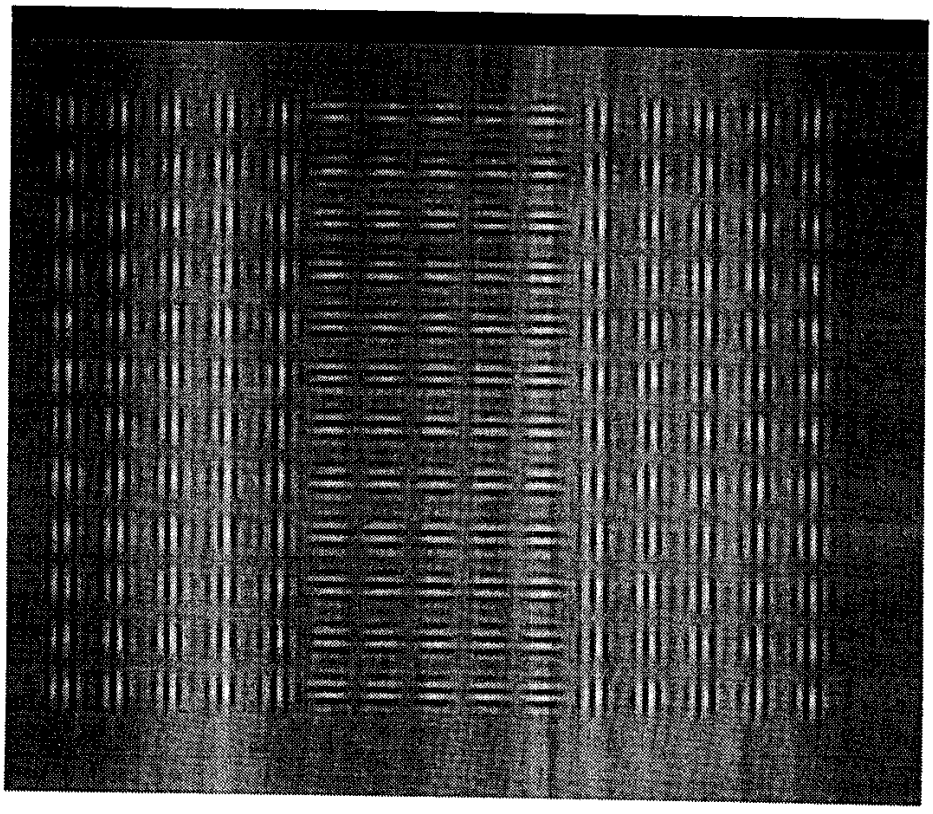

FIGURE 8(a). *Caption overleaf.*

(b)

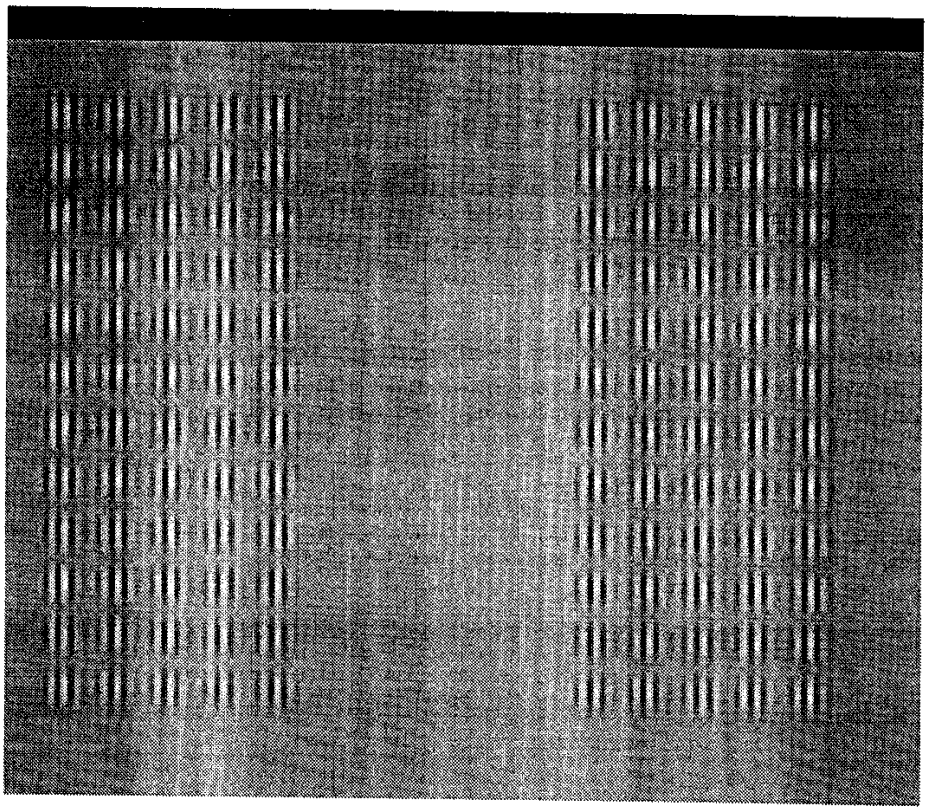

FIGURE 8. Examples of patterns in which the regions are distinguished by element type (each region containing only a single type of element).

Comparison of element-type and element-arrangement results

Figures 9 and 10 show the results with both element-arrangement (solid symbols) and element-type (open symbols) patterns. The d/r ratios (interference) are plotted against orientation difference between the two elements (Fig. 9) or against the spatial frequency of element 2 (Fig. 10).

For each observer, the curves for the element-type patterns are substantially narrower than those for element-arrangement patterns even though the absolute bandwidths vary greatly from one observer to another (particularly for the element-arrangement case). The results did not seem affected by the other manipulations done as controls: orientation of element 1, what set of orientation values was used, and whether element-arrangement and element-type patterns were in separate sessions or intermixed (represented by different symbols in Figs 9 and 10; see figure legends).

The results for the element-type patterns seem quite similar to those of Caelli and Moraglia (1985) who used similar patterns, although a direct comparison of the two sets of results is not possible because the psychophysical tasks were different (see e.g. Nothdurft, 1985; Sutter & Graham, 1992) and because Caelli and Moraglia kept the contrast in the two types of elements constant and thus discrimination could have been influenced by perceived contrast differences as well as by differences in perceived spatial frequency or orientation.

DISCUSSION

The results in Figs 9 and 10 can be viewed as estimates of relative sensitivity as a function of spatial frequency and orientation for the first stage of complex channels (the results from element-arrangement segregation—shown as solid symbols) and for simple channels (the results from element-type segregation—shown as open symbols). For all observers, therefore, the bandwidth estimated for the first stage of the complex channels was substantially broader than that for simple channels (by at least a factor of 2). For simple channels, the bandwidth estimate was rather like that from near-threshold psychophysical experiments, from 0.5 to 1 octaves on the

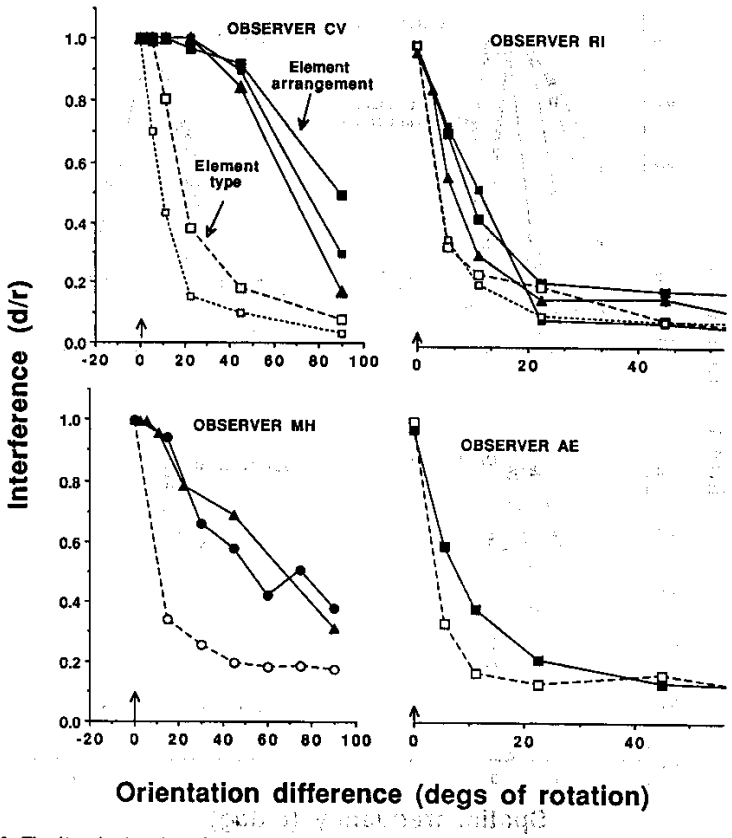

FIGURE 9. The *d/r* ratio plotted as a function of orientation difference for segregation of element-arrangement patterns (solid symbols) and element-type patterns (open symbols). These are estimates of the relative sensitivity of the first stage of complex channels and that of simple channels respectively. Different panels give results from different observers. Experiments including only one type of pattern (element-type or element-arrangement patterns but not both) are shown as triangles or circles. The results for element-arrangement segregation previously shown in Fig. 7 reappear in Figs 9 and 10 as solid triangles. A further element-arrangement experiment is shown as solid circles (observer MH in Fig. 9, run with a different set of orientations). Experiments where trials of both types of patterns are intermixed are shown as squares. In the case of observers RI and CV, we ran the experiment with two slightly different sets of patterns. The smaller square symbols show the case where the orientation of element 1 was vertical and the larger squares the case where it was horizontal. (The global orientation of the stripes was always vertical as shown in Figs 3 and 8.) Note that the horizontal axes on the righthand panels are more spread out than those on the lefthand panels.

spatial-frequency dimension and from 5 to 20 deg of rotation on the orientation dimension.

At least three different explanations, illustrated in Fig. 11, exist for why the bandwidths of the first stage of complex channels might be broader than the bandwidths of simple channels.

A straightforward possibility is that the receptive fields of the first stage of the complex channel might be shaped so as to produce a broader bandwidth.

A second possiblity is that the bandwidths of individual receptive fields in the first stage might be relatively narrow, but receptive fields with somewhat different preferred frequencies and orientations might all feed into the same second filter.

Thirdly, the bandwidth of individual receptive fields in the first filter might be relatively narrow but inhibition among channels might exist [as in a normalization network including channels tuned to different spatial frequencies and/or orientations—see e.g. Graham (1991), Graham *et al.* (1992a), Heeger (1991) and Robson (1988a,b)]. Such inhibition would widen the measured bandwidth in element-arrangement experiments like these (but not in element-type experiments) if the interchannel inhibition had some moderate spatial spread (as in Malik & Perona, 1990).

About the physiological substrate

The structure of the complex channel in Fig. 1 resembles that hypothesized for complex cortical cells (e.g.

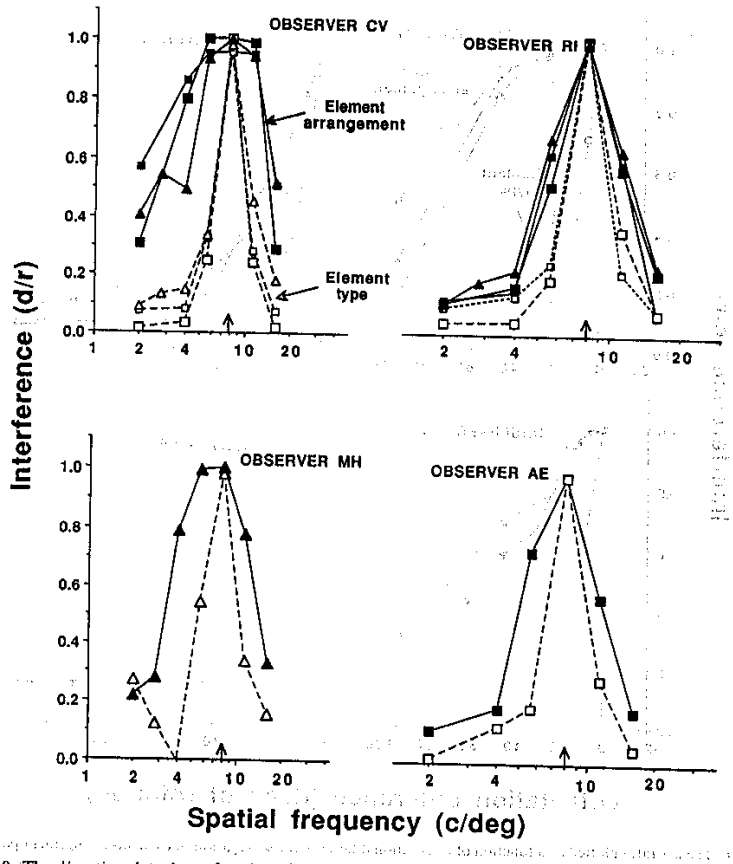

FIGURE 10. The *d/r* ratios plotted as a function of spatial frequency for the same four observers as in Fig. 9. See the legend of that figure for further description.

Hochstein & Spitzer, 1985; Heitger, Rosenthaler, von der Heydt, Peterhans & Kubler, 1992; Heeger, 1991). However, in the models of complex cortical cells, the linear filtering after rectification is generally a simple excitatory sum (i.e. the second filter in the cell models has only an excitatory region) whereas the second filter in the perceptual complex channels in Fig. 1 has an excitatory–inhibitory structure. [The model for cortical end-stopped cells proposed by Heitger *et al.* (1992) does have the feature of an excitatory–inhibitory difference between rectified outputs from smaller receptive fields, but it also includes other mechanisms beyond those in the complex channels of Fig. 1].

A possible terminological confusion. Because the second filter in these complex-cortical cell models is entirely excitatory, other authors in the texture literature have suggested a possible analogy to complex neurons that is somewhat different from the one implied by our use of the term "complex channels". Since the simple linear channels in others' models (as in ours) are followed by spatial-pooling which resembles an excitatory filtering, others (e.g. Bergen & Adelson, 1986, 1988; Bergen & Landy, 1991) have sometimes referred to this combination as being like complex cells. We, however, reserve the word "complex channels" to refer to a rectification-like-nonlinearity embedded between two stages of band-pass linear filtering (as in Fig. 1), where these channels all precede the pooling across spatial position and across channels that determines the observer's final response.

The best model for various cortical cells (including complex and end-stopped cells) as well as the relationship of these cortical complex cells to the processes underlying region segregation remain open questions.

The psychophysical results reported here do suggest that the physiological substrate for the first-stage filters of the complex channels involved in perceptual region segregation is not LGN cells since the first stage is very orientation-selective. The results here are probably consistent with a physiological substrate for the complex channels that is in V1 or any higher location. There is a certain amount of evidence that stimuli containing regions defined by different orientations (e.g. Knierim & van Essen, 1992; Lamme, Dijk & Spekreijse, 1992; Northdurft & Li, 1985; Van Essen, DeYoe, Olavarra,

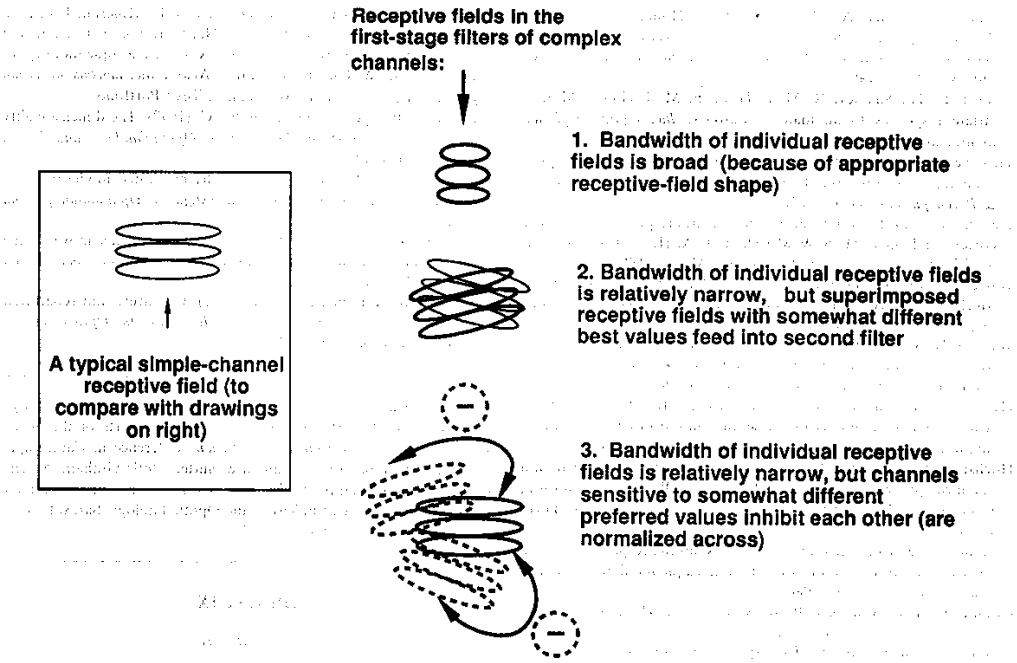

FIGURE 11. Diagram of three reasons why a complex channel's first stage might be more broadly tuned than a simple channel.

Knierim, Fox, Sagi & Julesz, 1989) are good stimuli for V1 cells. It is not yet known, however, whether the properties of cells in V1 match those necessary for the psychophysical results. Even if the substrate for the channels themselves were in V1, however, the substrate for region-growing or boundary-finding processes (that are essential for the perception of segregation although not studied here) might be at a higher level. Von der Heydt, Peterhans and Baumgartner (1989a, b) reported that illusory-border stimuli were effective in V2 but not in V1 [but see Grossof, Shapley and Hawken (1992) for a finding that other kinds of anomalous contours do seem effective in V1].

The results explained by complex channels here may be better explained eventually by an entirely different theoretical framework. This caveat is true for all results, of course, but may be particularly relevant here. The extent of the individual differences present in the perceived segregation of element-arrangement patterns (Figs 9 and 10) might suggest to some readers the action of higher-level processes. On the other hand, substantial differences exist between individual members of the same species in the number of cones in the retina and in the area V1 occupies in the cortex. It would be interesting to know how the functional properties of neurons in V1, V2, and higher areas compare from individual to individual in the same species.

REFERENCES

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.

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, Mass.: MIT Press.

Caelli, T. & Moraglia, G. (1985). On the detection of Gabor signals and discrimination of Gabor textures. *Vision Research, 5,* 671–584.

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.

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

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

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, Mass.: MIT Press.

Graham, N. & Sutter, A. (1991). Nonlinear processes in perceived region segregation: Spatial-frequency- and orientation-selectivity of the first stage of complex channels. Presented at *Contrasts in Vision,* Cambridge, England, September 1991. Abstract in *Ophthalmic and Physiological Optics.*

Graham, N. & Sutter, A. (1992). Spatial-frequency, and orientation-selectivity of simple and complex channels in region segregation. *Investigative Ophthalmology and Visual Science, 33,* 1354.

Graham, N., Beck, J. & Sutter, A. (1992a). Nonlinear processes in spatial-frequency channel models of perceived texture segregation: Sign and amount of contrast. *Vision Research, 32,* 719–743.

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

Grosof, D. H., Shapley, R. M. & Hawken, M. J. (1992). Monkey striate responses to anomalous contours? *Investigative Ophthalmology and Visual Science, 33,* 1257.

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

Heeger, D. J. (1991). Nonlinear model of neural responses in cat visual cortex. In Landy, M. S. & Movshon, J. A. (Eds), *Computational models of visual processing.* Cambridge, Mass.: MIT Press.

von der Heydt, R., Peterhans, R. E. & Baumgartner, G. (1989a). Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *Journal of Neuroscience, 9,* 1731–1748.

von der Heydt, R., Peterhans, R. E. & Baumgartner, G. (1989b). Mechanisms of contour perception in monkey visual cortex. I. Contours bridging gaps. *Journal of Neuroscience, 9,* 1749–1763.

Heitger, F., Rosenthaler, L., von der Heydt, R., Peterhans, E. & Kubler, O. (1992). Simulation of neural contour mechanisms: From simple to end-stopped cells. *Vision Research, 32,* 963–981.

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

Knierim, J. J. & Van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neuorphysiology, 67,* 961–980.

Lamme, V. A. F., van Dijk, B. W. & Spekreijse, H. (1992). Texture segregation is processed by primary visual cortex in man and monkey. Evidence from VEP experiments. *Vision Research, 32,* 797–807.

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

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

Nothdurft, H. C. & Li, C. Y. (1985). Texture discrimination, representation of orientation and luminance differences in cells of the cat striate cortex. *Vision Research, 25,* 99–113.

Pantle, A. (1992). Immobility of some second-order stimuli in human peripheral vision. *Journal of the Optical Society of America A, 9,* 863–867.

Robson, J. G. (1980). Neural images: The physiological basis of spatial vision. In Harris, C. S. (Ed.), *Visual coding and adpatability* (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, New York, June 1988. Abstract p. 5.

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. *Investigative Ophthalmology and Visual Science (Suppl.), 30,* 161.

Sutter, A. & Graham, N. (1992) Investigating the dynamics of region segregation. How immediate is it? *Investigative Ophthalmology and Visual Science (Suppl.), 33,* 956.

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

Sutter, A., Sperling, G. & Chubb, C. (1991). Further measurements of the spatial frequency selectivity of second-order texture mechanisms. *Investigative Ophthalmology and Visual Science (Suppl.), 32,* 1039.

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

Van Essen, D. C., DeYoe, E. A., Olavarra, J. F., Knierim, J. J., Fox, J. M., Sagi, D. & Julesz, B. (1989). Neural responses to static and moving texture patterns in visual cortex of the macaque monkey. In Lam, D. M. & Gilbert, C. (Eds), *Neural mechanisms of visual perception* (pp. 137–154). Woodlands, Tex.: Portfolio.

Werkhoven, P., Sperling, G. & Chubb, C. (1992). The dimensionality of motion-from-texture. *Investigative Ophthalmology and Visual Science, 33,* 1049.

Wilson, H. R. (1991). A psychophysically motivated model for two-dimensional motion perception. *Investigative Ophthalmology and Visual Science, 32,* 893.

Wilson, H. R. & Kim, J. (1992). Non-Fourier plaids move in the vector-sum direction. *Investigative Ophthalmology and Visual Science, 33,* 1049.

Wilson, H. R. & Richards, W. A. (1992). Curvature and separation discrimination at texture boundaries. *Journal of the Optical Society of America A, 9,* 1653–1662.

Acknowledgements—The research was supported by National Eye Institute Grant 1 RO1 EYO8459. Preliminary reports of this work were presented at the *Contrasts in Vision* conference in Cambridge, England, September 1991 (Graham & Sutter, 1991; Graham, Sutter, Venkatesan & Humaran, 1992b) and at the meetings of the Annual Association for Research in Vision and Ophthalmology, Sarasota, Fla, May 1992 (Graham & Sutter, 1992).

APPENDIX

About Models

About pooling rules

To fully specify these models, the outputs of the simple and/or complex channels need to be related to the observer's possible responses in an experiment. These relationships are briefly described below. Much fuller descriptions can be found in Graham *et al.* (1992a) and Sutter *et al.* (1989). The family of rules we consider includes the rules used by a number of other models of texture segregation.

Pooling across spatial position. The spatially-pooled output from a single channel (which will be referred to below as the spatially-pooled difference response from the channel) might be taken to be the following difference: the peak–trough amplitude in the channel's outputs in one region minus the peak–trough amplitude in its outputs in the other regions. Or it might be some other comparison between amount of modulated output in the different regions. Some alternative spatial pooling rules are given and investigated in Sutter *et al.* (1989) and Graham *et al.* (1992a). For example, the standard deviation of the outputs at different positions in one region (or, equivalently, the standard deviation of the outputs at different positions within one full period of one region) might be subtracted from the standard deviation of the outputs in the other. This particular rule is essentially equivalent to a local-energy computation like that used by others (e.g. Bergen & Adelson, 1986, 1988; Bergen & Landy, 1991). For the arguement here, all of the spatial-pooling rules we have considered behave the same way.

Pooling across channel outputs. The relationship between the spatially-pooled difference responses from the various channels and the observer's response need to be specified next; i.e. some rule for pooling information across different channels needs to be specified. The naive two-channel model presented below assumes the observer's response to be the maximum of the two channels' spatially-pooled difference responses. Some alternative rules for pooling information across channels were considered by Sutter *et al.* (1989), Graham (1991) and Graham *et al.* (1992a). For the argument here, these other rules do *not* behave exactly like the maximum rule, and their behavior is discussed further below.

Pooling rules assume knowledge of boundary position. The kind of computation just described assumes knowledge of the position of the boundary. This is an unrealistic assumption of course, and it should be seen as a simplified assumption that is suitable for this situation (where, e.g. the boundary itself is not manipulated). In reality, one

presumes there is a region-growing or contour-computing process that actually finds the boundary. The relationship of our assumption to models including such a process is briefly described in Graham *et al.* (1992a). The work described here is not helpful in further elucidating such processes.

The naive two-channel model

Equation (1) is repeated here from the main text.

$$\frac{d}{r} \approx \frac{S_A(2)}{S_A(1)} \times \frac{S_{obs}(1)}{S_{obs}(2)} \quad (1)$$

it can be derived from a naive, two-channel model. This model assumes that there are exactly two channels responsive to the patterns in question. These are two complex channels in the case of the element-arrangement patterns and two simple channels in the case of the element-type patterns. The general derivation is the same in both cases as are many details. Where the details differ below, they will be briefly spelled out.

In the case of the two simple channels assumed sensitive to the element-type patterns, one of the two channels is more sensitive to the first element than to the second element, the other channel is more sensitive to the second element than to the first (as shown in the left diagrams of Fig. A1), and the channels are symmetric [see equation (A6)].

In the case of the two complex channels assumed sensitive to the element-arrangement patterns, one of the two channels has a first stage that is more sensitive to the first element than to the second element, the other channel has a first stage that is more sensitive to the second element than to the first (as shown in the left diagrams of Fig. A1), and the channels are symmetric [as in equation (A6) below]. Further, the second-stage filter is maximally sensitive at the periodicity of the pattern, and the nonlinearity between the two filters is a full-wave rectification.

This model is very similar to the initial model used for summation-near-threshold experiments [e.g. comparing the threshold for a compound grating containing two spatial frequencies to the thresholds for the components alone—see, e.g. Graham (1989) Section 6.2]. Here, however, the two values (e.g. two spatial frequencies of Gabor-patch elements) *interfere* with one another's effects when they stimulate the same channel; in the near-threshold summation experiments, two values stimulating the same channel *augment* one another. The effect of some modifications of this naive two-channel model are discussed below.

As indicated by the diagrams in the left column of each row in Fig. A1, the top row shows the case of channels with non-overlapping sensitivities (as when the first-stages of the complex channels are so selective that each channel responds to only one of the two elements). The bottom row shows the case where the two channels have entirely overlapping sensitivities (as when the two elements are identical). And the middle row shows an intermediate case.

The middle column of Fig. A1 shows the responses of each channel and of the observer as a function of contrast in element 2 (when contrast in element 1 is held constant). The channel response is the difference between the spatially-pooled outputs in the different regions, and the observer's response is taken to equal the maximum of the two channels' responses. The right column shows an enlarged version of the diagram in the middle row of the middle column; this enlarged version is labeled to illustrate the derivation of equation (1) described below.

On plots like that on the right of Fig. A1, the vertical intercept of each channel's function depends on the channel's sensitivity to element 1 (since the contrast in element 2 is zero for the vertical intercept). The slope of the channel's function depends on the channel's sensitivity to element 2 (since the contrast in element 2 is varying on the horizontal axis while the contrast in element 1 is held constant). As the contrast in element 2 increases from zero, the channel's response will generally first decrease (as the increasing presence of element 2 interferes more and more with the channel's original ability to segregate on the basis of element 1). This function will reach zero when element 2 and element 1 are stimulating the channel equally. (An example of this in the case

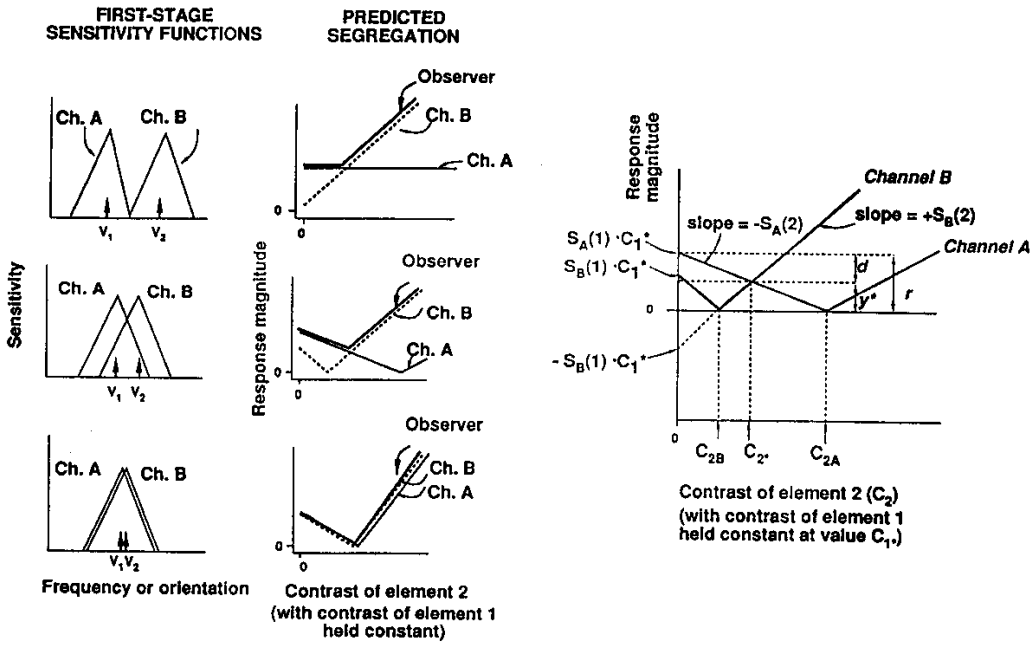

FIGURE A1. Diagram of naive two-channel model. Left column shows three possible sets of channel sensitivity functions. Middle column show their predictions for response magnitude as a funciton of contrast in the second element (when contrast in the first element is held constant at some relatively high value). Righthand panel is an enlargement of these predictions to allow annotation with the symbols used in the derivations.

of complex channels and element-arrangement textures is illustrated in the middle column in Fig. 4.) For further increases in the contrast of element 2, the channel's response will again increase. (That these increases and decreases are linear rather than curvilinear is by assumption. The simple channels are linear filters by assumption. In the case of complex channels, the intermediate nonlinearity is assumed to be a full-wave rectification in this naive model, and both filterings are assumed linear.)

Proof of equation (1) relating the d/r ratio to channel sensitivity

In symbols (see labels on Fig. A1), let C_1^* stand for the contrast in element 1 (which is fixed for a given curve) and C_2 for the contrast of element 2 (which varies and is plotted on the horizontal axis).

Let $R_A(C_2)$ stand for channels A's response at C_2 and C_{2A} stand for the contrast of element 2 at which $R_A(C_2)$ reaches its minimum (zero).

Let $S_A(1)$ and $S_A(2)$ stand for channel A's sensitivity to elements 1 and 2 respectively.

Then, for these two channels in the naive model, the *decreasing* part of channel A's function (which will be the part of interest) is:

$$R_A(C_2) = S_A(1) \cdot C_1^* - S_A(2) \cdot C_2 \quad (\text{for } C_2 < C_{2A}). \tag{A1}$$

And the *increasing* part of channel B's function is:

$$R_B(C_2) = S_B(1) \cdot C_1^* - S_B(2) \cdot C_2 \quad (\text{for } C_2 > C_{2B}). \tag{A2}$$

The observer's response is (according to this naive model) the envelope of the two channel's functions. Thus, as shown in the diagram on the right of Fig. A1, the decreasing part of channel A's function and the increasing part of channel B's intersect at the minimum of the observer's function. Let C_2^* be the contrast of element 2 at that intersection, and let y^* be the vertical coordinate there. Thus,

$$y^* = R_A(C_2^*) = R_B(c_2^*). \tag{A3}$$

By substitution of equations (A1), (A2) into (A3) and appropriate rearrangement one finds that:

$$C_2^* \cdot [S_A(2) + S_B(2)] = C_1^* \cdot [S_A(1) + S_B(1)]. \tag{A4}$$

Since the observer's response is assumed to be the maximum of the two channels' responses, the observer's sensitivity will be the maximum of the two channel's sensitivities:

$$S_{obs}(1) = S_A(1) \quad \text{and} \quad S_{obs}(2) = S_B(2). \tag{A5}$$

By the assumption that the channels are symmetric:

$$S_B(1) = k \cdot S_A(1) \text{ and } S_A(2) = k \cdot S_B(2). \tag{A6}$$

Then, by equations (A4)–(A6)

$$\frac{C_2^*}{C_1^*} = \frac{[S_A(1) + S_B(1)]}{[S_A(2) + S_B(2)]} = \frac{S_{obs}(1)}{S_{obs}(2)}. \tag{A7}$$

Also, since y^* is the value of both channel's responses at contrast C_2^*, then by equation (A1) for channel A applied to C_2^*

$$y^* = R_A(C_2^*) = S_A(1) \cdot C_1^* - S_A(2) \cdot C_2^* \quad (\text{for } C_2 < C_{2A}). \tag{A8}$$

And, the "range" r is the response of channel A when $C_2 = 0$. Thus

$$r = S_A(1) \cdot C_1^*. \tag{A9}$$

Also $d = r - y^*$. Hence by application of (A9), (A8), and (A7), one finds

$$\frac{d}{r} = \frac{r - y^*}{r}$$

$$= \frac{S_A(1) \cdot C_1^* - [S_A(1) \cdot C_1^* - S_A(2) \cdot C_2^*]}{S_A(1) \cdot C_1^*}$$

$$= \frac{S_A(2) \cdot C_2^*}{S_A(1) \cdot C_1^*} = \frac{S_A(2) \cdot S_{obs}(1)}{S_A(1) \cdot S_{obs}(2)}. \tag{10}$$

Thus equation (1) of the text has been proved.

A final nonlinearity

To account for the fact that the observer's responses were limited to a finite range, while the predictions of this model can increase to infinity, one might wish to use a final nonlinear function of the predictions. We tried to do so by estimating the final nonlinearity from the curves where only one element was present, that is, when $C_2 = 0$ (as in the dashed curves in Fig. 6) and then predicting the curves for the conditions where $C_2 > 0$. We found that the sensitivities estimated with this additional complexity did not in general differ very much (or systematically) from those estimated simply by the d/r ratio. However, these estimates were occasionally dramatically affected by what seemed to be minor changes in the fitting procedure or in the data. In trying to understand this erratic behavior we realized that some further kind of intensity-dependent nonlinearity was also necessary for a complete description of these results. Candidates include those described in Graham (1991) and Graham *et al.* (1992a)—either a local nonlinearity before the channels or an inhibitory interaction among the channels modeled by a normalization network. Thus, we did some calculations with these intensive nonlinearities that are briefly described at the end of the next section.

More sophisticated models

Three other modifications of the naive model in Fig. A1 are considered briefly in this section. These—especially the first two—are similar to modifications often discussed for the models of summation-near-threshold experiments (see Graham, 1989, Section 6.3). However, the effects on estimated sensitivity (and hence bandwidth) are in the opposite direction (since here interfering effects rather than summating effects are the signature of two values stimulating the same channel).

Pooling across channels. For one thing, the assumption that the observer's response is the maximum of the two channel's responses can be loosened. The observer's response is often considered instead to be some more equal combination of both of the channels' responses. For example, Sutter *et al.* (1989) and Graham *et al.* (1992a) used the Quick Pooling Formula with a variety of exponents lower than infinity, and a number of investigators have used an exponent of 2 (often called an energy measure, and equivalent to taking the standard deviation). On this kind of combination of channel outputs, the observer's response in diagrams like those on the right of Fig. A1 would not dip down quite as far as does the envelope of the two channels. Then the measured d/r ratio in the observer's response—if used in equation (1) to estimate the relative sensitivity of individual channels—would tend to *underestimate* that relative sensitivity and hence underestimate the bandwidth of channels.

We did some calculations with the naive model above altered to allow Quick pooling across the two channels and studied exponents in the range from 1 to 4. As expected, the d/r ratio now underestimated the true relative sensitivity of the channels to their less-preferred values. This effect tended to be greater (both absolutely and proportionally) for low values of true relative sensitivity, and the greatest absolute underestimate was about 0.15. For example, (i) a true relative sensitivity of 0.25 led to d/r ratios of about 0.10 when the exponent was 2.0, but (ii) a true relative sensitivity of 0.75 led to estimates of 0.713 or greater for all exponents in the range 1–4.

More than two channels. Second, if there are more than two channels active in the situation, more than two channels' functions would have to be drawn in the segregation vs contrast diagrams of Fig. A1. With the maximum rule, these extra functions would either have no effect on the d/r ratio or, by filling in the dip further, would cause the measured d/r ratio to again underestimate the relative sensitivity and bandwidth of individual channels. With other pooling-across-channel rules, the effects can be more subtle but will tend to cause the d/r ratio to underestimate the relative sensitivity and hence the bandwidth of individual channels.

Nonlinear effects on inidividual channels' responses. Third, perhaps the function for each individual channel should not be composed of straight line segments as in Fig. A1 but of curvilinear segments. (Within the context of summation-near-threshold experiments on the spatial-frequency or orientation dimensions, curved functions for individual channels can be produced by spatial probability summation or analogous spatial pooling.) In the context of these element-arrangement texture-segregation experiments, curved functions could be produced by postulating nonlinearities other than straight rectification (e.g. squaring) at the middle stage of the complex channels. They

would also be produced by the two candidate intensive nonlinearities in Graham *et al.* (1992a): an early local nonlinearity before the channels (which might be active in either the element-type or element-arrangement experiments) or intra-channel inhibition as in a normalization network (which, as indicated in the discussion section, might be expected to be more active in the element-arrangement than in the element-type experiments). If such curved functions are the correct description for individual channels, the relationship of the d/r ratio to the sensitivity of individual channels will in general be changed from that in equation (1). The direction of the change would depend on the kind of curvature in the individual channel's curves.

We did some sample computations of the possible effect of two candidate intensive nonlinearities like those in Graham *et al.* (1992a). As it turns out, there is very little effect at all with their particular models assuming compressive early, local, nonlinearity. With the models assuming normalization, however, the d/r ratios *overestimate* relative sensitivity (for all exponents in the pooling-across-channels rule) with the greatest absolute effect occuring for true relative sensitivities near 0.5 (e.g. a true value of 0.5 produces a d/r ratio of 0.58 to 0.60 for exponents in the 1–4 range and a d/r ratio of 0.61 for the maximum rule).

In future work, we will try to answer the question of which intensive nonlinearity is a better description. That would allow us to improve somewhat on the relative sensitivity estimates made here by the d/r ratio. Note, however, that the distortion introduced by our ignorance of the appropriate intensive nonlinearity (as indeed by our ignorance of exactly how many channels to include and the proper rule for pooling across channels) seems smaller than the differences among individual observers' results.