Available online at www.sciencedirect.com
SCIENCE DIRECT®
Vision Research 44 (2004) 3145–3175
Vision Research
www.elsevier.com/locate/visres

Is there opponent-orientation coding in the second-order channels of pattern vision?

Norma Graham *, S. Sabina Wolfson

Department of Psychology, Columbia University, Mail Code 5501, New York, NY 10027, USA

Received 19 December 2003; received in revised form 5 March 2004

Abstract

Is there opponency between orientation-selective processes in pattern perception, analogous to opponency between color mechanisms? Here we concentrate on possible opponency in second-order channels. We compare several possible second-order structures: *SIGN-opponent-only* channels in which there is no opponency between orientations (also called complex channels or filter-rectify-filter mechanisms); three structures we group under the name *ORIENTATION-opponent*; and finally *BOTH-opponent* channels which combine features of both SIGN-opponent-only and ORIENTATION-opponent channels but lead to predictions that are distinct from either of theirs. We measured observers' ability to segregate textures composed of checkerboard and striped arrangements of vertical and horizontal Gabor grating patches. The observers' performance was compared to model predictions from the alternative opponent structures. The experimental results are consistent with SIGN-opponent-only channels. The results rule out the ORIENTATION-opponent and BOTH-opponent structures. Further, when the models were expanded to include a contrast gain-control (inhibition among channels in a normalization network) the SIGN-opponent-only model was also able to explain a contrast-dependent effect we found, thus providing another piece of evidence that such normalization is an important process in human texture perception.

Keywords: Orientation opponency; Contrast gain control; Texture; Gabors; Normalization; Second-order channels; Filter-rectify-filter mechanisms

1. Introduction

Opponency between orientation-selective processes in pattern perception—somewhat analogous to opponency between color mechanisms—has been proposed explicitly or implicitly a number of times over the last 15 years (e.g. Arsenault, Wilkinson, & Kingdom, 1999; Bergen & Landy, 1991; Gray & Regan, 1998; Kingdom & Keeble, 1996, 1999, 2000; Kingdom, Keeble, & Moulden, 1995; Kwan & Regan, 1998; Landy & Bergen, 1991; Nothdurft, 1997; Prins & Mussap, 2000, 2001; Rubenstein & Sagi, 1993). The kind of opponency suggested has varied, sometimes being specified only vaguely, sometimes in detail. The question of whether or not there is orientation opponency of any particular kind remains open and several recent papers consider this question (Kingdom, Prins, & Hayes, 2003; Motoyoshi & Kingdom, 2003; Prins, Nottingham, & Mussap, 2003). Why is the possibility of orientation opponency so interesting? Locations in the visual field where change occurs are frequently important, and orientation opponency might help find the locations where orientation changes quickly, thereby providing a clue to important boundaries in the real world. Here we present results which cleanly distinguish among several of the most frequently-suggested types of orientation opponency.

In particular, we investigate the possibility of opponency between perpendicular orientations in

* Corresponding author.
E-mail addresses: nvg1@columbia.edu (N. Graham), sw354@columbia.edu (S.S. Wolfson).

0042-6989/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2004.07.018

second-order channels. Second-order channels consist of two stages of linear receptive fields with an intermediate pointwise nonlinearity inbetween. Or, less technically, second-order channels can be described as structures in which neurons having small receptive fields feed into neurons with large receptive fields. These second-order channels are necessary to account for many phenomena of texture perception and related perceptual tasks (e.g. see recent reviews in Landy & Graham, 2003, and in introduction to Graham & Sutter, 1998). Here we compare experimental results to the predictions from several possible types of second-order channels with and without orientation opponency. These channels are embedded in models allowing for other nonlinear processes which have been suggested as necessary to explain pattern perception. The sketch in Fig. 1 shows the overall framework of these models including: (i) both first-order and second-order channels preceded by an overall sensitivity setting stage; (ii) the pointwise nonlinearity—shown as expansive in the figure—between the two stages of linear filtering in the second-order channels; (iii) a compressive nonlinearity in the form of inhibition among channels in a normalization network; and (iv) the comparison, pooling, and decision processes (''comparison-and-decision stage'') incorporating the assumptions that relate the channel outputs to the observer's responses. Each feature in this model is discussed below in the main text as it becomes relevant. Further information, including all equations, is given in appendix.

1.1. Possible opponency in second-order channels

Diagrams of five possible structures for second-order channels are shown in Fig. 2. The name we will use for each type of structure is given at the top. In each of the diagrams, the large ovals represent the receptive field characterizing the second stage; the large plus and minus signs inside represent excitatory and inhibitory influences respectively. The small ovals (with small plus and minus signs inside) represent the receptive fields at the channels' first stage. The shorthand labels (e.g. −V +2V −V) are explained in the figure legend. The intermediate pointwise nonlinearity between the two stages is NOT represented in the diagrams here. We will consider a variety of possible intermediate nonlinearities in our models.

In the SIGN-opponent-only channel (left panel of Fig. 2), the second-stage receptive field has an excitatory center and inhibitory flanks. The first-stage receptive fields that feed into the second-stage excitatory center have the same orientation as those feeding into the second-stage inhibitory flanks. Thus there is no orientation opponency in this structure. (Note that we are NOT assuming that the orientation of the first-stage receptive fields is necessarily the same as the orientation of the second-stage receptive field.) Some other arrangements for this sign-opponent only structure are sketched in Fig. 3. One could also form receptive fields (not shown in Fig. 3) with a SIGN-opponent-only structure by reversing the signs from those shown so that the center was inhibitory and the flanks excitatory. Their predictions will always be identical to those with signs as shown and thus we will rarely mention them below.

SIGN-opponent-only channels have generally been called ''complex channels'' in our previous work (e.g. Graham, Beck, & Sutter, 1992; Graham & Sutter, 1998, 2000; Graham & Wolfson, 2001; Sutter, Beck, & Graham, 1989). They have also been called ''single-opponent'' (e.g. Prins & Mussap, 2001; Prins et al., 2003) and ''filter-rectify-filter'' (e.g. Kingdom et al., 2003; Mussap, 2001) processes.

In an *ORIENTATION-opponent-only* channel (second-to-left panel of Fig. 2), the first-stage receptive fields' feeding into the center of the second stage are perpendicular to those that feed into the second-stage flanks. Thus there is opponency between the orientations stimulating the second stage center and surround. However, both the second-stage center and surround are excitatory so there is no sign opponency. This kind of structure has been suggested by, for example, Rubenstein and Sagi (1993)

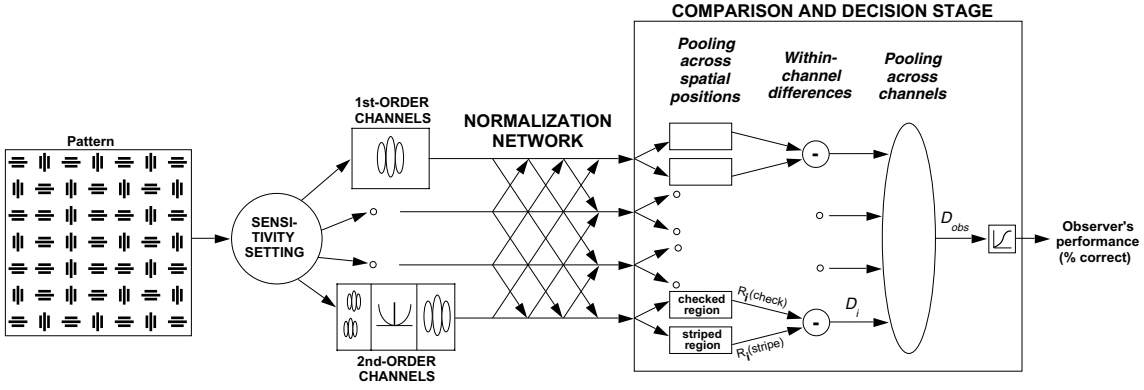

Fig. 1. Sketch of overall framework of the models used in this study. See text for explanation.

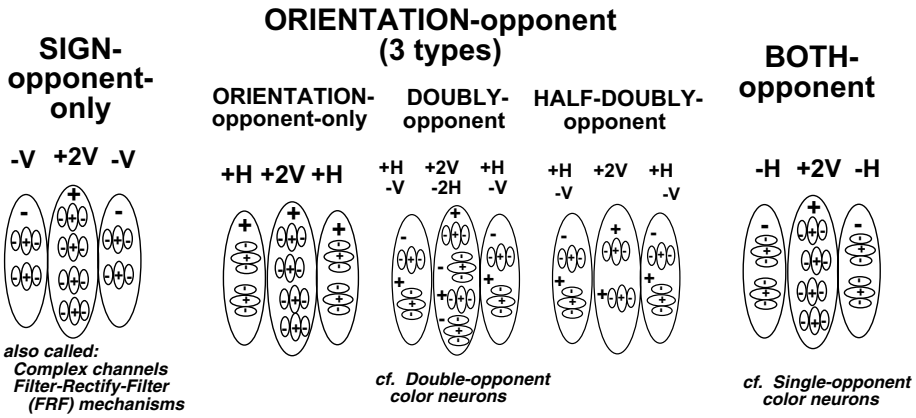

Fig. 2. Diagrams of five possible structures of second-order channels. A second-order channel consists of two stages, with relatively smaller receptive fields at the first stage and larger ones at the second, with a nonlinearity (e.g. a rectification) inbetween the two stages. In each of the five diagrams, the large ovals, large plus signs, and large minus signs represent the receptive field of the second stage. The pointwise nonlinearity that intervenes between the first and second stages is NOT represented here. The small ovals, small plus signs, and small minus signs represent the receptive fields of the first-stage filters that feed into each section of the second-stage receptive field. The words at the top of each diagram give a general name for the kind of opponency shown in the diagram, e.g. SIGN-opponent-only. Below the general name there is a label representing the structure of the particular example shown in the diagram. For example, in the symbol −V +2V −V in the left panel, the +2V for the center means that the input to the second-stage receptive-field center is excitatory, from vertically-oriented first-stage receptive fields, and of weight 2 (where the weight is in arbitrary units). Similarly the symbol −V in each flank means that the input to each second-stage receptive-field flank is inhibitory, from vertically-oriented first-stage receptive fields, and of weight 1. The other symbols should be interpreted analogously where H stands for horizontally-oriented first-stage receptive fields. The three middle structures in Fig. 2—ORIENTATION-opponent-only, DOUBLY-opponent, and HALF-DOUBLY-opponent—are grouped under the name ORIENTATION-opponent because all three make the same predictions for many experiments including those reported here. The results reported here are consistent with predictions from the SIGN-opponent-only channels but not with those from the other structures.

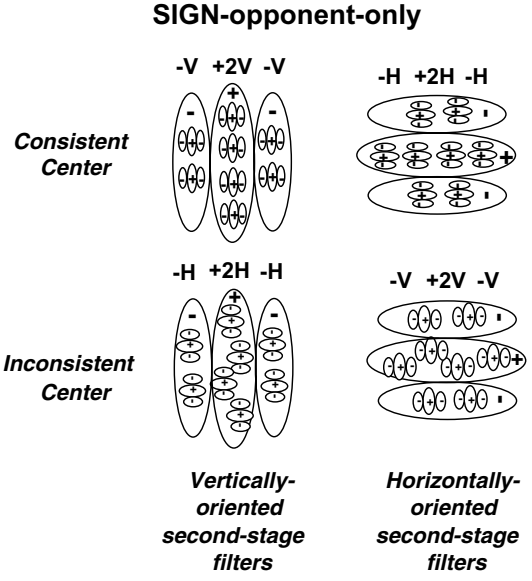

Fig. 3. Illustrating the various arrangements of orientations at the first and second stage that are possible for the SIGN-opponent-only structure. (The version from the left panel of Fig. 2 appears in the top left here. The version from Fig. 7 appears in the bottom left here.) One could also form receptive fields by switching signs of the second stage center and surround (not shown) but this will NOT change the predictions. Analogous possibilities (not shown) exist for the ORIENTATION-opponent and BOTH-opponent structures as well.

and Kingdom and Keeble (1996). There are many possible arrangements of this structure analogous to those in Fig. 3 for the SIGN-opponent-only structure. Also one could again switch signs to make both centers and flanks inhibitory rather than excitatory; this leads to predictions identical to those with signs as shown and thus will rarely be mentioned below.

In a *DOUBLY-opponent* channel (middle panel of Fig. 2), both vertical and horizontal first-stage receptive fields' outputs feed into both the center and surround regions of the second-stage receptive field with excitatory and inhibitory signs as shown in the figure. These DOUBLY-opponent channels are analogous to the double-opponent cells discussed in the color perception literature where horizontal and vertical here play the roles of the opponent colors (e.g. red and green) in the color literature (see, for example, Fig. 31-5 in Gouras (1991, Chapter 31)). (In the color-literature the second-stage receptive fields are often concentric rather than elongated as here.)

Still another form is shown is the *HALF-DOUBLY-opponent* (second-to-right panel of Fig. 2). This structure is just like the DOUBLY-opponent one shown to its left except that there is only one orientation of first-stage receptive field feeding into the second-stage center while both orientations still feed into the surround.

The DOUBLY-opponent channels of Fig. 2, or close relatives like the HALF-DOUBLY-opponent channels, have been considered a number of times in the pattern perception literature (e.g. Gray & Regan, 1998; Prins & Mussap, 2000, 2001).

For many experiments including the experiments reported here, one can show that all three structures—ORIENTATION-opponent-only, DOUBLY-opponent, and HALF-DOUBLY-opponent—make the same predictions. These three types of opponent structures will therefore frequently be referred to together here. And for lack of a better word, we will refer to all three of these structures together as *ORIENTATION-opponent* channels. (This word has been used in so many ways by different authors that we hesitate to use it, but we can not think of a better short alternative.)

In a *BOTH-opponent* channel (right panel Fig. 2) the small receptive fields of the neurons feeding into the second-stage center are perpendicular to those feeding into the flanks, and the center is excitatory whereas the surround is inhibitory. Therefore this receptive field shows both orientation opponency and sign opponency. There are many possible orientation arrangements of this structure (analogous to those in Fig. 3 for the SIGN-opponent-only). One could also form receptive fields by switching signs so that the center was inhibitory and the flanks excitatory, but this again leads to predictions identical to those with signs as shown and thus will rarely be mentioned below. BOTH-opponent channels are analogous to what have been called "single-opponent" neurons in the color literature (see, for example, Gouras (1991) Fig. 31-5). Orientation opponency of the BOTH-opponent type has been considered by Motoyoshi and Kingdom (2003).

As will be shown below, the three types of opponent structure just described—SIGN-opponent-only, ORIENTATION-opponent (containing three subtypes), and BOTH-opponent—lead to predictions that are very different from one another. And only the SIGN-opponent-only channels lead to predictions that are consistent with the results reported here from experiments using perpendicular orientations. Further, including normalization in the model of SIGN-opponent-only channels leads to correct prediction of not only the overall trend in the results but also of a contrast-dependent effect.

1.2. Overview of experiment

The segregatability of element-arrangement textures is the perceptual task used here because it affords a clean distinction among different kinds of opponency. Element-arrangement textures (first introduced by Beck, 1982; Beck, Prazdny, & Rosenfeld, 1983; see Wolfson & Graham, in press) are composed by arranging two types of elements either in a striped- or checkerboard-arrangement. Here we used element-arrangement textures where the two types of elements were both patches of sinusoidal grating (Gabor patches) identical in spatial characteristics but one type was oriented horizontally and the other vertically, like the patterns used by Graham et al. (1992); (Graham, Sutter, & Venkatesan, 1993). Fig. 4 shows three pieces of textures from the present study—all with a striped arrangement. The right panel shows a *compound* texture where the two element types have identical contrast. The left and middle panels of Fig. 4 show the corresponding *component* one-element-only textures where one element has contrast zero and thus is invisible. When the elements in a one-ele-

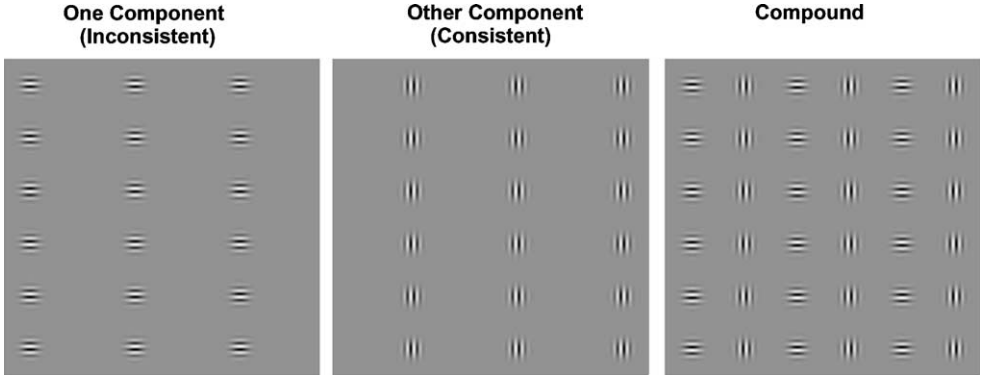

Fig. 4. Illustrating the compound and component patterns used in this study. This figure shows portions of the striped arrangement textures from three patterns used in the experiments here. The rightmost panel shows a *compound* texture containing both element types. The two kinds of elements are Gabor patches of identical spatial characteristics and contrast except their orientations are perpendicular. The left and middle panels show the two *components* of this compound. These components are *one-element-only* textures. (One type of element has contrast zero and thus is invisible.) When the elements in a one-element-only texture have the same orientation as the stripes, the pattern will be called the *consistent* component (middle panel). When the element and stripe orientations are perpendicular, the texture will be called the *inconsistent* component (left panel). The overall mean luminance of the patches is the same as that of the background. Notice that the contrast at each point in a compound texture can be considered to be the sum point-by-point of the contrast in two component one-element-only textures. The mean luminance of all the patterns was the same. For some of the experiments in this study, individual grating-patch elements were randomly chosen to be in either sine-phase or minus sign-phase (randomly centered at a positive or at a negative zero-crossing) as in these examples in Fig. 4 (although close scrutiny may be required for the observer to verify that they are indeed of variable phase). In the other experiments, the elements were all in the same phase (not shown here). The models tested here predict no effect of this phase manipulation, and we found none.

ment-only texture have the same orientation as the stripes, the texture will be called the *consistent* component (middle panel). When the element and stripe orientations are perpendicular, the texture will be called the *inconsistent* component (left panel).

On each trial the observer saw a pattern that contained two textures—one texture was a striped arrangement and one a checkerboard arrangement—of the same two elements. One of the textures filled a rectangle that was surrounded by the other texture. Fig. 5 shows an example of a full pattern where one element has contrast zero and the other is shown in caricature to ensure visibility after reproduction. This caricatured example shows an inconsistent component pattern; it is a component pattern because one type of element has zero contrast; it is an "inconsistent pattern" because the horizontal stripes (that form the center rectangle) are made with vertically-oriented Gabor patches; the region outside of the rectangle has a checkerboard arrangement. The observer responded by indicating the overall orientation of the center rectangle (elongated vertically is the correct response in Fig. 5). Percent correct on this task was measured and will be called the *segregatability* of the pattern.

In the experiments reported here, compound patterns were used in which the two components were of equal contrast. The components were also used by themselves. Trials of the compound at six different contrast levels were randomly intermixed with trials of its components at all those levels. For some of the experiments, the Gabor-patch elements were chosen randomly to be in either sine-phase or minus-sine-phase as in Fig. 4. In the other experiments, the elements were all in the same phase. The segregatability of each compound was compared to the segregatability of its components. The next section describes how the measured segregatability of these patterns can distinguish among the structures of second-order channel in Fig. 2.

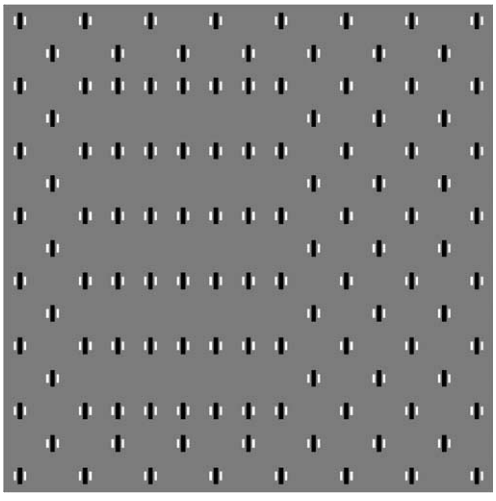

Fig. 5. Sketch of a pattern used in this study. The Gabor-patch elements are shown in caricature here to be easily visible. This example shows an inconsistent component pattern since there is only one element type visible (the other element type has contrast zero), and the striped region of the pattern has horizontal stripes with vertical grating patches (hence inconsistent). There is a checkerboard arrangement on the outside and a striped arrangement on the inside of the embedded rectangle. In the study the rectangle could be in various locations, and the checkerboard arrangement could be inside the rectangle or outside. For segments of striped-arrangement textures that contain more realistic portrays of the grating-patch elements, see Fig. 4.

2. Predictions

2.1. The predictions for the observer

Fig. 6 presents a summary of the models' predictions for the segregatability of a compound pattern and its two components. This subsection briefly summarizes these predictions, and this summary is all that is necessary for understanding the results of this study. For interested readers, more details about the assumptions of the models and the derivation of the predictions in Fig. 6 are given in Section 2.2 and appendix.

The three rows in Fig. 6 show the predictions from three different classes of structures. The top row shows the predictions from SIGN-opponent-only channels (left panel in Fig. 2). The middle row shows the predictions from any of several types of ORIENTATION-opponent channels (middle three panels in Fig. 2). The bottom row shows the predictions from BOTH-opponent channels (right panel in Fig. 2).

The left graph in each row of Fig. 6 shows predicted segregatability thresholds plotted on a *summation square*. The axes give the contrasts in the components of a compound pattern. In Fig. 6 these component contrasts are given relative to the observer's threshold for that component when presented alone. Thus the points $(0,1)$ and $(1,0)$ represent the observer's thresholds for the component patterns (the one-element-only patterns). The black area on each summation square shows the range within which the observer's segregatability thresholds for a compound pattern are predicted to fall. (A pattern is at threshold for an observer when it produces a criterion percent correct.) Points within the light-gray area (closer to the origin than the black area) represent compound gratings predicted to be below threshold. Points further from the origin than the black area (within the white area) represent compounds predicted to be above threshold.

The black area on each summation square includes all predicted thresholds from the range of models (containing channels of the given opponent structure) that seem reasonable on the basis of current knowledge. The straight line boundary of each solid black area is the prediction from the "simplest" model containing the given opponent structure. (See below and the appendix for more details about the full range of models leading to predictions in the black area.)

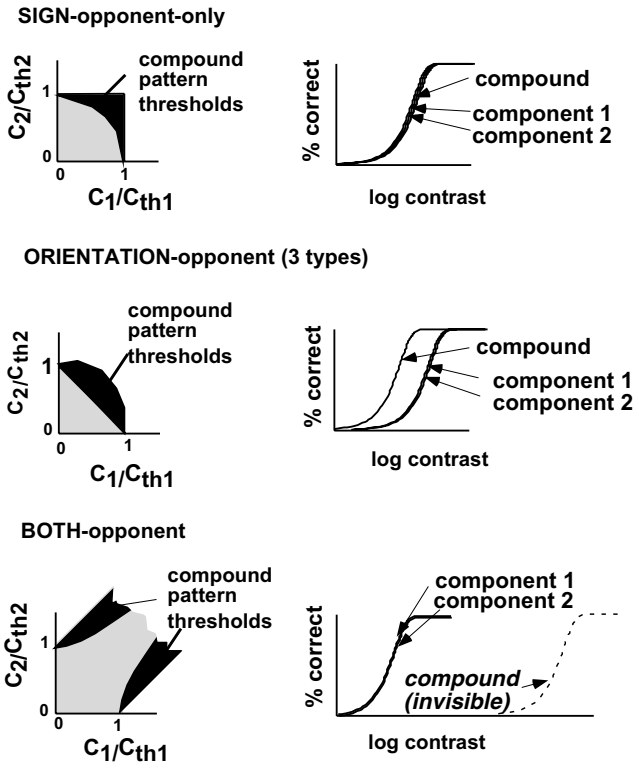

Fig. 6. Predictions for the observer's performance from the different types of opponent-structure. The top row shows the predictions from SIGN-opponent-only channels (left panel in Fig. 2). The middle row shows the predictions from any of the several types of ORIENTATION-opponent channels (Fig. 2, middle three panels). The bottom row shows the predictions from BOTH-opponent channels (right panel in Fig. 2). The left column shows the predictions plotted in summation-square form for thresholds of compound patterns (where the compounds could contain any ratio of contrasts in the two components). The horizontal axis is the contrast in one component of the compound pattern relative to the channels' threshold for that component presented in a one-element-only stimulus. The vertical axis is the relative contrast in the other component. Thresholds are predicted to fall within the solid black area. Points closer to the origin (in the light-gray area) correspond to patterns that are below threshold. Points further from the origin (in the white area outside the black area) correspond to patterns that are above threshold. The straight-edged boundaries on the solid black area are predictions from the "simplest" model containing each structure. (In the bottom summation square for the BOTH-opponent channel, the jagged edges of the black and gray areas indicate that those areas extend indefinitely out toward the upper right.) The solid black areas indicate the range of predictions that can come from a family of "reasonable" models tested. See text and appendix for details. The right column shows the full psychometric function predicted for a compound stimulus containing equally effective components (so the psychometric functions for the two components are the same).

If contrast is plotted in physical units rather than relative to component thresholds, the summation-square plots will not necessarily be exact squares but will be stretched in the horizontal or vertical direction depending on which component the observer is more sensitive to.

The right graph in each row of Fig. 6 shows the full predicted psychometric functions (plotted as percent correct versus log contrast) for the case where the two component patterns' psychometric functions are identical, and the compound contains components at equal contrast.

For the SIGN-opponent-only structure (top row in Fig. 6), the channels that respond to one component do NOT respond to the other component at all. As a consequence, the observer's thresholds are predicted to fall on the right and top outside edges of the summation square or slightly inside. (They will fall slightly inside if, for example, there is probability summation among the channels.) In other words, the compound is predicted to be visible only if at least one component would be visible (or just below threshold) by itself. When the predictions are presented as psychometric functions, the function for the compound pattern will always be much the same as the function for the most sensitive of the two components. (Just as there is a range of possible predicted thresholds, there is a range over which the predicted function for the compound might be slightly shifted to the left relative to the components' functions—but it is harder to show a range of possible positions for a function and we do not do so in Fig. 6.)

For the several ORIENTATION-opponent and the BOTH-opponent structures (unlike the SIGN-opponent structure), all the channels respond to both components (or to neither component). However, ORIENTATION-opponent and BOTH-opponent channels differ in how, when responding to a compound, the second stage combines the component responses. The ORIENTATION-opponent structures *add* the responses (that is, the response to the compound is the sum of the responses to the two components). But the BOTH-opponent structure *subtracts* them.

Thus, for the ORIENTATION-opponent structures (middle row in Fig. 6), the predicted thresholds for the compounds lie on the negative diagonal of the summation square or slightly outside it. (They will fall slightly outside if there is, for example, an expansive nonlinearity at the intermediate stage of the channels.) And the psychometric function for the compound case is predicted to be displaced toward lower contrasts by a factor of 2.0 or a little less (equivalent to 0.3 log units of contrast or somewhat less).

In the BOTH-opponent case (bottom row in Fig. 6), the predicted thresholds lie on lines of slope one going up from the components' individual thresholds or inside those lines. (They lie inside those lines if, for example, there is an expansive nonlinearity at the intermediate stage of the channels.) And any compound in the light gray area inbetween the black ranges is invisible because the difference between the responses to the components is so small. And thus the compound containing equally effective components is never visible so the predicted psychometric function for it (shown in Fig. 6) is moved rightwards toward infinity.

For the interested reader, the full ranges of assumptions used in making these predictions are described in the appendix. Here we will just summarize briefly by saying that the set of models leading to the ranges of predictions in Fig. 6 allows for: (i) the possibility of an intermediate expansive nonlinearity in the second-order channels (Graham & Sutter, 1998); and (ii) the possibility of probability summation across spatial positions and among channels (or equivalent nonlinear pooling) as has been demonstrated with many kinds of compound stimuli, e.g. review in Graham (1989). The possibility of normalization (inhibition among channels) is not fully represented in Fig. 6 but will be discussed later in this paper.

The reader willing to take the predictions in Fig. 6 without further explanation or justification can skip the next section (and the appendix)—going directly to Section 3 with little loss of continuity.

2.2. More about the predictions: channel outputs and the comparison-and-decision stage

The next three figures (Figs. 7–9) illustrate how three opponent structures respond to the component and compound patterns. Here, in this section of the main text, we first describe the general features of these three figures and introduce the assumptions of the comparison-and-decision stage. Following this general discussion of the three figures, we summarize the conclusions to be drawn from each figure individually. (The *appendix*, particularly Appendix A, presents the relevant equations and assumptions more formally.)

2.2.1. Introducing channel outputs and Figs. 7–9

At the top of each of Figs. 7–9 are a pair of sketches—the left-hand sketch shows the kind of receptive-field structure assumed for the channel in that figure, and the right-hand sketch shows the predictions from that structure for the channel's thresholds plotted in a summation square. (The reader might most conveniently read this section while referring to only one of the figures, e.g. Fig. 7.)

Now that it is necessary to be more technical, let us define "channel" explicitly: A channel is a collection of *units* (e.g. neurons), where all units in a channel have receptive fields that have identical characteristics except

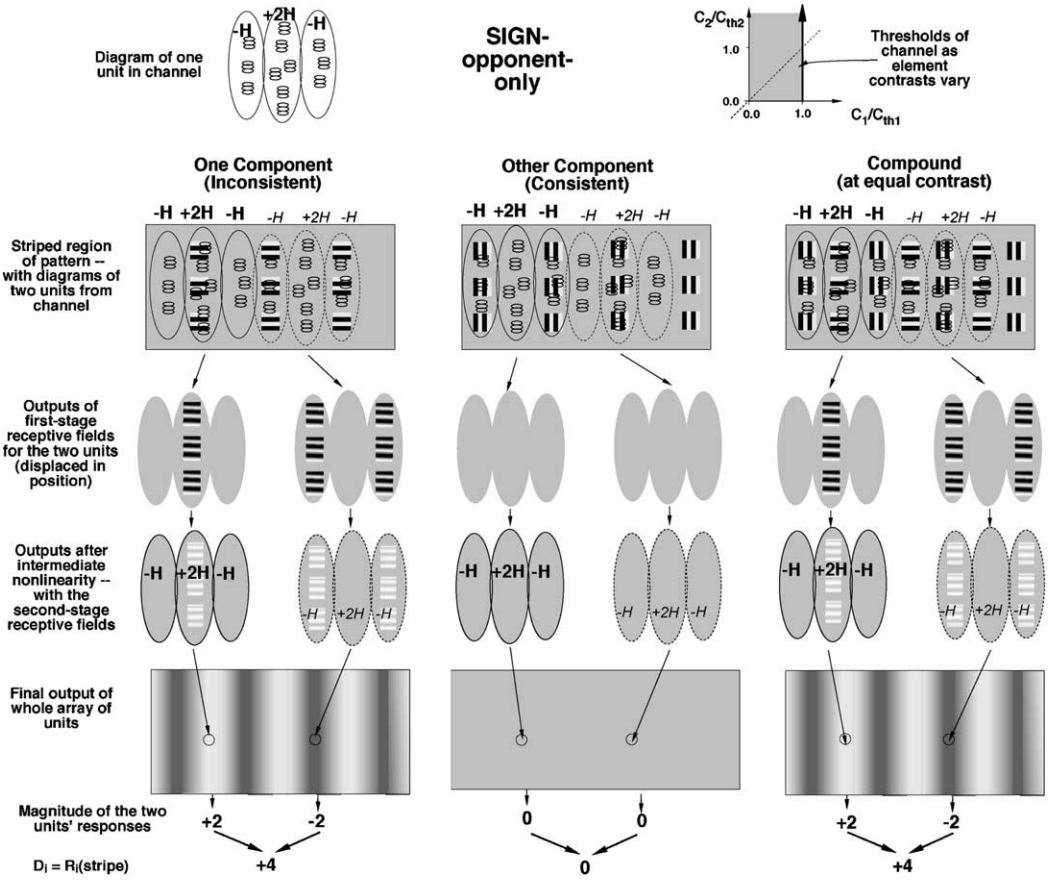

Fig. 7. Illustration of the output of a SIGN-opponent-only channel to the compound and component stimuli used in this study. The labels in the figure itself describe each part of the figure, and there is a fuller description of the figure in the text.

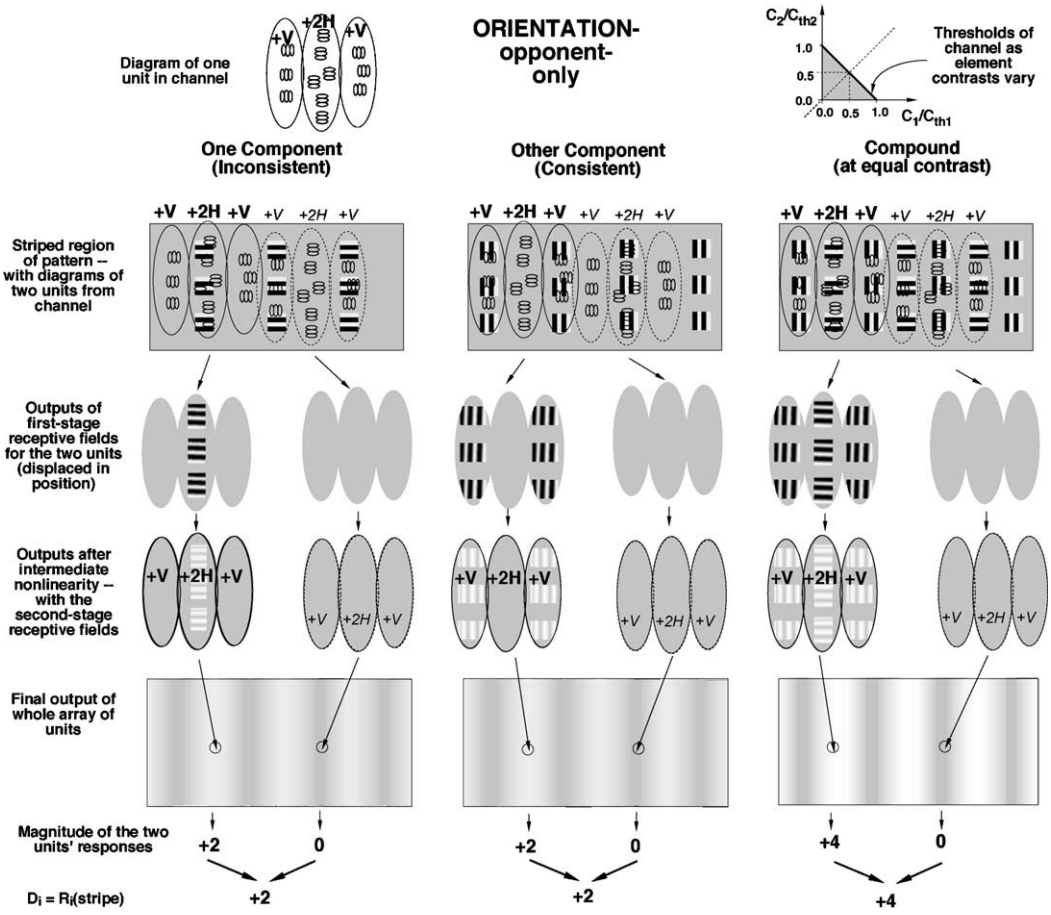

Fig. 8. Same format as Fig. 7 except illustrating the output of an ORIENTATION-opponent-only channel.

that different units are centered at different positions in the visual field.

Below the top pair of sketches in each figure is a row of three gray-level patterns, which illustrate portions of the striped-texture regions from the two component patterns (left and middle columns of the figure) and the compound pattern (right column). (These illustrations are stylized to make them reproduce easily.)

In Figs. 7–9 the phases of all the grating-patch elements are drawn to be identical. In some of the experiments here, however, as mentioned previously, the phases of the elements varied randomly between two phases 180° apart (e.g. examples in Fig. 4). As readers can verify for themselves as they go along, this phase variation does NOT affect any of the predictions in Figs. 7–9 since the grating patches were placed far enough apart that no receptive field of the first-stage filter could respond to two of the patches.

Superimposed on each of these pieces of striped texture are sketches of the receptive fields from two "extreme" units of the full channel. One extreme unit (outlined in black) is centered in the middle of a stripe of horizontal grating-patch elements. (These elements have high contrast in the inconsistent component and in the compound but zero contrast in the consistent compound.) The other extreme unit (dashed outline) is centered in the middle of a stripe of vertical grating-patch elements. (These elements have zero contrast in the inconsistent component but high contrast in the consistent component and the compound.) We refer to these two units as "extreme" units of the channel since their responses are at the extreme maximal or minimal end of the distribution of responses from units in that channel.

The second row of gray-level illustrations in Figs. 7–9 shows outputs from the first-stage receptive fields of the two "extreme" units. The gray level at each point indicates the response of the first-stage receptive field centered at that point: mid-gray corresponds to a response of zero; darker and brighter levels correspond to negative and positive values respectively. (The continuity of gray levels in the figure may seem to imply that the first-stage receptive fields are extremely dense, so dense that there is an infinity of first-stage receptive fields with one centered at every point in the visual field. This is not possible nor is it

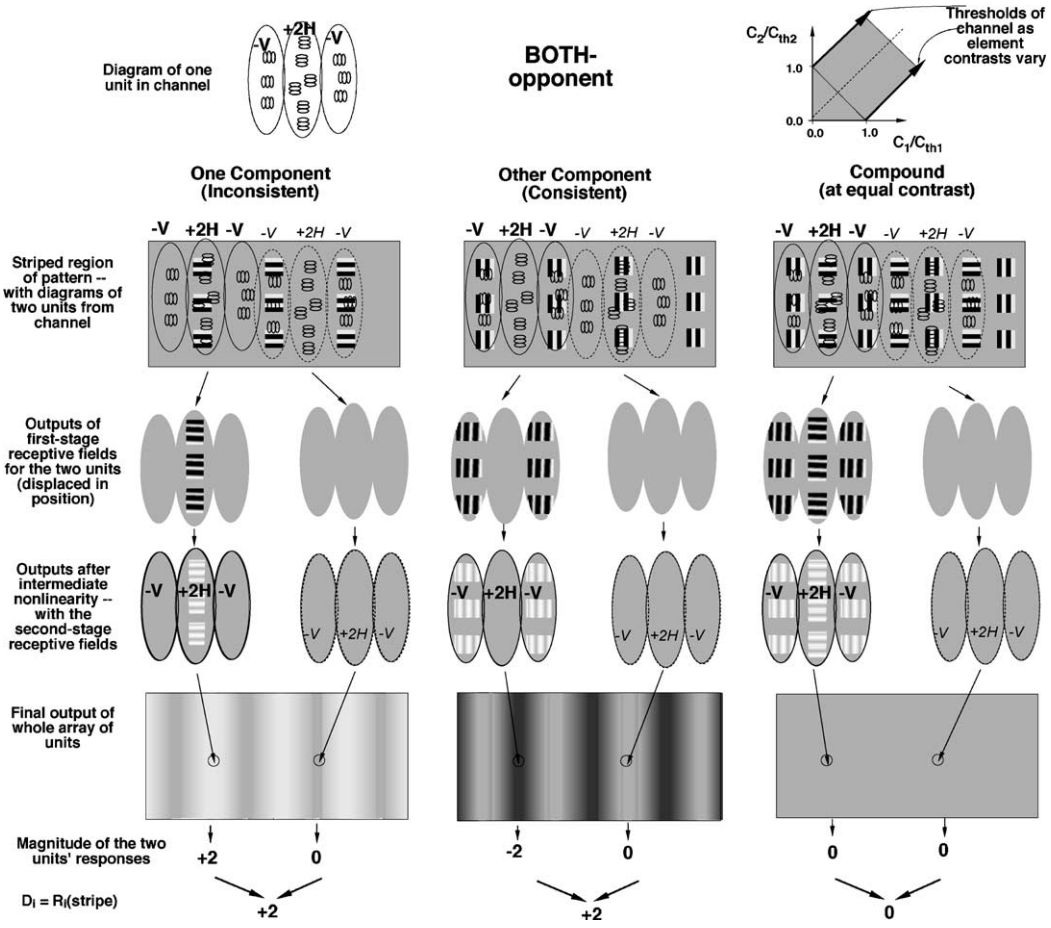

Fig. 9. Same format as Figs. 7 and 8 except illustrating the output of a BOTH-opponent channel.

necessary for the predictions. There only need to be several first-stage receptive fields within each region of the second-stage receptive field for these predictions to apply.)

The third row of gray-level illustrations in Figs. 7–9 shows the first-stage outputs after they have been acted upon by the intermediate rectification nonlinearity (where the superimposed expressions, e.g. +2H, indicate the action yet-to-be-taken by the second-stage of the receptive field). For Figs. 7–9, the intermediate nonlinearity has been assumed to be piecewise linear (that is, neither expansive nor compressive). This corresponds to the model producing predictions on the straight edges of the black areas in Fig. 6. The modifications in the predictions produced by relaxing this assumption are presented in Appendix B.

The bottom row of gray-level illustrations in Figs. 7–9 shows the channel's final output—the output after the second stage of filtering. If there is NO normalization among channels, then the channel outputs shown in the gray level illustrations of the bottom row are also the inputs to the comparison-and-decision stage of Fig. 1. The assumption that there is no normalization was made for Fig. 6's predictions for the observer. The ways in which including normalization changes these predictions will be presented in the discussion section (with further details in Appendix D).

2.2.2. The comparison-and-decision stage (Fig. 1)

The two arrows below the bottom gray outputs in Figs. 7–9 indicate the positions of the responses from the two "extreme" units that were illustrated in the previous rows. These units produce the largest and smallest magnitudes in the full channel output (e.g. +2 and −2 in the left column of Fig. 7). The absolute value of the difference between them (written at the bottom of each column, e.g. +4 in left column of Fig. 7), that is, the *channel's peak-to-trough response to the striped texture*, is called $R_i(stripe)$ in Figs. 7–9 as it is one member of the family of possible rules for "pooling across spatial positions" we use in our models. (The spatial-pooling rule in Figs. 7–9 is simply to find the maximum and minimum and subtract them. We consider a family of possible spatial pooling rules in our work, but the conclusions for this paper do not depend on what member of the family is used. This spatial pooling is shown in Fig. 1

as the first process in the comparison-and-decision stage.)

Only portions of the striped region are shown in Figs. 7–9, but it is easy to summarize the results for the corresponding checkerboard regions verbally without actually drawing and presenting figures. Namely, for the checkerboard region, the response from every unit in the channel (for the channels of Figs. 7–9) is approximately the same. Thus the channel's peak-trough response (or indeed the output of any other spatial-pooling rule we consider) is always approximately zero, that is, $R_i(check)$ equals zero for the channels shown in Figs. 7–9. (To see this more clearly, you could imagine the same kind of figures as Figs. 7–9, but with checkerboard portions of patterns. Then, because of the checkerboard arrangement, both horizontal and vertical Gabor patches would occur in approximately equal numbers within both the center and also within the surround of the second-stage receptive fields. Thus the effects of horizontal and vertical Gabor patches would always cancel out to a very good approximation. Thus the response from each receptive field in the channel would be almost identical and the final outputs of the whole array would be homogeneously gray.)

Since $R_i(check)$ equals zero for the channels in Figs. 7–9, the difference between a channel's responses to the striped and to the checkerboard regions (called D_i in Fig. 1) is just equal to its response to the striped region, that is $D_i = R_i(stripe)$. To make the explanations below less verbose, we will frequently refer to this within-channel difference for a channel as the "channel's response" although it depends not only on the channel's output but on the results of subsequent spatial pooling and within-channel differencing that occurs in the comparison-and-decision stage of Fig. 1.

For the intuitive explanations given in the next few sections, we will consider only the single channel most important for the task. This is the channel tuned to the characteristics of the patterns (first-stage receptive fields matched to the Gabor elements, and second-stage receptive fields matched to the spacing of the rows and columns) like the channels in Figs. 7–9.

To some extent, however, the whole set of channels could affect the response of the observer. In our models we consider a family of possible rules for pooling across multiple channels, that is for computing a decision variable for the observer (represented by D_{obs} in Fig. 1) from all the channels' responses (all the D_i). We then make the simplifying assumption: the greater the pooled value D_{obs}, the better the observer is able to segregate the checkerboard from the striped region in the pattern and therefore, operationally, the greater the observer's performance in the forced-choice experiments of this study. The range of possible effects expected from the existence of multiple channels and pooling across them is indicated in Fig. 6.

See Appendix C.1 for more discussion of the comparison-and-decision stage [1] and for the equations from which were derived the predictions for the results on a summation square that are shown in the upper right sketches in Figs. 7–9 (Appendix A.2), and for the observer in Fig. 6 (Appendices B and C.2).

In the next sections here in the main text, we try to provide insight into these predictions by drawing the reader's attention to aspects of the channel outputs in Figs. 7–9.

2.2.3. Outputs from SIGN-opponent-only channels (Fig. 7)

The SIGN-opponent-only channel (Fig. 7) has a nonzero response to only one of the two components gratings. (Remember that what we mean here by the channel's response is the within-channel difference to the striped versus checked regions, and that this number is identical to the number $R_i(stripe)$ at the bottom of the pairs of columns.) Further, the SIGN-opponent-only channel's response to the compound is identical to its response to that effective component. For example, in Fig. 7, the output to the consistent component (middle column bottom) is uniform and thus the channel response is zero; the full output to the compound (bottom right column gray-level illustration) is identical to the output to the inconsistent component (bottom left), and thus the channel responses to those two patterns are identical also. Thus, for an individual SIGN-opponent-only channel, a compound pattern is at threshold if and only if the contrast in the effective component of the compound pattern is at its own threshold (as shown by the vertical line on the summation square in the upper right of Fig. 7).

[1] *An aside about terminology*. For the purposes of this footnote, it will be useful to let F represent a linear filter and N represent a pointwise nonlinearity. Using these abbreviations, second-order channels are sequences FNF, and first-order channels are just single processes F. One special case of the processes called "Pooling across spatial positions" and "Within-channel differences" in our models (see Fig. 1) could be represented as a point-wise nonlinearity followed by a linear edge-detector filter, that is, as a sequence NF. Thus what Fig. 1 shows as a second-order channel followed by this special case of spatial pooling and within-channel differencing would be a sequence $FNFNF$. Such a sequence might be called a third-order channel. (Note that then the first-order channels shown in Fig. 1 would be part of sequences FNF which might be called second-order channels. See also Landy & Graham, 2003, p. 1111.) However, we think it would be a mistake to speak in these terms here since we feel that the "Pooling across spatial positions" and "Within-channel differences" processes (indeed, all the processes in the comparison-and-decision stage of Fig. 1) are at very different levels of processing from those labeled as "channels". In any case, our conclusions here are based on predictions that depend on assumptions about the possible alternative opponent structures of what is referred to in Fig. 1 as the second-order channels, and the particular details of the processes in the comparison-and-decision stage have little effect on the predictions.

Note that the channel shown in Fig. 7 corresponds to the bottom left sketch in Fig. 3. There is an analogous SIGN-opponent-only channel (shown in the top left of Fig. 3 and left of Fig. 2) that responds to the other component (the consistent component). For lack of a better term we refer to this other channel as the *fraternal twin* of the channel in Fig. 7.

The other two cases of SIGN-opponent-only receptive fields in Fig. 3 (the two in the right column) do not respond to the element-arrangement patterns (vertical stripes) shown in Fig. 7 at all and so do not need to be considered when discussing those patterns.

For an observer whose response is determined by only these SIGN-opponent-only channels, the threshold for the compound will be equal to the lowest threshold from the pair of channels responding to the two components (e.g. the channel in Fig. 7 and its fraternal twin) thus producing thresholds near both the right edge and top edge of the summation square (Fig. 6 top row). (Remember that the range of observer thresholds—the black area in Fig. 6 top row—extends inside the summation square due to other factors, e.g. possible probability summation among channels. See Appendices C.1.1 and C.1.2. Also, in Appendix C.1.3, the question of possible intrusions into the observer's decision from other channels with different spatial-frequency and orientation sensitivities is discussed.)

A note about possible single-region second-stage receptive fields. A reader may have noticed that the predicted channel outputs in Fig. 7 do not depend strongly on the inhibitory surround of the second-stage receptive field. In particular, imagine that the second-stage inhibitory surround were removed leaving a second-stage receptive field with only an excitatory center. Then the bottom row of channel outputs would look the same as in Fig. 7: the output to the compound would still be identical to the output to the most-effective component, and the output to the other component would still be uniform. (The exact magnitudes would be different: In particular, the pairs of numbers at the bottom of the three drawings would be (+2, 0) (0, 0) (+2, 0).) Thus this SINGLE-REGION structure would, in fact, lead to the same predictions for the observer as does the SIGN-opponent-only structure. This SINGLE-REGION channel structure is not considered further in this paper, however, as it has already been rejected by previous evidence that the second-stage filter is bandpass on the spatial-frequency dimension, e.g. Landy and Oruc's (2002) second-order summation experiments and Ellemberg, Hess, and Allen's (2004) simultaneous detection and identification experiments.

2.2.4. Outputs from ORIENTATION-opponent channels (Fig. 8)

Of the three types of ORIENTATION-opponent structures, only one is illustrated in these explanatory figures—namely the simplest one, the ORIENTATION-opponent-only structure—and it is shown in Fig. 8. The analogous illustrations for the other two types did not seem worth the space to present. While somewhat tedious to do, interested readers could sketch the analogous ones for themselves relatively straightforwardly, and could verify that the conclusions illustrated in Fig. 8 for the ORIENTATION-opponent-only case also apply to the DOUBLY-opponent and HALF-DOUBLY-opponent cases.

For an ORIENTATION-opponent-only channel, the output to the compound (Fig. 8 lower right) shows more modulation than to either component alone (Fig. 8 lower left and middle) because of the summing of center and surround areas by the second-stage receptive field. Indeed this kind of channel's response to the compound is the *sum* of its responses to the components separately. Thus the compound's thresholds for the individual channel plotted on a summation square (Fig. 8 upper right) lie on the negative diagonal which shows linear summation.

The range of observer's thresholds for this case (Fig. 6 middle row) extends out from the negative diagonal due to several factors, including the possible nonlinearity at the intermediate stage of the channel (see Appendix B).

Note that, for the ORIENTATION-opponent-only case we have just been discussing and also for the BOTH-opponent case (next section), there is no second channel structure (no "fraternal twin") that needs to be considered. (The other arrangements for each ORIENTATION-opponent-only and BOTH-opponent channels—analogous to the other arrangements in Fig. 3 for SIGN-opponent-only—either do not respond at all to the patterns in Figs. 8 and 9, or they respond just as the channel shown does and therefore do not need to be considered separately.)

2.2.5. Outputs from BOTH-opponent channels (Fig. 9)

A third distinct kind of behavior is shown by the third structure, the BOTH-opponent structure. Here the output to the compound (Fig. 9 lower right) shows less modulation than to either component because of the differencing of center and surround areas by the second-stage receptive field. Indeed this kind of channel's response to the compound is the *difference* of its responses to the components separately. Thus the compound's thresholds for the individual channel plotted on a summation square (Fig. 9 upper right) lie on the lines showing complete cancellation between components. The range of observer's thresholds for this case (Fig. 6 bottom row) extends inward from these lines due to several factors, including the possible nonlinearity at the intermediate stage of the channel (see Appendix B).

2.2.6. Final note on channel outputs

Remember that the examples in Figs. 7–9 have been worked through for the most unadorned version of a

channel of the given type. In particular, in these figures: each channel has an intermediate nonlinearity that is ordinary rectification (piecewise linear); there is no probability summation or equivalent nonlinear pooling among channels as we are looking only at a single channel; and there is no inhibition among channels of the normalization type (or any other kind of intensive nonlinearity). This simple unadorned model leads to the straight line predictions in Figs. 6–9. The general qualities apparent in Figs. 7–9 hold approximately over a wider range of assumptions, and the black regions in the summation squares of Fig. 6 allow for more complicated models as described briefly in the description of Fig. 6 above. The effects of allowing for normalization in an inhibitory network are not represented in Fig. 6. But, the effects are described in Section 4 as they turn out to be of substantial interest in interpreting the results of the experiments.

3. Experimental methods and procedures

The observer's performance on compound and component patterns was measured. Each pattern contained two regions: a rectangle region embedded in a background region (a caricatured example is shown in Fig. 5). One region contained a checkerboard arrangement and the other region contained a striped arrangement of the two element types: vertical and horizontal grating patches. The striped portions of one compound and its components are illustrated in Fig. 4.

3.1. Element details

Each element was a Gabor patch truncated to lie within a square of width 64 pixels so that neighboring elements did not overlap (64 pixels subtends 1° at a viewing distance of 86 cm, and 2° at 43 cm). The period of the sinusoid in each Gabor patch was 8 pixels, so the spatial frequency was 1/8 cycle per pixel, which was 8 c/deg at a viewing distance of 86 cm and 4 c/deg at 43 cm; as mentioned above, the grating-patch orientation was either vertical or horizontal. The full width at half height of the circular Gaussian envelope of each Gabor patch was 16 pixels (0.25° at a viewing distance of 86 cm, 0.5° at 43 cm), which equals two periods of the sinusoid. The center-to-center distance between the Gabor patches was 64 pixels, which equals eight periods of the sinusoid.

Two phase conditions were investigated. In the *constant-phase condition* the phase was the same in every Gabor patch; in particular, the harmonic oscillation was in positive-sine-phase with respect to the window. In the *random-phase condition*, the phase in each element was randomly chosen from two possibilities: the positive-sine-phase or negative-sine-phase. Note that, for these phases, the space-average luminance of the elements was always the same as the background luminance. For more details see Graham and Wolfson (2001).

3.2. Details of the arrangement of elements

Each pattern was a grid of 15×15 Gabor-patch elements with two regions: a rectangle region of 7×11 elements (or 11×7 elements) embedded in a background region. The full pattern subtended either 15° (at a viewing distance of 86 cm) or 30° (at a viewing distance of 43 cm). When vertically elongated, the rectangle could occur randomly in any of three overlapping positions: left of center, center, or right of center. When horizontally elongated, it could occur randomly in any of three overlapping positions: top, center, or bottom. The random assignment of rectangle position diminishes the probability that the observer can do the task by attending to a very few elements at a fixed location.

In half the patterns the rectangle contained the checkerboard arrangement of elements, and the background contained the striped arrangement (vertical or horizontal stripes with equal probability). In the other half of the patterns, the rectangle contained the striped arrangement (horizontal or vertical with equal probability) and the background contained the checkerboard arrangement.

The period at which either the checkerboard or striped arrangement repeated itself is 128 pixels (two elements). Thus the fundamental frequency of the element-arrangement patterns was 1/128 cycles per pixel (0.5 c/deg at a viewing distance of 86 cm and 0.25 c/deg at a viewing distance of 43 cm).

These element-arrangement patterns are analogous to the modulated-noise patterns used by a number of other investigators. The frequency and orientation in the Gabor patches here are analogous to the carrier frequency and orientation. And the fundamental frequency here is analogous to the modulation frequency.

3.3. Details of the equipment and mean luminance

Stimuli were presented on an Apple 17 in. ColorSync monitor (75 Hz refresh rate, 1280×1024 resolution) controlled by a Power Mac G3. The mean luminance of our patterns was approximately 35 cd/m^2. Stimuli were generated and presented using MathWorks' MATLAB with the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). The monitor's lookup-table was linearized.

3.4. Details of each trial

The subject's task in the experiment was to indicate the elongated orientation of the embedded rectangle.

To begin each trial, the observer pressed the "0" key (on the numeric keypad) which presented a fixation point (a low-contrast 20×20 pixel square) for 500 ms followed by a screen that was uniform at the mean luminance for 500 ms. Then the stimulus appeared for 100 ms, with abrupt onset and offset, followed by a uniform screen until the observer responded. The observer was forced to wait 1 s after the stimulus terminated before responding (the computer beeped to indicate when the observer could respond), a procedure we initiated to make sure that observers waited for appropriate processing before responding (Graham et al., 1993) which is particularly important when second-order channels are involved as evidence suggests they are rather slow (Sutter & Graham, 1995; Sutter & Hwang, 1999, but also see Motoyoshi & Nishida, 2001). The observers then pushed either the "8" key on the numeric keypad to indicate the rectangle was vertical or the "4" key to indicate the rectangle was horizontal. A high or low-pitched tone provided feedback as to the correctness of the response.

3.5. Structure of the sessions

Each session consisted of 864 trials, 144 at each of six contrasts (6%, 10%, 18%, 31%, 54%, or 94%). All patterns in a given session were either of constant phase, or all patterns in a given session were of random phase.

Of the 144 trials at each given contrast: one-third (48 trials) presented a compound pattern in which both element types were at that given contrast; one-third presented the inconsistent component in which the elements having nonzero contrast were at the given contrast; and one-third presented the other component (again the elements having nonzero contrast were at the given contrast).

Each third (48 trials) consisted of all combinations of: (i) the positions and elongation of the rectangle (six possibilities), (ii) the characteristics of the arrangements (checkerboard inside the rectangle and striped outside or vice versa—with either vertical or horizontal stripes—thus four possibilities), and (iii) whether the upper left element position was a vertical or horizontal Gabor patch in the compound patterns, or whether the upper left element position was blank or filled in the component patterns (two possibilities).

All four observer completed five sessions with the constant-phase patterns. Three of the observers (CC, JR, and MK) completed five sessions with the random-phase patterns. The combination of subject and phase condition will be referred to here as an "experiment". Thus there were seven experiments.

Observers ran in a dimly lit room. For observers CC and MK the viewing distance, with unrestrained head, was approximately 43 cm, and for observers AF and JR it was approximately 86 cm.

3.6. Subjects

All observers were paid undergraduates with previous experience in texture segregation experiments. Some previous results from CC, MK, and AF are reported in Graham and Wolfson (2001). Observers had normal or corrected-to-normal visual acuity.

3.7. Threshold calculation

For each pattern in each experiment, the observers results were pooled across incidental conditions (e.g. whether it was stripes or checkerboard inside the rectangle) and pooled across sessions to produce the psychometric functions from which the thresholds were then calculated.

3.7.1. Consistent versus inconsistent component analysis

For the results presented in Figs. 10 and 11, we considered component 1 to be the consistent one-element-only pattern and component 2 to be the inconsistent one-element-only pattern. Thus for the results we will present in these figures, we grouped trials so that those presenting stimuli having consistent structure in the stripes were grouped together, and those with inconsistent structure were grouped together. In more detail, note that (in this analysis) the patterns containing component 1 either had vertical stripes composed of vertical grating patches or horizontal stripes composed of horizontal grating patches; patterns containing component 2 either had vertical stripes composed of horizontal grating patches or horizontal stripes composed of vertical grating patches.

3.7.2. Vertical versus horizontal grating-patch-component analysis

Although we will not show them in figures, we also computed the results when we considered component 1 to be the one-element-only patterns containing horizontal grating patches and component 2 to be the one-element-only patterns containing vertical grating patches. That is, for this alternate analysis, we grouped trials so that those presenting stimuli containing vertical grating elements were grouped together, and those presenting stimuli containing horizontal grating elements were grouped together. In other words, for this alternate analysis, patterns containing component 1 either had vertical stripes containing vertical grating patches or had horizontal stripes containing vertical grating patches. Patterns containing component 2 either had vertical stripes containing horizontal grating patches or had vertical stripes containing horizontal grating patches. These results using vertical versus horizontal grating-patch components will be discussed below although not shown in figures.

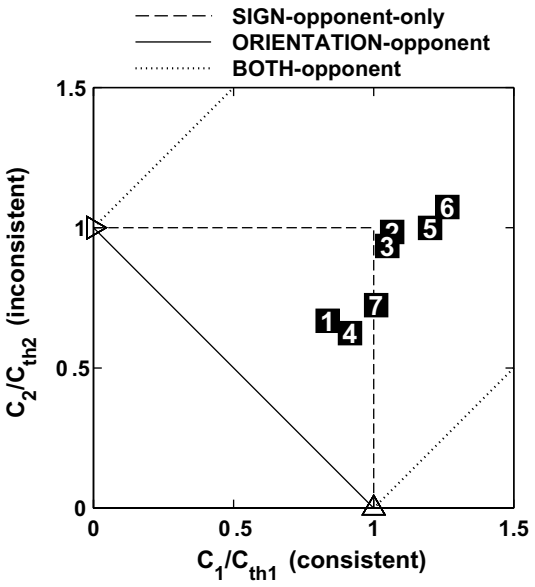

Fig. 10. Contrast thresholds from all seven experiments plotted on a summation square. Component contrast is plotted on each axis relative to the threshold for that component in a one-element-only stimulus. The horizontal axis shows the relative contrast in component 1 (the consistent component). The vertical axis shows the relative contrast in component 2 (the inconsistent component). The open triangles indicate the thresholds for the two one-element-only component stimuli used in each experiment. They plot at (0, 1) and (1, 0) because relative contrast is shown on the axes. The closed squares show the threshold for the one compound studied in each experiment (the compound in which the physical contrasts of the two components was equal). The number in each closed square indicates the observer and condition, see Fig. 11. See that figure legend for more detail. Threshold was calculated as the contrast level leading to 70% correct. This plot also shows the predictions from the simplest forms of models (see Fig. 6) containing SIGN-opponent-only channels (top and right edges of the summation square), ORIENTATION-opponent channels (the negative diagonal line), or BOTH-opponent channels (the pair of positively-sloped lines). The experimental results are closest to the predictions of the SIGN-opponent-only model.

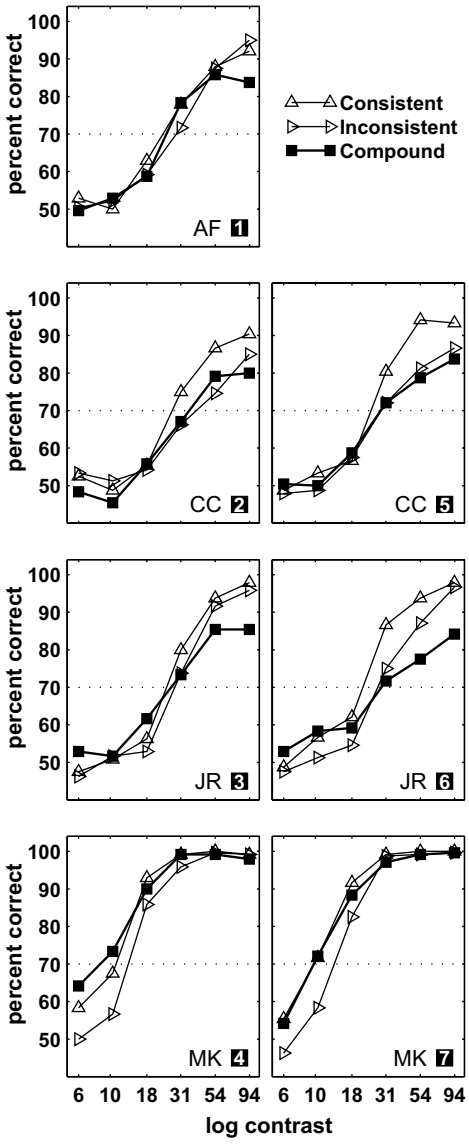

Fig. 11. Psychometric functions from the seven experiments. Percent correct is shown on the vertical axis and physical contrast (expressed in percent) on the horizontal axis for the seven experiments. The dotted horizontal line at 70% correct shows the criterion used to compute the thresholds plotted in Fig. 10. Each data point is the percent correct over 240 trials. Results for the compound stimuli are plotted as squares; those for the consistent component as up-pointing triangles; and those for the inconsistent component as right-pointing triangles. For experiments 1–4 (left column), the phase of all Gabor patches was the same odd-symmetric phase (the constant-phase condition). For experiments 5–7 (right column) the phases were a random mixture of the two 180° apart odd-symmetric phases (the random-phase condition). Observers CC and MK were at a viewing distance of 43 cms so that spatial frequency of the elements was 4 c/deg, and the fundamental frequency of the arrangement was 0.25 c/deg. Observers AF and JR were at a viewing distance of 86 cms so that the spatial frequency of the elements was 8 c/deg, and the fundamental frequency of the arrangement was 0.5 c/deg. The psychometric function for the compound is approximately the same as that for the most effective component, in line with the predictions in Fig. 6 from the SIGN-opponent structure. There is a small discrepancy at the high contrasts where, in many of the panels, the segregatability of the compound is less than that of the most segregable component.

3.7.3. Fitting psychometric functions

A Quick (Weibull) function was fit to each psychometric function in order to allow us to interpolate and find the threshold value—that is, to find the contrast level which corresponded to the criterion performance level (70%). We used the values interpolated from those fitted functions for the thresholds (e.g. those in Fig. 10 for consistent and inconsistent components).

3.7.4. Inside versus outside the rectangle

Although not shown here, we also analyzed separately the stimuli where the stripes were in the rectangle (and the checkerboard outside) versus those where the stripes were outside. The results do not differ and thus results averaged over both those conditions are plotted here.

4. Results and discussion

Fig. 10 shows the measured observer thresholds for the seven individual experiments. The thresholds are plotted on a summation square with relative contrast values on the axes so the component thresholds (triangles) fall at (0,1) and (1,0) for all seven experiments.

In this plot the compound thresholds (square symbols) fall near the top and right edges of the summation square which are the predictions from the SIGN-opponent-only channels. The compound thresholds are all outside the predicted range from the ORIENTATION-opponent channels (the negative diagonal line and the adjoining black region in Fig. 6 middle row), and they are nowhere near the predictions of the BOTH-opponent channels (the pair of positively-sloped lines and the adjoining black region in Fig. 6 bottom row).

Similarly, in Fig. 11 showing the full psychometric functions for the seven individual experiments, the psychometric function for the compound stimulus (squares) is approximately juxtaposed with the psychometric functions of the individual components as predicted by SIGN-opponent-only channels; the compound's function is certainly not far to the left of the components' functions as predicted by ORIENTATION-opponent channels nor far to the right as predicted by the BOTH-opponent channels. If the two components are not exactly equal in segregatability, the prediction of the SIGN-opponent-only channels is that the compound is approximately as segregable as the most segregable component. To a good first approximation, the empirical functions are consistent with this prediction of the SIGN-opponent-only channels.

Looking at the Fig. 11 more closely, however, shows that in detail the psychometric functions are not completely consistent with the Fig. 6 predictions of the SIGN-opponent only channels. In particular, consider the difference between the results at low and high contrast levels. At low contrast levels, performance on the compound (solid squares) may well be identical (except for variability) to that on the more effective component (whichever of the open triangles is higher) as predicted by the SIGN-opponent-only channels. However, at higher contrast levels, the psychometric function for the compound stimulus tends to veer away from its upward course so that the segregatability of the compound becomes distinctly *less* than that for the most segregable component. This is clear in five of the seven experiments while not being true for the two experiments using observer MK.

We had seen hints of this effect in some unpublished rating-scale results which were an extension of some published studies (Graham et al., 1993; Graham, Sutter, Venkatesan, & Humaran, 1992). However, we had not considered those results seriously because the subjectivity of rating-scale methods makes it easy to believe that the observers introduce various kinds of distortion that do not directly represent the perception they are being asked to rate.

We will return to possible explanations of the performance at high contrast levels after briefly looking at some subsidiary results that are not important for the overall argument of the paper but which may interest some readers.

4.1. Some subsidiary results

4.1.1. No effect of phase condition

Note that the results described above hold both for the constant-phase condition (left column in Fig. 11) and the random-phase condition (right column in Fig. 11). More generally the results for these two phase conditions are very similar. This replicates the previous result of Graham and Wolfson (2001) and is what would be expected if the grating-patch elements in our patterns are far enough apart that no individual first-stage receptive field responds to more than one grating-patch element.

4.1.2. Bias toward consistent textures

Also, the previously-reported bias (Graham & Wolfson, 2001) in favor of consistent over inconsistent one-element-only patterns is seen here: in the psychometric functions of Fig. 11, the upward-pointing triangles are generally higher than the rightward-pointing triangles—this is true for three of the four Ss, for six of the seven experiments. Subject AF who only participated in the same-phase condition, did not show this bias.

The existence of a small difference between the sensitivities of the channels to the consistent and inconsistent condition (or the vertical and horizontal grating elements, next paragraph) can be assimilated easily into the model framework above and does not substantially change the predictions although we presented the predictions for cases without a bias.

4.1.3. Horizontal grating-patch and vertical grating-patch components

In addition to using consistent versus inconsistent arrangements of elements as the two components (as done in Figs. 10 and 11), we analyzed the results using the vertical-grating-patch versus the horizontal-grating patch elements as the two components. Although of course the exact percent corrects were slightly different, the analysis led to the same conclusions about opponency as those from (Figs. 10 and 11). In particular, the results were still best described by the models containing SIGN-opponent-only rather than ORIENTATION-opponent or BOTH-opponent channels. And, further, there was still the decrease in performance on the compound relative to the components at high contrasts.

When analyzed this way, the results again show a slight difference in threshold between the two components but now this difference reveals a slight bias toward one orientation of grating patch compared to the other (rather than a bias toward consistent over inconsistent arrangements). In particular, for two observers the performance on horizontal grating-patch elements was slightly better than performance on vertical grating-patch elements; for one observer there was a small difference in the opposite direction; and for the fourth observer there was no discernible difference.

4.2. Normalization and the results at high contrast levels

Let us return to the slight discrepancy between the predictions of the SIGN-opponent-only channels model in Fig. 6 top row and the measured performances at high contrasts: for three of the subjects (AF, JR, and CC—five of the seven panels in Fig. 11), the segregatability of the compound at high contrasts (solid squares) was not as great as that of the most effective component (whichever of the open triangles is higher). Thus, at high contrasts, the observer's performance on the compound is below that predicted by the SIGN-opponent-only structure in Fig. 6 (although still agreeing better with that prediction than with the predictions from the other structures).

This decreased segregatability of the compound (relative to its components) at high contrast levels might result from some form of inhibition between SIGN-opponent-only channels sensitive to the two components; but it would need to be a kind of inhibition that is nonlinear and therefore does not occur at all contrast levels. [2] We and others have been led to invoke inhibition among channels in a normalization network to explain other phenomena in pattern perception.

4.2.1. About normalization

At least two different kinds of nonlinearities—one *intensive* in character and one more intrinsically *spatial*—have been useful in accounting for texture segregation and similar perceptual phenomena (e.g. Graham, 1991; Graham et al., 1992; Malik & Perona, 1990; Sperling, 1989; Wilson, 1993). The second-order channel structure itself embodies the spatial nonlinearity. But to study second-order channels properly, possible intensive nonlinearities need also to be considered. Some of these are *early local nonlinear transformations* depending on luminance (e.g. retinal light adaptation depending on average light level) or on contrast (early local contrast-gain-controls), but these are not very important in results using patterns like these (e.g. Graham & Sutter, 2000). Of possible significance here is the (nonlinear) inhibition among orientation- and spatial-frequency-selective channels that operates as a more global contrast-gain control and which has been modeled with success as a normalization network.

This normalization process has been shown to account for a number of characteristics of perceived segregation of textures like those we study here (e.g. Graham & Sutter, 1996, 2000) as well as characteristics measured in other kinds of perceptual tasks, e.g. suprathreshold discriminations of sine-wave gratings (e.g. Foley, 1994; Olzak & Thomas, 1999; Itti, Koch, & Braun, 2000; Teo & Heeger, 1994; Thomas & Olzak, 1997; Watson & Solomon, 1997).

Further, something is known about why the visual system might have evolved such a process: It may prevent overload on higher levels by repositioning the limited dynamic range to be centered near the ambient contrast level and at the same time preserve selectivity along dimensions like orientation and spatial frequency. (See discussions and references in, e.g., Bonds, 1993; Heeger, 1991; Lennie, 1998; Victor, Conte, & Purpura, 1997.) It has also been suggested that such normalization has the right properties to help encode natural images efficiently (Schwartz & Simoncelli, 2001; Simoncelli & Olshausen, 2001; Zetzsche, Krieger, Schill, & Treutwein, 1998).

4.2.2. Predictions from normalization for this study

Given the previous successes of explanations based on inhibition embodied in a normalization network, it seemed worthwhile to do the work necessary to add normalization to calculations of predictions that we had done for this study in order to see if normalization might explain the decreased segregatability at high contrast found here (Fig. 11). Thus we systematically explored the effect of adding normalization to the model. Here in the main text we briefly summarize our calculations and the predictions, and further information including equations forms Appendix D.

First, adding normalization does *not* change the predictions of models with either BOTH-opponent or ORIENTATION-opponent channels enough to make the predictions consistent with the results reported here

[2] There is an alternate explanation for this discrepancy that might occur to some readers and this footnote briefly explains why we do not think this alternate possibility likely. Suppose the first-stage receptive fields have broad enough bandwidth that they are somewhat sensitive to the orientation that is perpendicular to their best orientation. The channel would then show less response to the compound pattern than to the most effective component because the perpendicular elements would act like elements of the best orientation but at lower contrast thus reducing the modulation in the channel's output to the compound. (This is the explanation that Graham et al., 1993, used, in fact, for explaining their results with closer orientations.) However, this explanation predicts that effect would occur at all contrasts of the compound and components. Thus it is ruled out by the results here at low contrasts which do not show this effect although the results at high contrasts do.

(and this is true for any value of threshold criterion, that is, at any contrast level).

Second, adding normalization to the model with SIGN-opponent-only channels largely leaves the predictions much the way they are shown in Fig. 6. However, it does more than that: since the predictions of the model with normalization do depend on the threshold criterion, normalization does distort the shapes of the predicted psychometric functions to some extent, and, as it turns out, it can account for the performances at high contrast levels.

Some sample predictions of the model containing SIGN-opponent-only channels and normalization are shown in Fig. 12. The vertical axis shows the predicted value from the model of D_{obs}, the decision variable for the observer, a value which is assumed to be monotonic with percent correct as measured in the experiment. The horizontal axis shows contrast in arbitrary units. The top panel shows results when the two components are of approximately equal detectability (as was the case for observer AF and, to a lesser extent, JR). The bottom panel shows a case where those sensitivities are quite unequal (as for observer CC). Note that the predictions in Fig. 12 show the features of our results: in particular the performance on the compound at high contrasts drops below the performance on the most effective component.

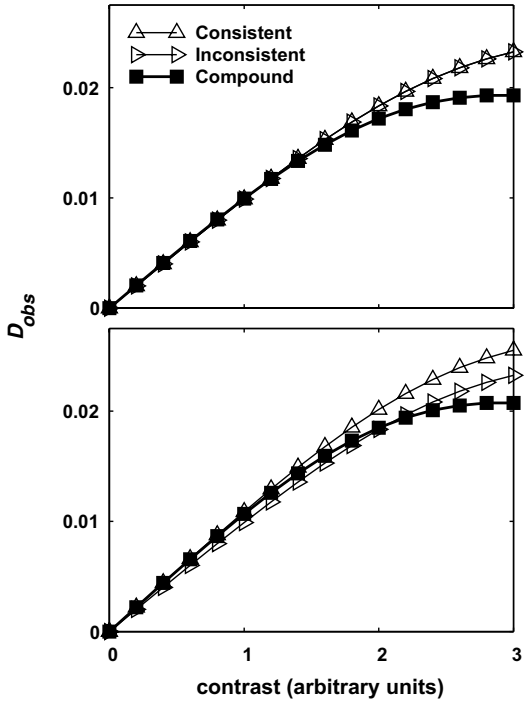

Fig. 12. Some predicted psychometric functions when a contrast-gain-control (inhibition among channels in a normalization network) is included along with SIGN-opponent-only channels in the model. The horizontal axis shows contrast in arbitrary units. The vertical axis shows the Channel Response D_{obs} in arbitrary units. The top panel shows a case where the observer's sensitivities to the two component stimuli are assumed equal. The bottom panel shows a case where those sensitivities are unequal. (See Appendix D for equations and the parameter values used to generate these predictions.) These psychometric functions show the important features of the observer's results in Fig. 11.

To see why the model with SIGN-opponent-only channels and normalization makes this prediction consider the following argument. Consider "other" channels that respond substantially to individual elements in the texture patterns but do not respond differentially to the checkerboard versus striped arrangements and thus are not able to segregate the textures. These "other" channels may include both simple linear channels and second-order channels. (See Graham & Sutter, 2000, and Appendix D.1.2 here for more discussion of possible "other" channels.) Although these "other" channels cannot segregate the striped versus checkerboard textures, they contribute to the normalization pool and thus to inhibition of the channels that can segregate the textures. More than that, they reduce the response of the channels able to segregate more for the compound stimulus than for its components for the following reason: In response to the compound stimulus, there are two distinct sets of these "other" channels; one set is responsive to one component and the other to the other component. In response to a component stimulus, however, there is only one set of "other" channels (those sensitive to the one component present). Thus there is more inhibition introduced for the compound than for the component pattern. Further, since normalization is a nonlinear process, the amount of reduction the "other" channels produce is not linear with contrast. If one considers the whole parameter space for the models containing normalization with SIGN-opponent-only channels, one can find predictions either of the result shown in Fig. 12 or the opposite (where the compound gets relatively more detectable at high contrasts). However, parameter values that seem reasonable on the basis of previous work (see Appendix D.3) produce predictions like those in Fig. 12, that is, predictions that are consistent with our empirical results.

Why one observer (MK) should seem immune from this effect of normalization is not clear, however. No observer has ever failed to show, for example, the signature of normalization in constant-difference-series experiments (Graham & Sutter, 1996, 2000). The effect here is, however, a subtler effect, and perhaps this observer's normalization parameters are not quite strong enough to produce the effect here.

We are not be surprised, in general, that there are differences among the four observers in the magnitude of this effect. Individual differences have occurred in many suprathreshold experiments (e.g. Cannon & Fullenkamp, 1993; Graham & Sutter, 1998, 2000; Morgan & Dresp, 1995) and given the possibilities open to observers

in suprathreshold experiments, such differences seem very likely to occur. (see Graham, 1989, p. 12).

4.3. Previous studies of opponency in texture channels

A model including SIGN-opponent-only (complex, filter-rectify-filter) channels and inhibition among these channels (in a normalization network) can explain our results here using textures composed of horizontal and vertical grating patches. Models having opponency (of any of several types producing quite different predictions) between perpendicular orientations cannot.

Our conclusion here is consistent with several recent studies which have rejected various types of opponency between perpendicular orientations and explained their results in terms of SIGN-opponent-only channels, often with some process like normalization included (Arsenault et al., 1999; Kingdom et al., 2003; Motoyoshi & Kingdom, 2003; Prins et al., 2003).

In the past, however, a number of studies (e.g. Bergen & Landy, 1991; Gray & Regan, 1998; Kingdom & Keeble, 1996, 2000; Rubenstein & Sagi, 1993) have been interpreted as suggesting more complicated opponency structures than the SIGN-opponent-only structure. Not only are such complications not necessary to explain the results here, but no channels of these forms having any substantial sensitivity to these patterns could exist without having intruded on the results here.

Are there tasks or stimuli for which these other more complicated channel types would be important? Certainly it is possible that for other tasks (e.g. visual search, object identification, etc.) and for other stimuli (perhaps containing more complicated structures) the observer's performance might reflect higher-order processes than the channels revealed here.

However, we think it rather likely that at least some of the previous results suggesting more complicated channel types are in fact consistent with the simpler SIGN-opponent-only structure especially when normalization is taken into account. To prove that this is so, however, takes substantial computational effort, and it is not clear that it would be possible to reach a compelling conclusion for the earlier studies in any case. One of the reasons so much work would be involved is that—for many of the patterns used—a large number of SIGN-opponent-only channels that are sensitive to various spatial frequencies and orientations will respond to the patterns, and the particular subpopulation responding will differ from stimulus to stimulus (leading, e.g. to "off-frequency, off-orientation looking"). It is exactly these factors which Prins and Kingdom (2002, 2003) and Kingdom et al. (2003) considered and incorporated in some computations which led them to conclude in favor of filter-rectify-filter mechanisms (SIGN-opponent-only) for these recent studies. These factors are minimal for the study here (for further details see Appendix C.1.3).

It is also possible that interactions between orientations closer than 90° might be different than those at 90°. Motoyoshi and Kingdom (2003) have presented some evidence of this. Their experimental paradigm is elegant and, in fact, does a better job of disentangling bandwidth effects at the first-stage receptive field from effects at the second stage than does our paradigm here. They interpret their results at perpendicular orientations as showing no opponency, but their results at much closer orientations as showing opponency of the BOTH-opponent kind (interference rather than summation). This may well be true. Again we wonder, however, if perhaps normalization (inhibition among channels) with SIGN-opponent channels could account for their results at both close and far orientations. As it happens BOTH-opponency mimics normalization to some extent (or vice versa) in that it predicts an interference between the two orientations. Perhaps in their study, the normalization pool was stronger for the close-orientation condition than for the far due to the exact contrasts involved and the construction of the normalization pool. Then one would see interference at close but not at far orientations as they did.

5. Summary and conclusion

SIGN-opponent channels ("complex channels", "filter-rectify-filter" mechanisms, see Fig. 2) lead to the correct prediction that the segregatability of compound element-arrangement textures (made up of both horizontal and vertical grating patches) is approximately as good as the segregatability of the most effective component. Models having opponency between perpendicular orientations lead to incorrect predictions. Such models predict that the segregatability of the compound should be substantially greater than that of the most effective component (if the opponency is of any of the ORIENTATION-opponent types in Fig. 2) or substantially less (if the opponency is of the BOTH-opponent type in Fig. 2).

Normalization (a contrast-gain control that results from inhibition among channels) must be added to the SIGN-opponent-only channels model in order to account for the contrast-dependent effect seen in our results: as the contrast of the patterns is increased, the compound becomes somewhat less segregable relative to the most effective component. That normalization proved necessary to account for the contrast-dependent results here is one more instance of the power of normalization—a process suspected now on a number of grounds briefly mentioned above—to account for behavioral as well as neurophysiological results.

Acknowledgments

Some of this work was presented in Wolfson and Graham (2003). This work was supported by National Eye Institute grant EY08459.

Overview of appendices

The following appendices supplement the material in the main text which introduced the model (in Fig. 1). The appendix provides further details of the assumptions, gives equations, and presents references to more extended explanations in previously published work.

Appendix A. Predictions for individual second-order channels (Figs. 7–9)

As was discussed in Sections 1 and 2 of the main text (and presented in Figs. 7–9), the three kinds of channel structure lead to very different predictions for the channel responses to the compound versus component stimuli. In this section of the appendix, we present the equations and some further discussion of the assumptions underlying them.

A.1. Equations for individual channels including spatial pooling and within-channel differencing

The two component patterns in Figs. 7–9 are of equal contrast. But it is straightforward to generalize these illustrations to unequal contrasts of the components, and it leads to the same relationship. Thus unequal component contrasts lead to the same equations given below for the three types of structure (Eqs. (1)–(3)).

Further, the areas and sensitivities of the centers and flanks of the receptive fields in Figs. 7–9 are shown as "balanced", that is, the sensitivity at each point in the center region is twice that at the corresponding point in the flanks but the area of the center is half that of the total of the two flanks. Thus the integrated sensitivity over the center area is the same as the integrated sensitivity over the whole flank area. However, it is again straightforward to generalize these illustrations to allow unbalanced receptive fields. This changes the relationship between the channel's threshold for one component, and its threshold for another, thus allowing the channels to explain biases like that shown for the consistent versus inconsistent components by many of our observers (and/or biases for the vertical vs horizontal patches). But these unbalanced channels' outputs still show the same qualitative relationships illustrated in Figs. 7–9 and thus are described by the same Eqs. (1)–(3).

As mentioned in the main text, the phases of all the grating-patch elements are drawn in Figs. 7–9 to be identical, but the same predictions hold for the random-phase condition, because the patches are assumed to be far enough apart that no single first-stage receptive field responds substantially to more than one patch. (That we find the same results in the variable-phase and constant-phase experiments, see Fig. 11, is the evidence for this assumption.) Thus the Eqs. (1)–(3) apply to the variable-phase condition also.

The exact relationships illustrated in Figs. 7–9 and embodied in Eqs. (1)–(3) do depend significantly, however, on one further simplifying assumption, namely the assumption that the intermediate nonlinearity is an ordinary (piecewise-linear) rectification (either full-wave or half-wave rectification) The effects of loosening this assumption are discussed below (in Appendix B) after

we have derived the equations under this simplifying assumption of piecewise linearity.

Further, in Figs. 7–9, the spatial pooling process is assumed to be a simple peak-trough difference, but in our work more generally we have considered a family of spatial pooling rules, and for patterns like those in this study, the spatial-pooling rule will make no difference to the equations. (The reader can find more detail about the spatial pooling rules in Graham, 1991, and Graham et al., 1992.)

To present these Eqs. (1)–(3) and further equations below, we need to define some further symbols. (The definitions of symbols in this appendix will be bordered so that the reader can refer back to them more easily when necessary. The definitions of symbols in the main text will be repeated here.) See Fig. 1 for overall model.

> Let *cmpd* stand for a compound stimulus, and *oeo1* and *oeo2* stand for the two component stimuli ("*oeo*" for one-element-only) composed of elements of type 1 and type 2, respectively.
>
> Let $R_i(stripe)$ and $R_i(check)$ stand for the spatially-pooled response of channel i in the striped region or checked region, respectively.
>
> Let $D_i = |R_i(stripe) - R_i(check)|$.
>
> For shorthand, we often refer to D_i simply as the channel's response.

We will first look at the predictions for individual channels' D_i for each of the three kinds of opponent structure before proceeding to consider the observer as a whole (in Appendix C).

A.1.1. Equations for outputs from SIGN-opponent-only channels

For the SIGN-opponent-only structure, the channel's response to one component pattern is substantial (see left column Fig. 7) because the channel's structure matches the component's structure (inconsistent in this case—the orientation of the first-stage receptive fields is perpendicular to that of the second-stage receptive field). The channel's response to the other component is zero (consistent, middle column Fig. 7). And the channel's response to the compound pattern (right column) is always identical to its response to the effective component (left column), which is component 1 here, so:

$$D_i(cmpd) = D_i(oeo1) \tag{1a}$$

And $D_i(oeo2)$ is always 0.

As mentioned in the main text, there is another SIGN-opponent-only channel exactly like that in Fig. 7 except having vertical instead of horizontal first-stage receptive fields (the channel referred to in the main text as the *fraternal twin* of the channel in Fig. 7). This twin is sketched in the top-left of Fig. 3. For this twin channel, which we will call channel j

$$D_j(cmpd) = D_j(oeo2) \tag{1b}$$

and $D_j(oeo1)$ which is always zero.

There can also be SIGN-opponent-only channels in which the centers are inhibitory rather than excitatory as shown. The interested reader can convince himself (by constructing or imagining a figure analogous to Fig. 7) that these channels lead to the same equations as the channels shown. And thus we can ignore them without loss of generality.

There are also many possible SIGN-opponent-only channels that do not respond to either component, e.g. those in the right column of Fig. 3. They can be ignored here.

A.1.2. Notes applying to both the next cases

Note that these last two comments—the first about ignoring channels that do not respond to either component and the second about ignoring channels in which the excitatory and inhibitory signs have been switched—also apply to the two other kinds of opponent structure (ORIENTATION-opponent and BOTH-opponent) and will not be repeated below.

Note also that, as described in the main text, one does *not* need to consider a "fraternal twin" in the case of the ORIENTATION-opponent or BOTH-opponent structures.

A.1.3. Equations for outputs from ORIENTATION-opponent channels

For the ORIENTATION-opponent case we will first consider the ORIENTATION-opponent-only structure and then generalize. For the ORIENTATION-opponent-only structure, illustrated in Fig. 8, a channel's outputs to the two component patterns are identical. For some intuition into why this is so, you might consider the following. There is a symmetry or interchangeability of roles between the vertical and horizontal sub-entities both in the channels (where the sub-entities are first-stage receptive fields) and in the stimuli (where the sub-entities are grating patches). Therefore, what occurs in the center of the second-stage receptive field for one of the component patterns (the left column of Fig. 8) occurs in the flanks of the second-stage receptive field for the other (the middle column of Fig. 8). And then, since the center and flanks are both of the same sign in this structure, the effects in the center and flanks are summed in the response to the compound stimulus (right column of Fig. 8). Thus the response of an ORIENTATION-opponent-only channel to the compound stimulus is literally the *sum* of its responses to the components:

$$D_i(cmpd) = D_i(oeo1) + D_i(oeo2) \tag{2}$$

The prediction of either the DOUBLY-opponent or the HALF-DOUBLY-opponent structure is again Eq. (2) as

the reader could verify by sketching diagrams like those in Figs. 7–9 for these cases.

A.1.4. Equations for outputs from BOTH-opponent channels

For the BOTH-opponent structure, it is again true (as for the ORIENTATION-opponent-only structure) that what happens in the center of the second-stage receptive field for one of the components patterns (left column of Fig. 9) happens in the flanks of the second-stage receptive field for the other (middle column of Fig. 9). But for this structure, the center and flanks of the second-stage receptive field have oppositely-signed effects. Thus, in response to the compound stimulus (right column of Fig. 9), the effects in the center and flanks cancel each other out (right column of Fig. 9). More precisely, the response of the BOTH-opponent channel to the compound is the *difference* between its responses to the two components:

$$D_i(cmpd) = |D_i(oeo1) - D_i(oeo2)| \quad (3)$$

A.2. Equations for the predictions of individual channels' thresholds on summation squares

A channel's segregation threshold for a pattern is defined as the contrast in the pattern that produces a channel response D_i of criterion magnitude. For the models we have been discussing so far, the predicted thresholds are the same no matter what the value of the criterion magnitude so we will let $D_i = 1$ at threshold to make expressions simpler. To make predictions for the thresholds, we need to find the contrasts which produce a criterion-magnitude response. (We consider all contrast values to be positive in this paper. The 180° phase shift which might correspond to a negative contrast is spoken of explicitly as a phase shift here.)

Let the symbols $C_1(pattern)$ and $C_2(pattern)$ represent the contrasts in elements of type 1 and of type 2, respectively, in the pattern called *pattern*.

Let $C^*_{i1}(pattern)$ and $C^*_{i2}(pattern)$ represent the contrasts in elements of type 1 and elements of type 2 of *pattern* when *pattern* is at threshold for channel *i*, that is, when $D_i(pattern) = 1$. When the identity of the pattern is clear from context, we will often shorten these to C^*_{i1} and C^*_{i2} to make the equations more readable.

Let $C_{ith1} = C^*_{i1}(oeo1)$ and $C_{ith2} = C^*_{i2}(oeo2)$. That is, C_{ith1} and C_{ith2} are the contrasts in component patterns *oeo*1 and *oeo*2, respectively, when they are at the *i*th channel's threshold.

Note that, in these symbols, the ratio C^*_{i1}/C_{ith1} and C^*_{i2}/C_{ith2} are the relative contrasts in components 1 and 2, respectively, of the pattern under discussion when that pattern is at the *i*th channel's threshold.

For any channel having a piecewise-linear (ordinary) rectification between the two stages of receptive fields—as assumed in Figs. 7–9—the responses $D_i(oeo1)$ and $D_i(oeo2)$ are proportional to the contrasts in *oeo*1 and *oeo*2, respectively. And the constants of proportionality can be expressed as the reciprocal of the channel's thresholds for the components. That is, remembering that we are letting $D_i = 1$ at threshold, the equations are:

$$D_i(oeo1) = C_1(oeo1)/C_{ith1} \quad (4a)$$

$$D_i(oeo2) = C_2(oeo2)/C_{ith2} \quad (4b)$$

We now go on to consider the predicted thresholds for the compound pattern—both as graphed and as equations.

The locus of the channel's predicted thresholds for all compound patterns are plotted as the solid lines in the summation squares shown in the upper right sketches of Figs. 7–9. The shaded area represents compounds which are below threshold and therefore invisible to the channel, and the white area represents compounds which are above threshold. In general (there is one technicality mentioned below), each summation-square axis plots the relative contrast in one of the components of the compound pattern, where relative contrast equals the physical contrast divided by the *channel*'s threshold for that component by itself. These predictions are briefly discussed in more detail in the following paragraphs for each of the three types of structure.

For a SIGN-opponent-only channel illustrated in Fig. 7 and described by Eq. (1a), the channel's response D_i to the compound stimulus equals the response to the effective component. The response to the other component is approximately zero.

For the SIGN-opponent-only channel illustrated in Fig. 7, the segregation thresholds for compound stimuli fall on a straight vertical line going upward from (1, 0), the point representing the threshold of the effective component. Or in symbols, Eqs. (4a) and (1a) lead to

$$1 = (C^*_{i1}/C_{ith1}) \quad (5a)$$

(The technicality: if the response were exactly zero, the threshold C_{ith2} would be infinite, and the relative contrast, C_2/C_{ith2} plotted on the vertical axis of the summation square in Fig. 7 would always be zero. One can, without consequence, assume instead that C_{ith2} is not quite infinite and then relative contrast can take on a range of values as the figure implies. Alternately, the axis could have been labeled by physical contrast C_2 but we preferred to use relative contrast to make it the same as the summation squares in Figs. 8 and 9.)

The fraternal twin of the channel in Fig. 7 has the receptive field sketched in the top left of Fig. 3. As follows from Eqs. (4b) and (1b), it has thresholds that fall

on a straight horizontal line going rightward from (0, 1) and described by the equation

$$1 = (C^*_{j2}/C_{jth2}) \tag{5b}$$

Let us now consider the pair of twin channels leading to (5a) and (5b) together. (They would have to both be present in any reasonable model of an observer or else one component would be invisible.) The following equation serves for the pair of twins together:

$$1 = \min\{(C^*_{i1}/C_{ith1}), (C^*_{j2}/C_{jth2})\} \tag{5c}$$

The ORIENTATION-opponent and BOTH-opponent cases are somewhat more straightforward since there is not a fraternal pair of channels to be considered.

For any of the three types of ORIENTATION-opponent structure, remember that the channel's response D_i to the compound stimulus equals the sum of its two responses to the two components (Eq. (2) and illustrated in Fig. 8 for the ORIENTATION-opponent-only case.) Then Eqs. (4) and (2) lead to

$$1 = (C^*_{i1}/C_{ith1}) + (C^*_{i2}/C_{ith2}) \tag{6}$$

Thus an ORIENTATION-opponent channel's thresholds for compound stimuli fall on the negative diagonal exhibiting linear summation, and the compound is much more segregatable than either component.

For the BOTH-opponent structure (Eq. (3) and Fig. 9), the channel's response D_i to the compound stimulus equals the absolute value of the difference between its responses to the two components. Thus a compound stimulus is above threshold for the channel if and only if the difference between the two responses to the components is above threshold. In symbols, Eqs. (4) and (2) lead to

$$1 = |(C^*_{i1}/C_{ith1}) - (C^*_{i2}/C_{ith2})| \tag{7}$$

Therefore, when plotted on relative contrast axes, the thresholds fall on lines of slope one that go upwards from the component thresholds plotted at (0, 1) and (1, 0). Notice that such a channel could never segregate a compound stimulus made up of two equally-effective components. More generally, the shaded region in the summation square—representing invisible compounds—extends out in the upward/rightward direction indefinitely.

Appendix B. Allowing for an expansive intermediate nonlinearity in the channels

Graham and Sutter (1998) presented evidence that the intermediate function in complex channels (SIGN-opponent-only channels) is NOT a piecewise-linear rectification as we have been assuming up to this point. Instead it is expansive. This possibility is briefly mentioned in the main text and entered into the calculation of the black areas (the ranges of uncertainty) in Fig. 6 showing predictions for the observer. Here in the appendix we present the equations that led to those predictions.

Graham and Sutter (1998) also showed that there were at least three ways that such an expansive nonlinearity can be inserted into the general two-stage filters structure of a complex channel. These three ways all can be described as three different versions of second-stage pooling. All three are consistent with the Graham and Sutter (1998) results. As part of another study, Graham and Sutter (2000) produced substantial evidence against the second version, but that still leaves two candidates: Versions #1 and #3.

Since the results of this current study will not be able to discriminate among any of the three versions, we will not say anything in detail about any of the versions here. We will simply re-state the summation-square predictions above generalized to include the possibility of expansive nonlinearities in any of the three versions. To do so, we will assume for simplicity's sake that the pointwise function at the intermediate stage of the channel is described by a power function:

$$g(x, y) = a \cdot |f(x, y)|^{k_m} \tag{8}$$

where $f(x, y)$ is the input at point (x, y) to the intermediate nonlinearity in the channel and $g(x, y)$ is the output; k_m is the exponent describing the power function, and the parameter a simply scales quantities and can be ignored.

If $k_m = 1$, then g is a piecewise-linear function as in conventional full-wave rectification which is assumed in Figs. 7–9. If $k_m > 1$, then g is an expansive function. If $k_m < 1$, then g is compressive. We might have considered a family of functions larger than that in Eq. (8) but this family proved sufficient for our purposes. Indeed the Graham and Sutter (1998) results for seven subjects constrain this function to be expansive with an exponent in the range of 2–4.

The results of the Graham and Sutter (1998) study hold on the assumption that they were working with complex channels, that is, to the SIGN-opponent-only case. But here we also considered the possibility that the other types of opponent structures, if they existed, would have an expansive nonlinearity.

One can derive straightforwardly the predictions of all the opponent structures pictured in Fig. 2 when considered in conjunction with each of the three versions for incorporating an expansive nonlinearity from Graham and Sutter (1998). We will not go through these derivations here for the sake of space, but the resulting equations are given in the next few paragraphs. Illustration of these functions will be postponed until after we have gotten to the predictions for the observer's (rather than an individual channel's) threshold.

For the SIGN-opponent-only channels, for all three versions of incorporating the expansive nonlinearity, there is no effect of the intermediate nonlinearity. That is, the following equation, which is identical to Eq. (5c), gives the predicted thresholds for channel i and its twin j regardless of the value of k_m

$$1 = \min\{(C^*_{i1}/C_{ith1}), (C^*_{j2}/C_{jth2})\} \quad (9)$$

For the ORIENTATION-opponent channels, for all three versions of incorporating the expansive nonlinearity, the equation giving the predicted thresholds of an individual channel i is of the following form:

$$1 = (C^*_{i1}/C_{ith1})^B + (C^*_{i2}/C_{ith2})^B \quad (10)$$

where B is a parameter that depends on both the version of incorporating the expansive nonlinearity and on the value of the exponent in that expansive nonlinearity.

In particular, for ORIENTATION-opponent channels, for Versions #1 and #2, the value of $B = k_m$. However, for Version #3, the value of $B = 1$ no matter what value of k_m characterizes the intermediate nonlinearity. Notice that, when $B = 1$, this Eq. (10) reduces to Eq. (6) above as it should.

For a BOTH-opponent channel i, the equation is

$$1 = |(C^*_{i1}/C_{ith1})^B - (C^*_{i2}/C_{ith2})^B| \quad (11)$$

where (just as for the ORIENTATION-opponent case) $B = k_m$ for Versions #1 and #2 but for Version #3, $B = 1$. When $B = 1$, this equation reduces to the Eq. (7).

Appendix C. From channels to observers

To go from the responses of single channels as discussed in the preceding section to the response of the observer who is relying on multiple channels is not, in general, trivial. These problems are discussed in Graham (1989) for multiple channels (analyzers) generally and will be briefly mentioned here as they concern the study here. In Fig. 1, these processes are in the comparison-and-decision stage, and they were briefly introduced in Section 2 of the main text.

C.1. Discussion of assumptions

C.1.1. Pooling across channels and the decision variable

In our work we generally consider a family of possible rules for pooling across multiple channels, that is for computing a decision variable (represented by D_{obs}) for the observer from all the channels' responses (all the D_i). We then make sure that the conclusions in our experiment hold for all these rules (e.g. Graham & Sutter, 2000). This is the stage represented by the "Pooling across channels" box in the comparison-and-decision stage of Fig. 1.

When the observers in previously published work have been asked to rate the degree to which the checkerboard and striped textures effortlessly and immediately segregate (e.g. Graham & Sutter, 2000), we have made the simple assumption that the larger the value of D_{obs}, the larger the rated segregatability. In the study reported here, however, we ask the observer to identify the orientation of the embedded rectangle and we take the observer's performance in this identification task as an indicator of ability to segregate. Here, therefore, we make the simple assumption that the greater the pooled value D_{obs}, the better the observer is able to segregate the checkerboard from the striped region in the pattern and therefore, operationally, the greater the observer's performance in the forced-choice experiments of this study.

We do not consider these assumptions–or any of our other assumptions about the comparison-and-decision-stage—as anything but reasonable simplifying assumptions with which to approximate, for the purposes of this study, the action of ALL the higher stages of visual processing actually necessary to perform this task.

We do not know very much about these higher stages but one can imagine, for example, that not only does the inner region of texture have to be segregated from the outer (is this accomplished by "edge finding" or "region growing" or something else entirely? see Wolfson & Landy, 1998), but also the shape of the inner region has to be computed so that the observer can answer the question of whether the inner region is horizontally or vertically elongated. All of these processes (or something like them—this outline of processes may not even be correct) enter into the task on each trial, but none of them is explicitly represented in our model framework. We do not think enough is known about all these stages to make any more explicit statements in our model, nor would the results from these experiments provide any test of them.

Thus we opt for these families of simple assumptions represented in our comparison-and-decision stage and do explicit calculations to make sure that our conclusions do not change depending on the member of the family (e.g. Graham, 1991, pp. 276–277; Sutter et al., 1989, p. 317–318).

C.1.2. More about the pooling rules

In the family of pooling-across-channels rules, one important member is the simple assumption which amounts to saying that the observer's segregation threshold for a pattern equals the lowest of all the channels' segregation thresholds for that pattern. This assumption is frequently described as the assumption

that there is no probability summation or any other equivalent forms of deterministic nonlinear pooling.

To consider other possible assumptions that allow for probability summation (or some equivalent form of nonlinear pooling) we use the "Quick Pooling Model" that has been found so useful on many dimensions of pattern vision.

The family of pooling rules we consider for pooling across spatial positions is very similar to that for pooling across channels.

See Graham, 1989, for an introduction to these issues, in particular see pp. 167 onwards. And see our earlier papers (e.g. Sutter et al., 1989, and Graham, 1991; Graham et al., 1992 for more description of them as used here.)

C.1.3. What channels exist

The observer's response may be influenced by any and all channels that have any significant ability to do the task in question. Thus one initially needs to consider the full range of such channels that might reasonably be thought to exist on the basis of existing theory and that might contribute to the task. To make the problem tractable, however, one hopes to cut down the set of possible channels to a small number that need to be explicitly considered in the calculations. We do so in the next few paragraphs although many details are skipped.

C.1.3.1. Off-frequency off-orientation looking is not a problem. In many experiments involving spatial patterns, one finds that the channels which "intuitively" seemed like the ones that were being studied in the experiment—the channels that are "tuned" to the patterns in question (e.g. the channels in Figs. 7–9)—are NOT the only channels at issue. This happens because there are other channels tuned to quite different frequencies and orientations which play important roles and, unfortunately, quite complicated roles because their relative responsivities to different patterns in the study is quite different from that of the tuned channels. The observer can often perform better on certain stimuli by "looking at" these off-frequency and off-orientation channels than by looking at the tuned channels. (For a recent example of an elegant explanation using such off-frequency and off-orientation looking, see Kingdom et al., 2003 and Prins & Kingdom, 2002, 2003.)

However the intrusion of these off-frequency off-orientation channels is minimal for the study here. The fact that the intrusion is minimal is partly because the elements making up the patterns are spatially separate and partly due to the discrimination the observer is being asked to make. One can argue for this on various grounds, but perhaps most directly we also confirmed the arguments by doing explicit filtering of these patterns to calculate the responses of off-frequency and off-orientation channels. These responses were minimal. Thus we will ignore such channels here.

C.1.3.2. Multiple channels that are tuned to the patterns. Here we will consider only the "tuned" channels, where tuned channels in our study are those having first-stage receptive fields approximately matching the characteristics of the grating elements and second-stage receptive fields approximately matching the spacing and orientation of the stripes and checkerboard arrangements. But even considering only "tuned" channels leaves a multiplicity of potential channels that need to be at least briefly considered.

(i) One can start to simplify the situation by pointing out that if the multiple tuned channels all have the same balance (the same ratio of sensitivities to the two components), they can be safely ignored. For—if all of a group of channels have the same balance of sensitivities to components and if one uses the Quick Pooling Model as we do—then it turns out that one can model the result of the whole group of channels as a single channel.

(ii) Second, although it may not be obvious, this stipulation in the first paragraph applies to all SIGN-opponent-only channels since for any SIGN-opponent-only channel, the ratio of sensitivities to the two components is either zero or infinity. (Remember again that we are assuming that no first-stage receptive field is sensitive both to horizontal and to vertical receptive fields.) Thus all the channels having ratios of zero can be grouped into the same group as the channel in Fig. 7, and all those having ratios of infinity can be grouped with its fraternal twin. Therefore: All SIGN-opponent-only channels reduce to a single pair.

(iii) Third, some further consideration shows that ORIENTATION-opponent and BOTH-opponent channels with unbalanced sensitivities also can be ignored although with slightly more caution. The possibility of a spread of sensitivity ratios for different tuned channels needs to be considered slightly further for ORIENTATION-opponent and BOTH-opponent structures, because the predictions from such a set of tuned channels would not be precisely identical to the predictions shown in Fig. 6.

In fact, one can show quite easily graphically the predictions from a set of unbalanced ORIENTATION-opponent channels on a summation square. The prediction for each channel would be a line running from that channel's threshold on the vertical axis to that on the horizontal axis, and the lines would all have different slope. The prediction for the observer would be some-

thing like the inside envelope of this set of lines (or something further inside if there were probability summation among channels).

If the set of unbalanced ORIENTATION-opponent channels contains a channel that is among the channels most sensitive to the two orientations and is also rather balanced, then the predictions for the set of channels will be dominated by that particular channel will be very near a negative diagonal (and far from the experimental results).

If, however, the set of unbalanced ORIENTATION-opponent channels includes only channels that are very unbalanced, then all the individual channels' lines will be nearly horizontal or nearly vertical and the prediction for the observer will be very like that from the SIGN-opponent channels and will agree well with the experimental results. However, although we are calling these ORIENTATION-opponent channels, in a way that is a misleading name: since their predictions are almost vertical or almost horizontal they are actually so unbalanced as to effectively contain only one orientation, and thus to be essentially equivalent to SIGN-opponent channels anyway.

The exactly analogous point can be made about BOTH-opponent channels.

One way of describing the bottom line here is to say that, when we conclude in the main text that it is SIGN-opponent only channels, we should add the caveat: or else OPPONENT-orientation or BOTH-orientation channels that are so unbalanced they are essentially sensitive to one orientation only and therefore equivalent to SIGN-opponent-only channels.

Thus we will not consider further the question of unbalanced ORIENTATION-opponent or BOTH-opponent channels.

Conclusion. With little loss of generality, we can talk as if there were only a single channel of the ORIENTATION-opponent and BOTH-opponent type or a single fraternal pair of the SIGN-opponent-only type.

C.2. Predictions of the observer's performance (Fig. 6)

So, under these assumptions, lets look at the predictions for the observer—and the effect of the assumptions about the comparison-and-decision stage on these predictions—from the various opponent structures in Fig. 2. To describe the observer's behavior, the same notation will be used as for individual channels except that *obs* will be used in the subscripts rather than the letter indicating the individual channel. The use of a subscript '*th*' will still imply a threshold for a component by itself (*th*1 or *th*2 referring to *oeo*1 or *oeo*2, respectively). The use of a superscript * will imply the threshold for a pattern which is, in general, a compound pattern. The name of that pattern will be clear from context or given as an argument. In particular:

Let $D_{obs}(pattern)$ be the value of the variable determining the observer's response. A pattern is at threshold for the observer when $D_{obs}(pattern) = 1$. When the pattern is clear in context, we will use D_{obs}.

Let $C^*_{obs1}(pattern)$ and $C^*_{obs2}(pattern)$ be the contrasts in element type 1 and element type 2 of a pattern when that pattern is at threshold for the observer, that is, when $D_{obs}(pattern) = 1$. We will use C^*_{obs1} and C^*_{obs2} when the pattern is clear from context.

Let $C_{obsth1} = C^*_{obs1}(oeo1)$ and $C_{obsth2} = C^*_{obs2}(oeo2)$. So, C_{obsth1} and C_{obsth2} represent the contrasts in component patterns *oeo*1 and *oeo*2, respectively when they are at the observer's threshold.

The parameter k_d will represent the extent of nonlinear pooling across channels.

If the nonlinear pooling across channels is due to probability summation k_d is the exponent of the Quick function (Weibull function) that describes the psychometric function; alternately, k_d can just be taken as an exponent describing a more general form of nonlinear pooling.

For the SIGN-opponent-only channels (top row in Fig. 6), if there is no probability summation, the observer's thresholds are just equal to the threshold of whichever member of the pair of SIGN channels is most sensitive, and thus the observer's thresholds are predicted to fall on the right and top outside edges of the summation square (e.g. Fig. 6, top row, the straight edges of the black region). In symbols, it is the following equation, which is just like Eq. (9) except now it is for the observer rather than for the individual channels:

$$1 = \min\{(C^*_{obs1}/C_{obsth1}), (C^*_{obs2}/C_{obsth2})\} \tag{12a}$$

If, on the other hand, there is probability summation (or equivalent nonlinear pooling), the observer's thresholds near the upper right corner of the square will be reduced relative to the most sensitive channel's because there are two independent chances (of approximately equal probability) for the observer to detect the compound (the channel or channels sensitive to either component might detect it). More generally, the standard derivation from the Quick Pooling model shows that the thresholds will fall on a the locus described by the following equation:

$$1 = (C^*_{obs1}/C_{obsth1})^{k_d} + (C^*_{obs2}/C_{obsth2})^{k_d} \tag{12b}$$

The curved inside edge of the black area in Fig. 6 (top row) is approximately the curve from equation (12b) with $k_d = 2$, which is about the extreme of the range compatible with case of probability summation on other pattern dimensions.

(Some readers might be interested to note that equation (12a) is a subcase of equation (12b); it is the limit when k_d approaches infinity.)

For the ORIENTATION-opponent and BOTH-opponent cases, as discussed in Appendix C.1.3, there is effectively only one channel that needs to be considered (although it may be a proxy for a group of channels with similar balance of sensitivities in which some channels may overall be less sensitive than others). Thus, whether or not there is probability summation across channels, the locus of the predicted segregation thresholds for the observer has the same form as for the individual channel when plotted on relative contrast axis.

In particular, for the ORIENTATION-opponent channels (middle row, Fig. 6), the observer's segregation thresholds lie on a locus specified by an equation of the same form as Eq. (10) for an individual channel of this type, namely:

$$1 = (C^*_{obs1}/C_{obsth1})^B + (C^*_{obs2}/C_{obsth2})^B \quad (13)$$

For the BOTH-opponent channels (bottom row, Fig. 6), the observer's thresholds lie on a locus specified by an equation of the same form as Eq. (11) for an individual channel of this type, namely

$$1 = |(C^*_{obs1}/C_{obsth1})^B - (C^*_{obs2}/C_{obsth2})^B| \quad (14)$$

As indicated in the description of Eqs. (10) and (11) the value of B is determined jointly by which of the three versions from Graham and Sutter (1998) is used to incorporate the expansive nonlinearity and by the exponent of the intermediate nonlinearity.

The smallest reasonable value for B is 1, which leads to the solid straight edge of the black area in the summation square of Fig. 6 (middle and bottom row). The value $B = 1$ is predicted either when the intermediate stage has a piecewise-linear rectification (for any of the three versions) or when Version #3 is used regardless of the exponent of the intermediate nonlinearity.

The largest reasonable value for B is in the range 2–4. This value is predicted for Versions #1 and #2 with an expansive nonlinearity of exponent 2–4 (which is the exponent of the expansiveness found in Graham & Sutter, 1998). The curved edges of the black region in Fig. 6 (middle and bottom row) shows the case of B equal to approximately 2. If B were as high as 4 the dark ranges would extend somewhat further out for the ORIENTATION-opponent case (or further in for the BOTH-opponent case), but would not get far enough out or in to include the experimental results (as the interested reader could easily compute).

Appendix D. Allowing for normalization among channels

This section of the appendix presents some details for computing the effects, described in Section 4 of the main text, of inhibition among channels. This inhibition is modeled by a normalization network that has been sufficient to account for other phenomena in the perception of these patterns (Graham & Sutter, 2000).

D.1. Equations for predicting observer performance

This section of the appendix requires more extensive notation than the earlier sections because, once normalization is included, the channels' thresholds depend on the threshold criterion level of the decision variable D_i. And we will want to discuss the full psychometric functions. Here therefore we cannot just present equations in terms of channel thresholds as we did in earlier sections. We need to present equations here in terms of the magnitudes both of channels' responses and of the Decision Variable determining the observer's response. For these last few sections of the appendix, a subscript *NORM* will be added to indicate explicitly that the model with normalization taken into account is being used. And a subscript NO-NORM will similarly explicitly indicate the model without any normalization. Thus:

> Let $D_{obs,NORM}(pattern)$ be the value of the decision variable determining the observer's response to *pattern* once normalization has been taken into account. This will be shortened to $D_{obs,NORM}$ when the pattern is clear from context.
>
> Let $D_{i,NORM}(pattern)$ or $D_{i,NORM}$ represent the ith channel's response with normalization taken into account.
>
> Similarly, we will let $D_{obs,NO\text{-}NORM}(pattern)$ or $D_{obs,NO\text{-}NORM}$ be the value of the decision variable determining the observer's response in a model in which there is NO normalization.
>
> And we will let $D_{i,NO\text{-}NORM}(pattern)$ or $D_{i,NO\text{-}NORM}$ represent the ith channel's response in a model in which there is no normalization.

Using this notation one can write an equation allowing for an indeterminate number of distinct channels and allowing for probability summation among them:

$$D_{obs,NORM} = \{D^{k_d}_{1,NORM} + D^{k_d}_{2,NORM} + \cdots\}^{1/k_d} \quad (N1)$$

where, as in Eq. (12b), the exponent k_d represents pooling across channels that occurs via probability summation or equivalent nonlinear pooling in the Quick Pooling Model, and the symbol '$\cdots$' indicates the possibility of more terms of the same form as the two terms already written. Note that with equations like this, one can first pool over subsets of the whole set and then pool over these intermediate quantities, and the answer is the same as if one had pooled over the whole set to begin with. Thus one can take the individual terms in

Eq. (N1) to be distinct subgroups of channels rather than individual channels. (For this fact expressed in equations, see p. 726 of Graham et al., 1992.)

D.1.1. Equations for channels' response magnitudes for the three opponent structures

A channel's response with normalization taken into account $D_{i,NORM}$ is equal to the response before normalization is taken into account $D_{i,NO\text{-}NORM}$, divided by the normalization pool, which will be defined below and which will be denoted POOL:

$$D_{i,NORM} = D_{i,NO\text{-}NORM}/POOL \tag{N2}$$

See earlier publications for further justification and discussion of this modeling approach to normalization, particularly Graham et al., 1992, and Graham and Sutter, 2000, and their appendices.

The D_i will be as previously discussed for the channel types (see Eqs. (1)–(3)) but we need to rewrite them here in a convenient form for this section (using sensitivity rather than threshold notation the way we did above). To do so

Let w_{i1} and w_{i2} represent the (positive) sensitivities of the ith channel to element types 1 and 2, respectively.

Remember that $C_1(cmpd)$ and $C_2(cmpd)$ are the contrasts in elements of type 1 and 2, respectively, in the pattern *cmpd* (and that contrast is always positive in this study).

Remember that, with little loss of generality, we can ignore all channels except a twin pair of channels for the SIGN-opponent-only case and all channels except a single channel for each of the ORIENTATION-opponent and the BOTH-opponent cases (see Appendix C.1.3). Therefore

For the SIGN-opponent-only case, consider a twin pair of channels i and j. Then $D_{i,NO\text{-}NORM(cmpd)} = w_{i1} \cdot C_1(cmpd)$ and $D_{j,NO\text{-}NORM(cmpd)} = w_{j2} \cdot C_2(cmpd)$ To have more compact notation, we will drop the *cmpd* in what follows, so:

$$D_{i,NO\text{-}NORM} = w_{i1} \cdot C_1 \quad \text{and}$$
$$D_{j,NO\text{-}NORM} = w_{j2} \cdot C_2 \tag{N3}$$

For the ORIENTATION-opponent case, there is only one distinct channel, and

$$D_{i,NO\text{-}NORM} = (w_{i1} \cdot C_1)^B + (w_{i2} \cdot C_2)^B \tag{N4}$$

For the BOTH-opponent case, there is only one distinct channel, and

$$D_{i,NO\text{-}NORM} = |(w_{i1} \cdot C_1)^B - (w_{i2} \cdot C_2)^B| \tag{N5}$$

where B has the same meaning as for equations (13) and (14) above and can vary between 1 and about 4 reasonably.

D.1.2. The equation for the normalization pool

The value of POOL will depend on the total contributions of all channels responding to the textures including potentially many channels that do not contribute to the segregation judgment (channels for which the value of $D_{i,NORM} = D_{i,NO\text{-}NORM} = 0$). Following our earlier usage (e.g. Graham & Sutter, 2000) these channels that do not contribute to segregation but do contribute to the normalization pool will be called "other" channels in what follows.

Let the symbol R_{oi} represent the (always positive) response of the ith "other" channel (where o in the subscript indicates "other") as it enters into POOL.

For the channels that do contribute to the segregation, their contribution to normalization (for the patterns we are using in these experiments) can be shown to be the same as their contribution to segregation and is thus equal to the values D_i given above. So, writing a general formula that includes an indefinite number of channels that contribute to segregation as well as an indefinite number of "other channels", one gets

$$POOL = \{\sigma + D_{1,NO\text{-}NORM}^{k_n} + D_{2,NO\text{-}NORM}^{k_n} + \cdots + R_{01}^{k_n} + R_{02}^{k_n} + \cdots\}^{1/k_n} \tag{N6}$$

where

The parameter k_n is the exponent describing the pooling of different channels responses in the normalization pool.

The parameter σ is the parameter that sets the extent of the linear range in Eq. (N2) and keeps the denominator in that equation above zero. Varying σ is a way to control the strength of the normalization.

We need to derive an expression to be used in our prediction calculations to represent the magnitude of response from these "other" channels. To do so, we need to specify the "other" channels in more detail—that is, to specify what "other" channels might be responding to our patterns and therefore need to be modeled. As discussed in the next several paragraphs, we considered two very different cases of "other "channels in the set of simulations reported here, and we think these two cases cover the extremes of all likely cases of "other" channels for this study. For each of these two cases we considered both second-order opponent channels (those we have been considering) and simple linear spatial-frequency and orientation-selective channels. Such simple channels cannot do the task (segregate these second-order

patterns) and thus have not been considered so far in this description. But they are responding to the individual elements in both texture regions and thus will contribute substantially to the normalization pool.

D.1.2.1. First case of "other" channels. At one extreme we consider "other" channels that are equally responsive to both element types. For the current study, there are no such simple (linear) channels, because no simple channel is thought to be responsive to both vertical and horizontal elements. Further there are no second-order channels of the SIGN-opponent-only type that are responsive to both vertical and horizontal elements. However, when we do the predictions for second-order channels of the BOTH-opponent or ORIENTATION-opponent types, they do respond to both orientations of element, and therefore perhaps some of them are "other" channels. Perhaps some are unable to segregate the textures (because they have inappropriately-sized second-stage receptive fields), but they still respond to individual elements (because they have the appropriately-sized first-stage receptive fields) and thus contribute to the normalization pool. This possibility is incorporated into an equation identical to that used for the complex "other" channels in Graham and Sutter (2000), namely

$$R_{OXi} = \{(w_{OXi} \cdot C_1)^{k_{sp} \cdot B} + (w_{OXi} \cdot C_2)^{k_{sp} \cdot B}\}^{1/k_{sp}} \quad \text{(N7)}$$

where the subscript ox is used to indicate a second-order "other" channel. Thus R_{OXi} is the response of the ith second-order "other" channel, and w_{OXi} is its sensitivity to element type i.

The parameter k_{sp} is an exponent describing pooling across spatial position within the output from any single channel.

In practice k_{sp} might well be taken to equal k_d since the exponent describing nonlinear pooling across spatial position might well be equal to the exponent describing nonlinear pooling across different channels (see further description of spatial pooling and k_{sp} in Graham & Sutter, 2000).

B has the same meaning as in Eqs. (13) and (14) above for ORIENTATION-opponent and BOTH-opponent structures and can vary between 1 and about 4 reasonably. For SIGN-opponent structures it is always equal to 1.0. Remember that all contrasts are taken to be positive here.

Second case of "other" channels. At the other extreme, we considered "other" channels which are sensitive to only one of the two types of elements. These can be either simple or second-order potentially. The relevant simple "other" channels, those responding, respectively, to element type 1 (only) and element type 2 (only), have responses that can be well approximated as

$$R_{OS1} = w_{OS1} \cdot C_1 \quad \text{and} \quad R_{OS2} = w_{OS2} \cdot C_2 \quad \text{(N8)}$$

where OS is used in the subscript to indicate these are simple "other" channels.

The second-order "other" channels, for which one needs to allow for the possibility of the expansive nonlinearity, have responses that can be well approximated as

$$R_{OX1} = w_{OX1} \cdot C_1^B \quad \text{and} \quad R_{OX2} = w_{OX2} \cdot C_2^B \quad \text{(N9)}$$

where B has the same meaning as in discussion of Eqs. (13) and (14) above for ORIENTATION-opponent and BOTH-opponent structures and can vary between 1 and about 4 reasonably. For SIGN-opponent structures it is always equal to 1.0.

D.1.3. Assembling the equations into equation for observer

We will illustrate the assembling of the above equations into the model we simulated for the SIGN-opponent-only channels. From Eqs. (N2) and (N3) we get

$$D_{1,NORM} = (w_{11} \cdot C_1)/POOL \quad \text{and}$$
$$D_{2,NORM} = (w_{22} \cdot C_2)/POOL \quad \text{(N10)}$$

Then substituting Eq. (N10) into (N1) gives

$$D_{obs,NORM} = [\{(w_{11} \cdot C_1)^{k_d} + (w_{22} \cdot C_2)^{k_d}\}^{1/k_d}]/POOL \quad \text{(N11)}$$

Finally, we can get the expression for POOL by using Eq. (N6) for the general form of POOL and substituting into it by: (a) using Eq. (N3) for the channels that contribute to the pool but can also segregate and (b) using Eqs. (N8) and (N9) but not (N7) for the "other" channels. Remember that in Eq. (N9) the parameter B will be 1 for the SIGN-opponent-only case and thus can be ignored. Eq. (N7) is not used here because there are no second-order "other" channels responsive to both element types in the case of SIGN-opponent-only channels. These substitutions give:

$$POOL = \{\sigma + (w_{11} \cdot C_1)^{k_n} + (w_{22} \cdot C_2)^{k_n} + (w_{OS1} \cdot C_1)^{k_n} + (w_{OS2} \cdot C_2)^{k_n} + (w_{OX1} \cdot C_1)^{k_n} + (w_{OX2} \cdot C_2)^{k_n}\}^{1/k_n} \quad \text{(N12)}$$

Eqs. (N11) and (N12) together specify the model for the SIGN-opponent-only case with normalization allowed.

The equations for the ORIENTATION-opponent and BOTH-opponent cases with normalization can be derived similarly.

There are some steps missing here that need to be taken to justify our implicit assumption below that the normalization POOL is the same for all the channels

entering into a given prediction even though one imagines that this POOL has a limited spatial integration range and may also contain only a limited range of orientations and spatial frequencies.

For further explanation of the normalization model, see Graham and Sutter (2000) and previous papers.

D.2. Description of calculations from models including normalization

In order to understand the effect of normalization on the predictions for this study we computed predictions from the model including normalization for all three cases—SIGN-opponent-only, ORIENTATION-opponent, and BOTH-opponent—exploring a range of parameters very much like that explored by Graham and Sutter, 2000 (e.g. Fig. 14 in that paper for an overview) and which covers the reasonable ranges of the parameters in the context of these models.

We used contrasts in the range from 0 to 3 arbitrary units. And we set the parameters w_{11} and w_{22} to be 1. We considered values of σ from a minimum of 1 (which in these contrasts units produces the maximum possible normalization, see Fig. 14, right column, of Graham & Sutter, 2000) to a value so large that the predictions were identical to those without normalization. Notice that σ and contrast tradeoff so that increasing σ is the same as decreasing contrast. The values and ranges used here cover the interesting range of predictions for the model.

We considered values of the normalization pooling exponent $k_n = 1$ (linear summation within the pool), $k_n = 2$ (the value which makes our normalization pool much like that in the models of Heeger, 1991, 1992a, 1992b, and others), and $k_n = 8$ (just to check what happens as you go higher although no one has ever suggested such exponents to our knowledge).

We used B (reflecting the action of the intermediate nonlinearity as described above) equal to 1 and 3.

For all three cases we considered "other" channels described by Eqs. (N8) and (N9). For simplicity, we set $w_{OX1} = w_{OX2} = w_{OX}$ and considered $w_{OX} = 0$ and 4. Similarly we set $w_{OS1} = w_{OS2} = w_{OS}$ and considered $w_{OS} = 0$ and 4.

For the BOTH-opponent and ORIENTATION-opponent cases, we also considered "other" channels described by Eq. (N7) with $w_{OX} = 0$ and 4 and $k_{sp} = 1, 2$ and 4.

For the SIGN-opponent-only case, which is the only case where k_d matters, we considered values of the decision pooling exponent k_d of 2,4 and 30 (30 being a stand-in for the infinity that would produce no probability summation).

In order to consider the predictions for the thresholds, we also needed to consider a range of possible criterion threshold values (since once normalization is involved the predictions can depend on contrast values), and, in fact, we allowed the criterion threshold value to vary throughout the whole range of response magnitudes generated by the models.

D.3. Results of the calculations from models including normalization (Fig. 12)

First, after adding normalization to the models, the predictions from ORIENTATION-opponent and BOTH-opponent structures are still inconsistent with the experimental results. More particularly,

For the ORIENTATION-opponent channels, over the whole region of the explored parameter space, the predictions remains in the black area of Fig. 6 or are close enough to that black area to allow one to say conclusively: the predictions even with normalization included do not agree with our experimental results. There was one part of parameter space where the predictions moved out quite far toward the corner actually (toward 0.94, 0.94). But this is not really far enough to account for the data and, further, the value k_n was 1, which means linear summation inside the normalization pool, which has not been suggested by other authors.

For the BOTH-opponent channels, all the predictions extended so far to the upper right (as in the black area of Fig. 6 although sometimes of different shape) that the compound pattern containing equally-effective components would never be detectable at all. This is not consistent with our results.

Second, adding normalization to models with SIGN-opponent-only channels leads to an interesting result. In particular, a model using SIGN-opponent-channels and normalization can explain not only the general results (e.g. the Fig. 10 summation square) but also the difference between low contrast and high contrast results shown in the psychometric functions of Fig. 11. Some more detailed comments follow.

For the SIGN-opponent-only case, over the whole parameter space, the predictions for low to moderate threshold criteria remain consistent with the data in Fig. 10, that is, the threshold for the compound would be the same as that for the most effective component.

At high values of threshold criteria, however, the predictions from normalization are different from those at low or middle values.

Some of the predictions from normalization were in the opposite direction from the experimental results but this occurred only when the parameters in the predictions had values that seem to be inconsistent with the full set of results in this field. In particular, k_n (the normalization pool exponent which is taken by most authors to be 2) would have to be greater than k_d (the decision exponent that controls probability summation among the channels) which is rarely considered to be less than 4 and, considering how little probability

summation is shown at low contrasts in these data here you might think it would be even larger.

The predictions from normalization when the parameters were in other parts of the parameter space we explored were in the direction of the experimental results (the compound becoming less detectable than the components). Examples of these predictions are shown in Fig. 12, examples that capture many features of the empirical results. But the exact values of these parameters should not be taken very seriously as many other combinations would do. For the record, however, for these predictions these parameter values were as follows (identical for both panels except where indicated):

$$k_m = 3, \quad B = 3, \quad k_n = 1, \quad k_d = 30, \quad k_{sp} = 2, \quad \sigma = 99$$

$w_{11} = 1$ for the top panel;

$w_{11} = 1.1$ for the bottom panel

$$w_{22} = 1, \quad w_{OS1} = w_{OS2} = 0, \quad w_{OX1} = w_{OX2} = 1$$

References

Arsenault, A. S., Wilkinson, F., & Kingdom, F. A. A. (1999). Modulation frequency and orientation tuning of second-order texture mechanisms. *Journal of the Optical Society of America A, 16*, 427–435.

Beck, J. (1982). Texture segmentation. In J. Beck (Ed.), *Organization and representation in perception* (pp. 285–317). Hillsdale, NJ: Erlbaum.

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.

Bergen, J. R., & Landy, M. S. (1991). Computational modeling of visual texture segregation. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 253–271). Cambridge, MA: MIT Press.

Bonds, A. B. (1993). The encoding of cortical contrast gain control. In R. M. Shapley & D. M. Lam (Eds.), *Contrast sensitivity* (pp. 215–230). Cambridge, MA: MIT Press.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 443–446.

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

Ellemberg, D., Hess, R. F., & Allen, H. A. (2004). Evidence for spatial frequency and orientation labelled detectors in second-order visual processing. *Vision Sciences Society* (Abstract #E68, p. 137).

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

Gouras, P. (1991). Color vision. In E. R. Kandel, J. H. Schwartz, & T. M. Jessell (Eds.), *Principles of neuroscience* (third ed., pp. 467–480). New York: Elsevier.

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 M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 273–290). Cambridge, MA: MIT Press.

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

Graham, N., & Sutter, A. (1996). Effect of spatial scale and background luminance on the spatial and intensive nonlinearities in texture segregation. *Vision Research, 36*, 1371–1390.

Graham, N., & Sutter, A. (1998). Spatial summation in simple (Fourier) and complex (non-Fourier) channels in texture segregation. *Vision Research, 38*, 231–257.

Graham, N., & Sutter, A. (2000). Normalization: Contrast-gain control in simple (Fourier) and complex (non-Fourier) pathways of pattern vision. *Vision Research, 40*, 2737–2761.

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

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

Graham, N., & Wolfson, S. S. (2001). A note about preferred orientations at the first and second stages of complex (second-order) texture channels. *Journal of the Optical Society of America A, 18*, 2273–2281.

Gray, R., & Regan, D. (1998). Spatial frequency discrimination and detection characteristics for gratings defined by orientation texture. *Vision Research, 38*, 2601–2617.

Heeger, D. J. (1991). Nonlinear model of neural responses in cat visual cortex. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 119–133). Cambridge, MA: MIT Press.

Heeger, D. J. (1992a). Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9*, 181–197.

Heeger, D. J. (1992b). Half-squaring in responses of cat striate cortex. *Visual Neuroscience, 9*, 427–433.

Itti, L., Koch, C., & Braun, J. (2000). Revisiting spatial vision: toward a unifying model. *Journal of the Optical Society of America A, 17*, 1899–1917.

Kingdom, F. A. A., & Keeble, D. R. T. (1996). A linear systems approach to the detection of both abrupt and smooth spatial variations in orientation-defined textures. *Vision Research, 36*, 409–420.

Kingdom, F. A. A., & Keeble, D. R. T. (1999). On the mechanism for scale invariance in orientation-defined textures. *Vision Research, 39*, 1477–1489.

Kingdom, F. A. A., & Keeble, D. R. T. (2000). Luminance spatial frequency differences facilitate the segmentation of superimposed textures. *Vision Research, 40*, 1077–1087.

Kingdom, F. A. A., Keeble, D. R. T., & Moulden, B. (1995). Sensitivity to orientation modulation in micropattern-based textures. *Vision Research, 35*, 79–91.

Kingdom, F. A. A., Prins, N., & Hayes, A. (2003). Mechanism independence for texture-modulation detection is consistent with a filter-rectify-filter mechanism. *Visual Neuroscience, 20*, 65–76.

Kwan, L., & Regan, D. (1998). Orientation-tuned spatial filters for texture-defined forms. *Vision Research, 38*, 3849–3855.

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

Landy, M. S., & Graham, N. (2003). Visual perception of texture. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (pp. 1106–1118). Cambridge, MA: MIT Press.

Landy, M. S., & Oruc, I. (2002). Properties of second-order spatial frequency channels. *Vision Research, 42*, 2311–2329.

Lennie, P. (1998). Single units and visual cortical organization. *Perception, 27*, 889–935.

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

Morgan, M. J., & Dresp, B. (1995). Contrast detection facilitation by spatially-separated targets and inducers. *Vision Research, 35*, 1019–1024.

Motoyoshi, I., & Kingdom, F. A. A. (2003). Orientation opponency in human vision revealed by energy-frequency analysis. *Vision Research, 43*, 2197–2205.

Motoyoshi, I., & Nishida, S. (2001). Temporal resolution of orientation-based texture segregation. *Vision Research, 41*, 2089–2105.

Mussap, A. J. (2001). Orientation integration in detection and discrimination of contrast-modulated patterns. *Vision Research, 41*, 295–311.

Nothdurft, H. C. (1997). Different approaches to the coding of visual segmentation. In M. Jenkins & L. Harris (Eds.), *Computational and psychophysical mechanisms of visual coding* (pp. 20–43). New York: Cambridge University Press.

Olzak, L. A., & Thomas, J. P. (1999). Neural recoding in human pattern vision: model and mechanisms. *Vision Research, 39*, 231–256.

Pelli, D. G. (1997). The video toolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Prins, N., & Kingdom, F. A. A. (2002). Orientation- and frequency-modulated textures at low depths of modulation are processed by off-orientation and off-frequency texture mechanisms. *Vision Research, 42*, 705–713.

Prins, N., & Kingdom, F. A. A. (2003). Detection and discrimination of texture modulations defined by orientation, frequency and contrast. *Journal of the Optical Society of America A, 20*, 401–410.

Prins, N., & Mussap, A. J. (2000). Alignment of orientation-modulated textures. *Vision Research, 40*, 3567–3573.

Prins, N., & Mussap, A. J. (2001). Adaptation reveals a neural code for the visual location of orientation change. *Perception, 30*, 669–680.

Prins, N., Nottingham, N. K., & Mussap, A. J. (2003). The role of local grouping and global orientation contrast in perception of orientation-modulated textures. *Vision Research, 43*, 2315–2331.

Rubenstein, B. S., & Sagi, D. (1993). Effects of foreground scale in texture discrimination tasks: Performance is size, shape, and content specific. *Spatial Vision, 7*, 293–310.

Schwartz, O., & Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control. *Nature Neuroscience, 4*, 819–825.

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience, 24*, 1193–1216.

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

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

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

Sutter, A., & Hwang, D. (1999). A comparison of the dynamics of simple (Fourier) and complex (non-Fourier) mechanisms in texture segregation. *Vision Research, 39*, 1943–1962.

Teo, P. D., & Heeger, D. J. (1994). Perceptual image distortion. In *Human vision, visual processing, and digital display V* SPIE Proceedings (Vol. 2179, pp. 127–141).

Thomas, J. P., & Olzak, L. A. (1997). Contrast gain control and fine spatial discriminations. *Journal of the Optical Society of America A, 14*, 2392–2405.

Victor, J. D., Conte, M. M., & Purpura, K. P. (1997). Dynamic shifts of the contrast-response function. *Visual Neuroscience, 14*, 577–587.

Watson, A. B., & Solomon, J. A. (1997). Model of visual contrast gain control and pattern masking. *Journal of the Optical Society of America A, 14*, 2379–2391.

Wilson, H. R. (1993). Nonlinear process in visual pattern discrimination. *Proceedings of the National Academy of Sciences, USA, 90*, 9785–9790.

Wolfson, S. S., & Graham, N. (2003). Exploring the opponent structure of complex (second-order) channels. *Journal of Vision, 3*(9), 615a.

Wolfson, S. S., & Graham, N. (in press). Element-arrangement textures in multiple objective tasks. *Spatial Vision*.

Wolfson, S. S., & Landy, M. (1998). Examining edge- and region-based texture analysis mechanisms. *Vision Research, 38*, 439–446.

Zetzsche, C., Krieger, G., Schill, K., & Treutwein, B. (1998). Natural image statistics and cortical gain control. *Investigative Ophthalmology and Vision Science, 39* (Abstract # 1978).