Pergamon

0042-6989(95)00030-5

Vision Res. Vol. 35, No. 19, pp. 2685–2695, 1995

0042-6989/95 $9.50 + 0.00

Discrimination of Compound Gratings: Spatial-Frequency Channels or Local Features?

HIROMI AKUTSU,*† GORDON E. LEGGE*

Received 14 March 1994; in revised form 25 January 1995

Models based on spatial-frequency channels and local features provide alternative explanations for suprathreshold pattern discrimination. We compared psychophysical discrimination data with the predictions of the Wilson and Gelb channel model and three local-feature models. The features were peak–valley local contrast, peak–peak local contrast, and luminance gradients. We measured visual sensitivity for discriminating compound gratings ($F + 3F$ or $F + 5F$, in peaks-add or peaks-subtract phases) whose component contrasts were yoked together so that a contrast increment in one component was accompanied by an equal decrement in the other. The Wilson and Gelb model accounted for the results with peaks-add gratings, but failed to predict those with peaks-subtract gratings. None of the local-feature models explained the results by themselves. Most of the data fell close to an envelope composed of the lowest thresholds of the three feature-detector models, although there were important exceptions. Our findings are consistent with the view that suprathreshold pattern discrimination is mediated by mechanisms responsive to spatially localized features and that more than one type of feature is used.

Pattern discrimination Contrast discrimination Spatial frequency channels Local features

INTRODUCTION

Descriptions of spatial patterns can be given in either the space domain or the spatial-frequency domain. Corresponding theoretical approaches to pattern discrimination focus on either spatially localized features of luminance waveforms or the outputs of spatial-frequency filters. We compared psychophysical discrimination of suprathreshold compound gratings with predictions of local-feature models and the Wilson and Gelb channel model.

The appearance of the sum of two sine-wave gratings (a compound grating) changes when the contrasts of the components change (see the luminance profiles in Fig. 1). Observers decribe these differences in several ways, such as changes in the fuzziness or width of the bars, or the brightness or darkness of the bars. An observer may use changes in one or more of these *local features* to decide whether two compound gratings are the same or different. Several local features have been proposed as useful in discrimination (usually without specification of the underlying sensory mechanisms). These include the local contrast between adjacent luminance peaks and valleys (Badcock, 1984a, b), luminance gradients (Campbell, Johnstone & Ross, 1981; Hess & Pointer, 1987), the separation between luminance peaks (Burbeck, 1987), the width of a bright or dark bar (Bennett & Banks, 1991), the location of a luminance centroid or the average lightness in a given area (Bennett & Banks, 1991). Two local-feature models provide more details of the underlying mechanisms (Watt & Morgan, 1985; Morrone & Burr, 1988); local features, such as bar or edge location, are computed from the outputs of a set of spatial-frequency filters.

In the absence of a principled way of identifying a complete set of local features, we selected three features that have proven successful in previous work on suprathreshold pattern discrimination—*peak–peak local contrast*, *peak–valley local contrast*, and *gradient*.‡ These features convey to the observer percepts of the relative brightness (or darkness) of an area, the "contrast" of local regions, and the fuzziness or sharpness of contours.

It is logically possible, of course, that some alternative feature, not yet identified in the research literature, might provide a better account of the data than any of the features we studied.

Two of the features are based on relations between local peaks and valleys in the *luminance* waveform and are termed "local-contrast" features. In Fig. 1, the luminances of adjacent peaks and valleys (such as a and b, or b and c) can be used to define peak–valley local-contrast features. A given waveform may have

*Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, MN 55455, U.S.A.

†Present address: College of Optometry, University of Houston, 4901 Calhoun Boulevard, Houston, TX 77204–6052, U.S.A.

‡We also considered *width*, defined as distance between two inflection points. Discrimination predictions based on this feature were very poor and will not be considered further.

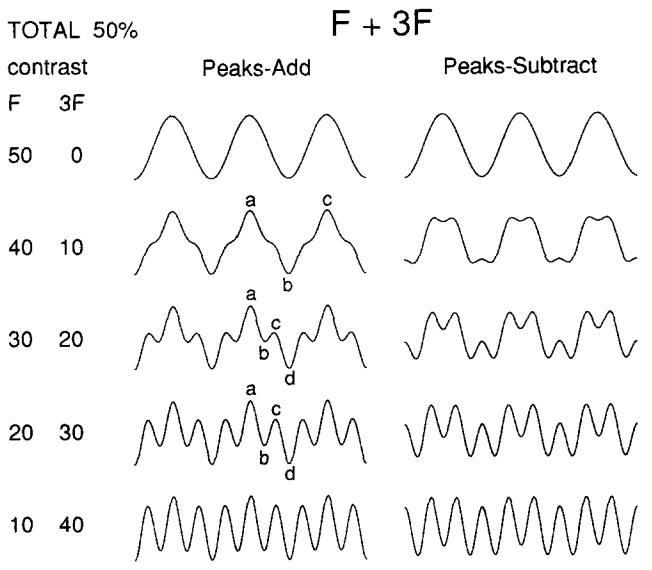

FIGURE 1. Luminance profile of compound gratings consisting of F and $3F$. Five contrast combinations are shown for peaks-add and peaks-subtract phases. The sum of the component contrasts is held fixed at 0.5. Note small shoulders around the mean luminance level in $F = 30$.

several distinctly different local contrasts. The Michelson (or some other) definition of contrast can be used to compute the local contrast from the luminances of the adjacent peaks and valleys. Similarly, the luminances of adjacent peaks (or adjacent valleys), such as a and c, or b and d for $F = 30\%$ and $3F = 20\%$ in Fig. 1, can be used to define peak–peak local-contrast features. Local-contrast features were proposed by Badcock (1984a, b) to account for pattern discrimination.*

Our third local feature is the luminance gradient, previously proposed by Campbell *et al.* (1981), and Hess and Pointer (1987). Luminance gradient features of a waveform are defined to be the maxima (unsigned) of the first derivative (occurring at points of inflection). This is closely related to the concept of zero crossing, proposed on computational grounds, by Marr and Hildreth (1980).

For the most part, theories of pattern discrimination based on spatial-frequency filters are more detailed than feature detection theories (Wilson & Gelb, 1984; Klein & Levi, 1985; Watson, 1983). These models are based on the spatial-frequency channel concepts introduced earlier (Campbell & Robson, 1968; Graham & Nachmias, 1971). Filter models have been successful in accounting for many types of threshold pattern detection, and for some types of suprathreshold pattern discrimination including spatial-frequency discrimination (Wilson & Gelb, 1984), vernier acuity (Wilson, 1986), bisection acuity (Klein & Levi, 1985), curvature discrimination (Wilson & Richards, 1989), and contrast discrimination (Bowne, 1990). The Wilson and Gelb model has the most detailed development and has been applied most widely to suprathreshold pattern discrimination. The quantitative properties of its underlying mechanisms are physiologically plausible (Wilson, Levi, Maffei, Rovamo & DeValois, 1990).

The Wilson and Gelb model contains six spatial-frequency channels, each of which is composed of the difference of two or three Gaussian-shaped spatial weighting functions. These channels encode a pattern at each point of the visual field (Wilson & Gelb, 1984; Wilson, MacFarlane & Phillips, 1983; Wilson, 1986). The model specifies the non-linear contrast response of the channels and imposes rules for computing discriminability between two stimuli.

*Badcock used two definitions for local contrast. One is equivalent to our peak–valley definition, and the other uses three adjacent peaks and valleys as $(a - c)/(a + b + c)$, similar to our peak–peak local contrast.

METHODS

Apparatus and stimuli

Vertical sine-wave gratings (single or compound) were presented on a Joyce Electronics CRT display by Z-axis modulation. The display had a raster frequency of 100 kHz, a non-interlaced frame rate of 100 Hz, a P31 phosphor, an unmodulated luminance of 170 cd/m^2, and a dark surround. Luminance levels were kept within the CRT's linear range.

Separate sinusoidal waveforms were computed digitally for the two components of the compound gratings on an LSI-11/23 computer. The digital sine waves were converted to voltages by two 12-bit D/A converters, passed through programmable dB attenuators, electronically added, and then fed to the Z-axis input.

The stimuli were compound sine-wave gratings consisting of the sum of a fundamental (F) and the third harmonic ($3F$) or fifth harmonic ($5F$). The phase relationship was peaks-add or peaks-subtract. The component contrasts were yoked together so that a contrast increment in one component was accompanied by an equal decrement in the other. As a result, the sum of the component contrasts remained constant at 0.5. In the peaks-add phase, the yoking ensured that the overall Michelson contrast of the stimuli remained constant and did not provide a cue for discrimination. Figure 1 shows a series of these patterns. In this paper, we use the Michelson definition of contrast of the sine wave components: $C = (L_{max} - L_{min})/(L_{max} + L_{min})$.

Procedure

Designating the sine-wave components A and B, and their contrasts C_A and C_B, the subject's task was to discriminate between compound gratings with contrast pairs of (C_A, C_B) and $(C_A - \Delta C, C_B + \Delta C)$. The experiment consisted of measuring just-noticeable differences in the yoked contrast. For example, if a stimulus $(C_A = 0.4, C_B = 0.1)$ was just noticeably different from another stimulus with $(C_A = 0.38, C_B = 0.12)$, the discrimination threshold was 0.02. Measurements of this kind provided a good test bed for examining the models because, as will become apparent, they yielded substantially different predictions.

In the first experiment, the compound gratings were

$F + 3F$. The selection of components was constrained by the frequency ratio of 1:3 and a requirement of about equal contrast sensitivities. These constraints were satisfied by choosing frequencies straddling the peak of the contrast sensitivity function, 2.1 and 6.3 c/deg for two observers, and 1.7 and 5.1 c/deg for a third. Separate yoked-contrast thresholds were measured for 10 contrast pairs of F and $3F$ in steps of 0.05 as follows (0.50, 0.00), (0.45, 0.05), ..., (0.05, 0.45). Separate measurements were made for peaks-add and peaks-subtract phases.

In the second experiment, we used compound gratings with a 1:5 frequency ratio to study performance with a wider frequency separation. In the third experiment, the compound gratings also had a 1:5 frequency ratio, but with low frequencies of 0.13 and 0.65 c/deg. The third experiment addressed the claim that detection and discrimination at low spatial frequencies are mediated by a gradient feature (Campbell *et al.*, 1981).

The display was split in two by a 1-cm-wide vertical black bar on the screen. The left and right halves subtended 3.3 (H) × 4.0 (V) deg at a viewing distance of 230 cm. Pattern discrimination thresholds were measured by a temporal two-alternative forced-choice (2AFC) staircase procedure. In each observation interval, a pattern was presented on each half screen. In one interval, the two patterns were identical compound gratings with component contrasts (C_A, C_B), termed reference patterns. In the other interval, a reference compound grating (C_A, C_B) appeared on one half screen, and a target compound grating on the other. The target was identical to the reference except its component contrasts were ($C_A - \Delta C$, $C_B + \Delta C$). The high frequency component was always incremented, and the lower frequency component decremented. The observer identified the interval in which the two patterns were different, and received feedback. The target appeared with equal probability in the first or second interval and on the left or right half screen. The stimuli were presented for 750 msec with abrupt onsets and offsets. The stimulus intervals were separated by 300 msec. The phases of the waveforms were fixed relative to the edges of the screen.

A three-down one-up adaptive staircase was used to find threshold values (79%-correct criterion) of ΔC (Levitt, 1971). The step size was 1 dB. The staircase stopped after nine reversals. The values of ΔC at the last four reversals were averaged to estimate a pattern discrimination threshold. Data points in the figures are the geometric means of four thresholds. The error bars show 95% confidence intervals. The confidence intervals were calculated from a pooled SE, since the variances among the data points were uniform. (For observer JR, we computed SEs for each data point because the variances weren't uniform.)

In a single session of the yoked-contrast experiment, thresholds were measured for each combination of component contrasts in random order. For a given experiment, each observer participated in eight sessions: four repetitions of the threshold measurements for each of two phase relations.

Prior to the yoked-contrast experiment, contrast sensitivity functions (CSFs) were measured for each observer (1–12 c/deg) using the Quest procedure (Watson & Pelli, 1983). The CSFs were used to select pairs of frequencies with equal sensitivities for the yoked-contrast experiment. Once this selection was made, contrast discrimination functions were measured for each of the two sine-wave gratings. The pedestal contrasts were 0, 0.01, 0.02, 0.04, 0.08, 0.16, 0.32, and 0.50. (In the $F + 5F$ experiment, the maximum pedestal contrast was either 0.25 or 0.32.) The procedures and apparatus were identical to those described above for the yoked-contrast experiment.

Observers

There were six observers. Three participated in the $F + 3F$ and $F + 5F$ experiments, and one in the low-frequency $F + 5F$ experiment. HA is an author. The others were naive to the purposes of the experiments. All were emmetropic or had corrected-to-normal visual acuities. All observers had more than 10 hr of practice before final data collection. Viewing was binocular with natural pupils in a dark room.

Model predictions

The predictions rely on data from contrast discrimination functions (CDFs) for sine-wave gratings. Because CDFs are very similar for stimuli with equal detection thresholds, we chose a pair of sine waves with the appropriate frequency ratio and about equal contrast sensitivities. This selection was based on CSF measurements. The pair of frequencies spanned the peak of the CSF. We then measured CDFs, ΔC vs C, for these sine waves. The same sine-wave gratings were used in the compound-grating experiments, and the CDFs were used to predict thresholds from the four models as follows:

Peak–valley local-contrast model. A pattern has a sequence of luminance peaks and valleys, L_a, L_b, L_c, etc. We can compute the local contrast for all adjacent pairs (i.e. for a and b, b and c, c and d, etc, see Fig. 1). For sine-wave targets all the local contrasts are identical. The CDFs for sine waves can be used to identify the increment threshold ΔC associated with any particular local contrast C. In the case of compound gratings, we computed all peak–valley local contrasts across the stimulus (Michelson contrast). The smallest value of yoked-contrast change ΔC yielding a threshold change for any local contrast value was taken as the model's discrimination threshold. (We based the prediction on the CDFs of the higher-frequency component, but since the CDFs were very similar for the two components, the choice is not critical.) In some cases, a yoked-contrast change resulted in the creation of a local-contrast feature not present in the stimulus prior to the change (e.g. emergence of the shoulder, and the notch evident in Fig. 1). The emergent feature was considered to have zero peak–valley contrast prior to the increment, and the threshold prediction was based on the contrast detection thresholds for the component sine waves.

Peak–peak local-contrast model. For sine waves, all peaks have the same luminance, but for a compound grating an adjacent pair can have different luminances (see Fig. 1). The difference can be used as a cue for pattern discrimination. Subjectively, this amounts to basing the discrimination on the relative brightness of adjacent bars. We computed peak–peak local-contrast values (Michelson definition) from the luminances of adjacent peaks in the waveform. For purposes of deriving predictions from sine-wave CDFs, we assumed that the discrimination properties of the peak–peak local-contrast mechanism were the same as the peak–valley mechanism.

Gradient model. Two just-discriminable gratings have contrasts of C and $C + \Delta C$, but they also have maximum luminance gradients of G and $G + \Delta G$. On the assumption that gradient features determine performance, the CDF provides the relationship between ΔG and G. For the compound gratings, we calculated all local gradient maxima across the waveform, and the change in yoked contrast required to produce a threshold change in these gradients. When the compound gratings were dominated by the fundamental component, F (i.e. had the same number of peaks and valleys), gradient predictions were based on the ΔG values from the corresponding sine-wave CDF. When the compound grating was dominated by the $3F$ (or $5F$) component (i.e. same number of peaks and valleys), ΔG values were obtained from the CDFs for these higher frequencies.

Wilson and Gelb (1984) model. We implemented this model in a computer simulation written in C on a Macintosh computer. We chose parameters of the model with the goal of optimizing its fit to our data.

For each observer, we customized sensitivity parameters of the six mechanisms [A in equation (2) of Wilson and Gelb (1984, p. 125)], and a parameter of the non-linear transducer function [ε in equation (6) of Wilson and Gelb (1984, p. 125)]. The rest of the parameters were taken from Wilson and Gelb (1984). The sensitivity parameters were selected as follows: a parabola was fit to the CSF data of each observer. From this curve, we obtained contrast sensitivities at the peak frequencies of the six mechanisms. We adjusted sensitivity parameter A for each mechanism so that the model would yield an accurate prediction for the observer's contrast sensitivity at the mechanism's peak frequency. We used the spatial summation rule described by Bergen, Wilson and Cowan (1979, their Fig. 9). For pooling across mechanisms, the exponent Q was set at 6, based on Wilson and Gelb (1984).*

The values of ε given by Wilson and Gelb produced lower slopes for CDFs than we obtained. Accordingly, we adjusted ε to give a best fit (least-squares criterion) to our sine-wave CDFs in the range 0–50% contrast. We obtained values of ε between 0.5 and 0.7 for all mechanisms, similar to values used by Yager and Kramer (1991). Figure 2 shows illustrative fits of the model to contrast-discrimination data, and indicates that the model was well calibrated for these observers.

Parameters derived from the CSF and CDF data were fixed and used to predict the results of the compound-grating experiments.

*In fitting the compound-dating data, we tried varying Q from 1 to 100. Although fits with higher values of Q were sometimes slightly better, the difference from the fit with 6 was negligible.

RESULTS

$F + 3F$ condition

Figure 3 shows data and predictions for three observers. The spatial frequencies of the components F and $3F$ were 2.1 and 6.3 c/deg for HA and KS, and 1.7 and 5.1 c/deg for RA. Discrimination threshold ΔC is plotted as a function of the contrast of the $3F$ components. (Recall that the sum of the contrasts of F and $3F$ is fixed at 0.50.) There are separate panels for peaks-add and peaks-subtract data. There was a significant effect of the contrast condition for all observers ($P < 0.01$).

The results are clearest for HA's data, shown in the top two panels of Fig. 3. The Wilson and Gelb model provided a fairly good fit to his peaks-add data, but not to the peaks-subtract data. This model predicted much lower thresholds in the peak-subtract case where it undershot the data across most of the range. The model's peaks-subtract curve had two deep dips (i.e. very low thresholds) for which there was no evidence in the data.

None of the local-feature models accounted for the full range of HA's peaks-add data. The gradient model fit the data for $3F$ contrasts from 0 to 10%, the gradient and peak–valley models fit the data from 20% to 35%, and the peak–peak model fit the data from 35% to 45%. Near 15%, the data varied smoothly, but all three local-contrast models had major excursions.

Except near 15% contrast, it is roughly the case that HA's peaks-add data follow the lower envelope of the local-feature curves. This raises the possibility of a multiple local-feature model in which the observer has simultaneous access to all three local features.

Similarly, for HA's peaks-subtract data, none of the local-feature models worked across the entire range. Predictions of the gradient and peak–valley models were similar and accounted for the data for $3F$ contrasts from 0 to 25%. The peak–peak model fit the data from 35% to 45%. Although the data roughly follow the lower envelope of the feature models, the points lie below all the models near 30%.

The results were qualitatively similar for observer RA (Fig. 3). The Wilson and Gelb model fit the peaks-add data quite well but not the peaks-subtract data. None of the local-feature models handled all the data. For peaks-add, the gradient model again worked at low contrasts of $3F$, and the peak-peak model worked for high contrasts (over a wider range than for HA). For peaks-subtract, all of RA's data were quite close to the peak–peak prediction. Perhaps RA relied primarily on a single local feature, the peak–peak local contrast, reverting to the gradient when the peak–peak cue was not present.

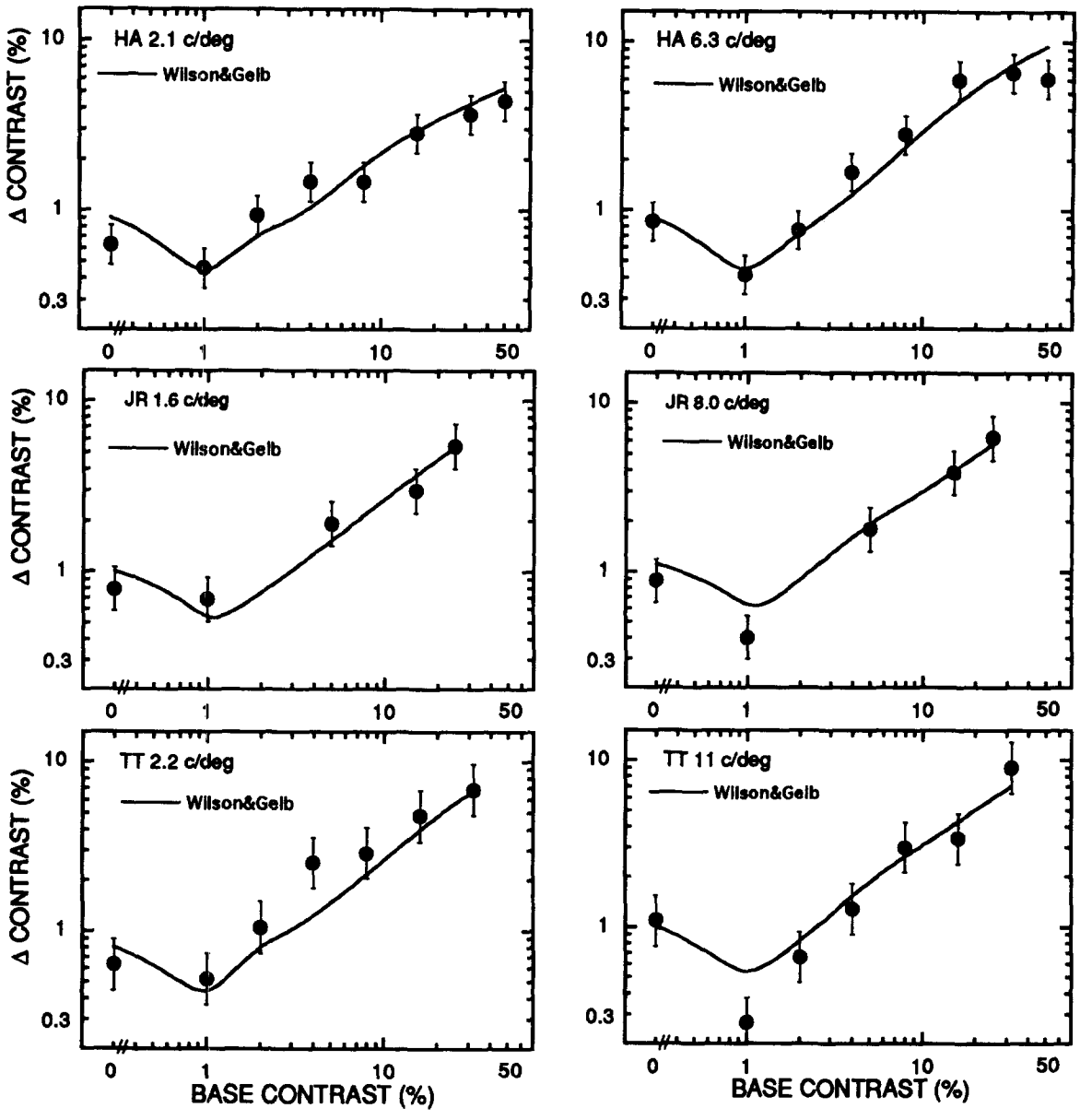

FIGURE 2. Contrast discrimination functions (●) and the predictions from the Wilson and Gelb model are shown for three observers (HA, JR and TT). Plots are given for simple sine-wave gratings at two spatial frequencies. The error bars indicate 95% confidence intervals.

Observer KS's data differ from the others in two ways. The peak–peak model does not provide a good fit to the peaks-add data for contrasts of 30–40%. For peaks-subtract contrasts of 30–40%, thresholds are lower than predicted by any of the local-feature models.

F + 5F condition

Figure 4 shows data and predictions for three observers. The spatial frequencies of the components F and $5F$ were 1.7 and 8.5 c/deg for HA, 1.6 and 8.0 c/deg for JR, and 2.2 and 11.0 c/deg for TT. Discrimination threshold ΔC is plotted as a function of the contrast of $5F$. There was a significant effect of the contrast conditions for all observers ($P < 0.01$).

The results are similar to those for $F + 3F$. The Wilson and Gelb model provided better fits to the peaks-add data than the peaks-subtract data, although the difference was not so striking. None of the local-feature models accounted for the full range of data. But, as in $F + 3F$ compounds, data for low contrasts of $5F$ were roughly consistent with the gradient model and data for high contrasts with the peak–peak model. There were individual variations in the contrast ranges over which these two local-feature models worked.

F + 5F condition—low spatial frequencies

One observer (LW) was tested with low component frequencies of 0.13 and 0.65 c/deg. His data are shown in Fig. 5 along with the predictions of the models. The data points are the means of two threshold measures rather than four. The results are qualitatively similar to those in the other conditions: the Wilson and Gelb model fit the peaks-add data better than the peaks-subtract data; none of the local-feature models could

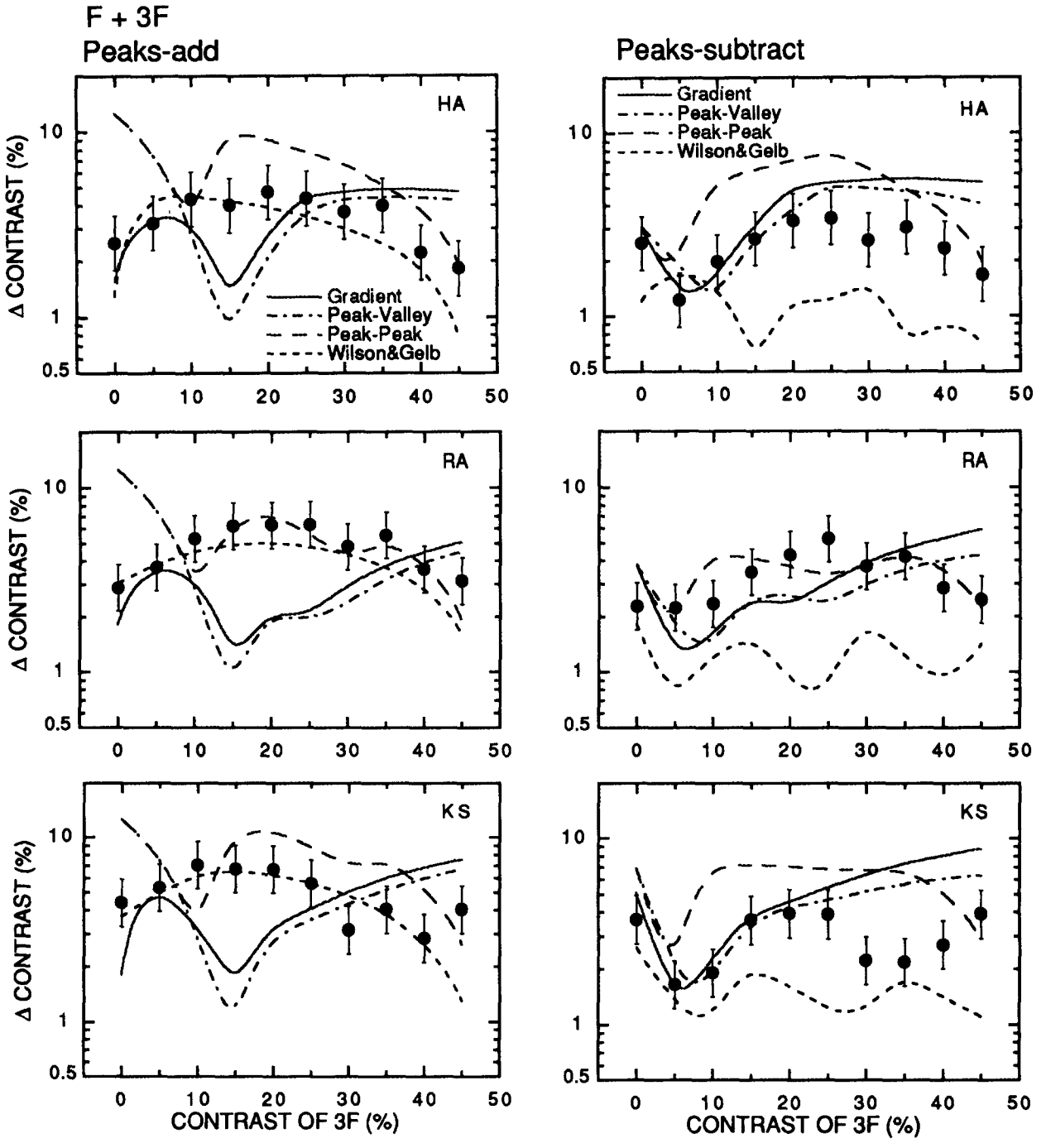

FIGURE 3. Discrimination thresholds for compound gratings consisting of F and $3F$. The sum of the component contrasts was constant at 0.5. The vertical scale shows the yoked threshold change in contrast (increment of $3F$ and decrement of F). The data in peaks-add and peaks-subtract phases are shown separately for each observer, along with predictions from the Wilson and Gelb (1984) model and the three local-feature models. Error bars indicate averaged 95% confidence intervals of all yoked conditions.

account for all of the data, but the gradient and peak–peak models together covered most of the data. The data don't support the view that the gradient feature is more salient at very low frequencies.

DISCUSSION

The Wilson and Gelb model

The Wilson and Gelb model accounted for the results with peaks-add gratings, but failed to predict those with peaks-subtract gratings. Why this difference?

Recall that the Wilson and Gelb model assumes three spatially adjacent sensors for each channel with the middle one aligned with the center of the pattern. For peaks-add patterns, the two components of the compound grating superimpose additively for all sensors, producing a strong response in all of them. The sensor responses are high, yielding fairly high thresholds,

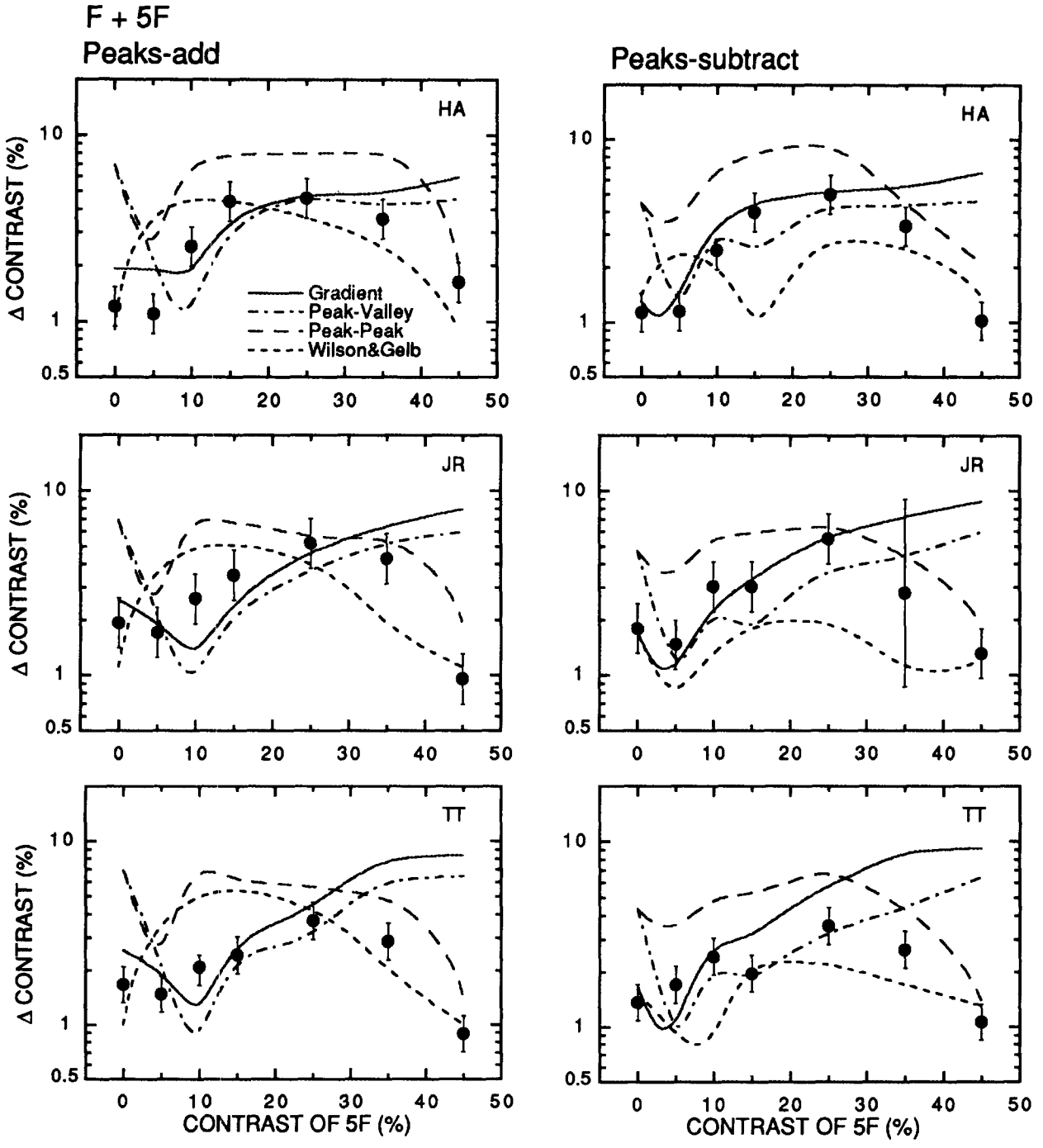

FIGURE 4. Discrimination thresholds for compound gratings of F and $5F$. The data and predictions from the models are shown for three observers (HA, JR and TT). Error bars indicate averaged 95% confidence intervals of all yoked conditions (for JR the values are not average).

consistent with the data. On the other hand, in the peaks-subtract conditions, the two components (F and $3F$) elicit responses in opposite directions (positive and negative responses) for sensors which are sensitive to both components. There are certain contrast combinations at which the net responses of the sensors are low. Low responses of the model sensors occur in the accelerating portion of the nonlinear transducer function where small contrast differences yield large response differences. These sensors account for the model's low contrast-discrimination thresholds for some contrast combinations in the peaks-subtract condition, viz. the ($1F = 25\% + 3F = 25\%$) and ($1F = 10\% + 3F = 40\%$) conditions. The data do not exhibit these very low thresholds.

Even for peaks-add phase, there are portions of the waveform of a $1F + 3F$ compound grating in which the components have opposite signs. For sensors centered at these points and responsive to both components, the opposing inputs would yield low responses, and low contrast-discrimination thresholds in the model. This reasoning led us to predict that if the sensor arrays in the

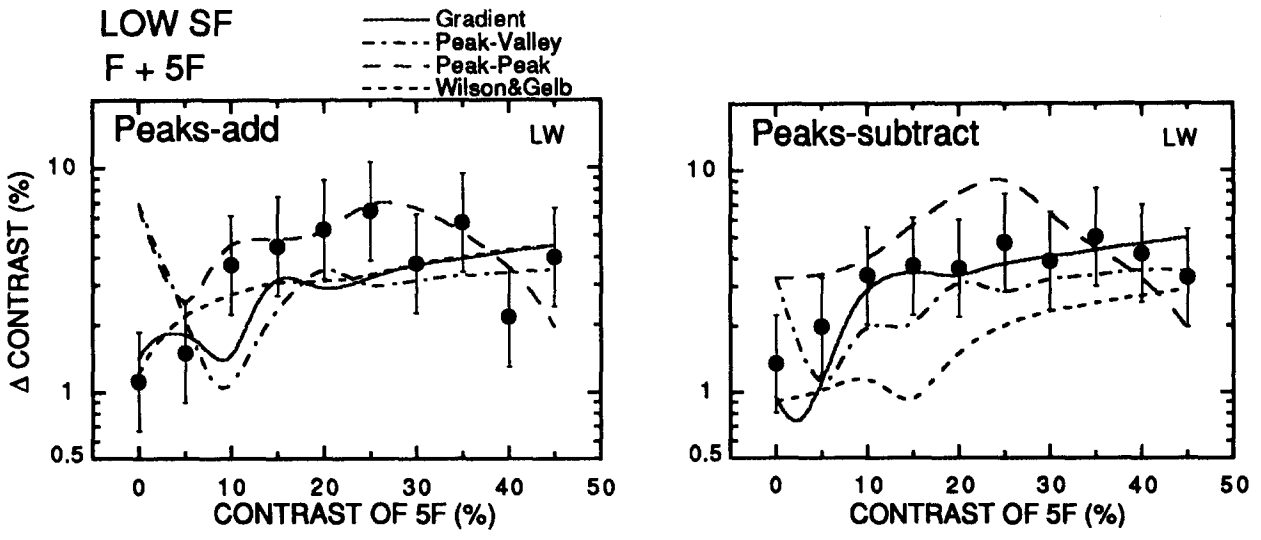

FIGURE 5. Discrimination thresholds for compound gratings of F and $5F$ with low spatial frequencies (0.13 and 0.65 c/deg) for observer LW. The data and predictions from the models are shown. Error bars indicate averaged 95% confidence intervals of all yoked conditions.

Wilson and Gelb model were extended to cover a full cycle of the $1F$ component (instead of three sensors clustered near the peak), the undesirable low thresholds might also appear in the peaks-add condition. We simulated this extension of the model. We increased the number of sensors from three to five, and also 11, using the same sensor spacing parameters used in the Wilson and Gelb model. Five sensors are sufficient to cover a half cycle at the peak frequency of the sensor, with one sensor at the center and two on either side. Eleven sensors cover a full cycle.

For both the five- and 11-sensor cases, predicted discrimination thresholds were low for both peaks-add and peaks-subtract conditions (between 0.5% and 2.0% contrast) and the pattern of threshold change was not consistent with our observed data.

Another way of amending the Wilson and Gelb model is to include sensors in sine phase: adding sensors with sine-phase receptive fields (90 deg out of phase with the cosine-phase receptive fields of the standard model). These additional sensors don't solve the problem of low thresholds in the peaks-subtract case. This is because all mechanisms are assumed to be independent and the contrast-discrimination threshold is still determined by the cosine-phase sensors with opposed inputs from the two components.

It appears that to avoid the unacceptably low thresholds predicted by the model, some form of inhibition between sensors, possibly including contrast gain control, would have to be included.

Local-feature models

None of the local-feature models explained the results by themselves. However, most of the data fell close to an envelope composed of the lowest thresholds of the three feature-detector models, with some exceptions. Our results immediately raise two questions about the local-feature models.

First, why does a particular local-feature model fit the data for only a limited range of contrast combinations? Adopting a minimum threshold principle, when two or more feature detectors have different thresholds for a given stimulus, the feature producing the lowest threshold determines performance. We would expect the envelope of lowest thresholds of the feature detectors to determine discrimination performance. This would require an observer to switch features in the discrimination task, depending on which one was most sensitive in a given condition. This envelope model accounts for most of the results in Figs 3 and 4.

Second, why doesn't any local-feature model account for some data points, such as $3F = 0.15$ in peaks-add phase? To address this question, we turn to contrast gain control.

Contrast gain control

A problem that is common to the channel model and the local-feature models is the prediction of very low discrimination thresholds that are not observed empirically. This behavior reflects the existence in the models of accelerating portions in the contrast–response functions when the sensors are weakly responding. The models are built like this because contrast-discrimination functions for sine-wave gratings (and other patterns) show a facilitation effect for near-threshold pedestals (Nachmias & Sansbury, 1974; Foley & Legge, 1981). In the Wilson and Gelb model, sensors may operate on the accelerating portion of the contrast–response function when the two components of the compound grating produce roughly equal and opposite responses in the front-end linear filter. In the local-feature models, a similar problem occurs when a low-contrast emergent feature appears in the compound waveform. One way of preventing model sensors from achieving high-sensitivity operation under these conditions is to introduce some form of mutual inhibition among sensors. Models of

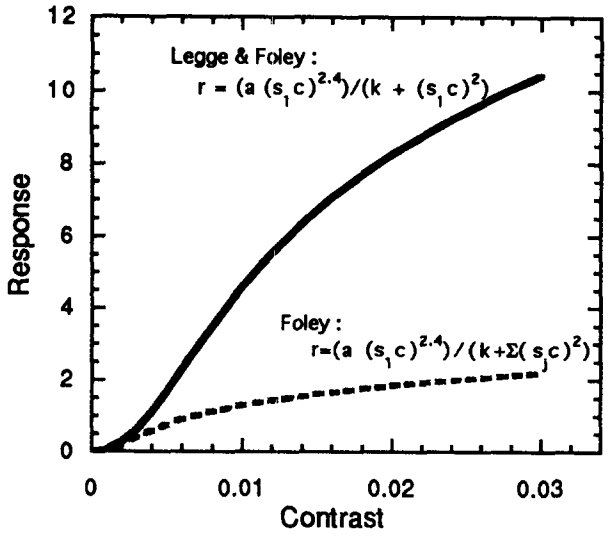

FIGURE 6. Contrast responses of a channel are shown for an independent channel model (Legge & Foley, 1980) and a mutually inhibitory channel model (Foley's model 2, 1994). In the equations, R is the channel response, C is stimulus contrast, and a and k are constants with values taken from Legge and Foley (1980). The additional parameters s and j in the Foley's equation are sensitivity parameters of the channels: $s_1 = 1$, $s_2 = 2$, $s_3 = 2$. (We chose three channels for illustration purposes only.) Note that the two models have the same equation with the exception of the pooling term in the denominator of the Foley model.

contrast-gain control may achieve this goal (Albrecht & Geisler, 1991; Foley, 1994; Heeger, 1991; Solomon, Sperling & Chubb, 1993; Wilson & Humanski, 1993). In the following paragraphs, (1) we show that contrast-gain control models can reduce or eliminate the high-sensitivity associated with weakly responding sensors; and (2) we provide empirical evidence for inhibitory effects of the sort required by the models.

Foley (1994, p. 1711) has suggested that when a linear channel's output is modified by divisive inhibition from the summed outputs of other channels, an accelerating contrast–response function can become negatively accelerated.

Figure 6 shows two contrast–response functions. The one from the Legge and Foley (1980) model is similar to the Wilson and Gelb model and does not include contrast gain control. The other is from Foley (1994) and includes divisive inhibition. The parameters of the models are described in the figure caption. In the Legge and Foley model, like the Wilson and Gelb model, the channel's contrast–response function is independent of the responses of other channels. In the contrast gain control model, such as Foley's (1994) model, a channel's response is inhibited by pooled responses from other channels. The key point to note in Fig. 6 is that conditions can be selected in which parameters yielding an accelerating nonlinearity in the absence of divisive inhibition convert to a negatively accelerated nonlinearity in its presence. This could account for the desensitization of sensors for patterns in which other sensors are highly responsive.

Contrast gain control may also help the local-feature models. The largest discrepancy between predictions and observed data occurred at $3F = 15\%$ contrast in peaks-add phase (see Fig. 3); the peak–valley and gradient models predicted very low thresholds. The luminance profile of this pattern is close to the luminance profile in Fig. 1 for $1F = 30 + 3F = 20$, peaks-add. The peak–valley local contrast or gradient values in the region b–c in the figure determine the model thresholds for this contrast combination. For example, small differences in the low peak–valley contrast is enough to exceed the model's threshold. But notice that this low-contrast local feature is flanked by a large luminance peak (a) and a large valley (d). Strong signals from adjacent high-contrast local features may desensitize the peak–valley detector responding to b–c.

In a brief control experiment, we looked for evidence of lateral masking of this sort in human performance. We measured contrast discrimination thresholds with and without flanking bars for a stimulus with similar luminance profile. The target for discrimination was one cycle of a vertical sine wave. The flanking bars consisted of half cycles of a cosine (positive on the target's left,

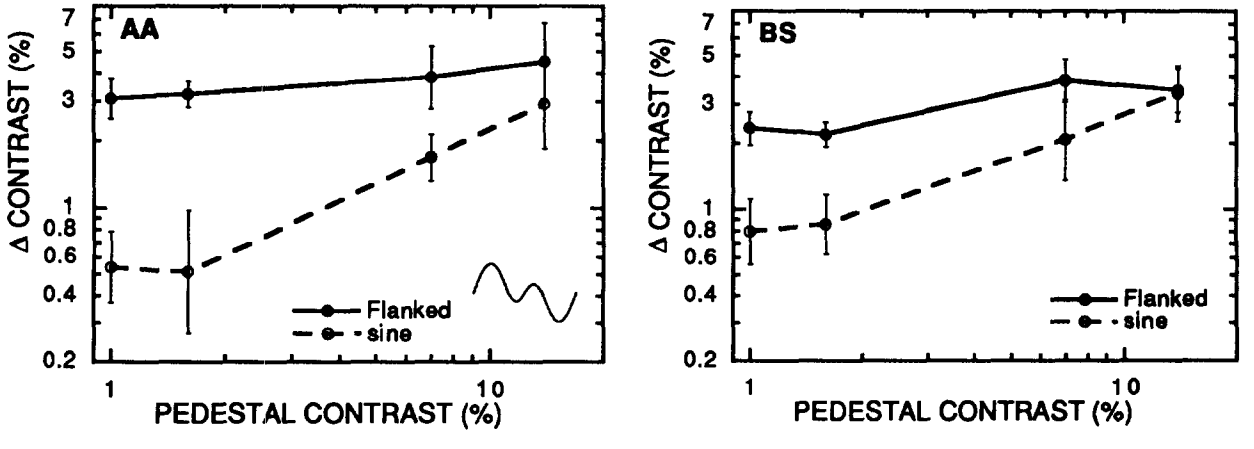

FIGURE 7. Contrast-discrimination thresholds for a one-cycle sine-wave pattern with and without flanking bars. The flanking bars were half cycles of a 50%-contrast cosine pattern (see the inset). Error bars indicate 95% confidence intervals.

negative on the right). See the inset in Fig. 7. The spatial frequencies were 2 c/deg for the flanking cosine, and 4 c/deg for the central sine. The choice of frequencies was governed by the attempt to approximate the local region of waveform a, b, c, d (Fig. 1) in the $1F = 30 + 3F = 20\%$ condition. The contrast of the flanking cosine bars was fixed at 50%. The pedestal contrast of the sine wave was varied from 1.0% to 14%.

Figure 7 shows the results for two observers. The results clearly exhibit the threshold elevation due to the flanking grating in the case of low pedestal contrasts (which correspond to our $1F = 35 + 3F = 15\%$, and $1F = 30 + 3F = 20\%$ compounds). The results of this brief experiment illustrate that if local features are responsible for pattern discrimination, the corresponding neural representations are vulnerable to lateral masking effects from spatially adjacent pattern features. The local-feature models must invoke some form of lateral interaction, possibly contrast gain control.

CONCLUSION

In all three of our experiments, the Wilson and Gelb channel model accounted fairly well for the peaks-add data but not for the peaks-subtract data. None of the three local feature models by itself fit the data. A simple summary is that the gradient model (or peak–valley model) fit the compound-grating data for a range of low contrasts of $3F$ or $5F$ (and high contrasts of F). The peak–peak model fit the data for high contrasts of $3F$ or $5F$ (and low values of F). The results were less consistent for compound gratings having components of nearly equal contrast. The results are qualitatively consistent with a multiple-cue model in which observers can switch from reliance on one feature to another. This interpretation is consistent with the findings of Hess and Pointer (1987).

All of the models showed a characteristic failure; they predicted very low thresholds for some compound gratings that were not present in the data. These were cases in which the models used weakly responding mechanisms, operating in a sensitive portion of their response-vs-contrast curves. Humans do not appear to have access to these sensitive mechanisms in suprathreshold pattern discrimination. One possibility is that these mechanisms are desensitized by an inhibitory signal from neighboring mechanisms with strong responses to high contrast pattern features.

REFERENCES

Albrecht, D. G. & Geisler, W. S. (1991). Motion selectivity and the contrast response function of simple cells in the visual cortex. *Visual Neuroscience, 7*, 531–546.

Badcock, D. R. (1984a). Spatial phase or luminance profile discrimination? *Vision Research, 24*, 613–623.

Badcock, D. R. (1984b). How do we discriminate relative spatial phase? *Vision Research, 24*, 1847–1857.

Bennett, P. J. & Banks, M. S. (1991). The effects of contrast, spatial scale, and orientation on foveal and peripheral phase discrimination. *Vision Research, 31*, 1759–1786.

Bergen, J. R., Wilson, H. R. & Cowan, J. D. (1979). Further evidence for four mechanisms mediating vision at threshold: Sensitivities to complex gratings and aperiodic stimuli. *Journal of the Optical Society of America, 69*, 1580–1586.

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

Burbeck, C. A. (1987). Position and spatial frequency in large scale localization judgments. *Vision Research, 27*, 417–427.

Campbell, F. W. & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology, London, 197*, 551–566.

Campbell, F. W., Johnstone, J. R. & Ross, J. (1981). An explanation for the visibility of low frequency gratings. *Vision Research, 21*, 723–730.

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

Foley, J. M. & Legge, G. E. (1981). Contrast detection and near-threshold discrimination in human vision. *Vision Research, 21*, 1041–1053.

Graham, N. & Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: A comparison of single-channel and multiple-channels models. *Vision Research, 11*, 251–259.

Heeger, D. (1991). Nonlinear model of cat striate cortex. In Landy, M. & Movshon, J. A. (Eds), *Computational models of visual processing* (pp. 111–135). Cambridge, Mass.: MIT Press.

Hess, R. F. & Pointer, J. S. (1987). Evidence for spatially local computations underlying discrimination of periodic patterns in fovea and periphery. *Vision Research, 27*, 1343–1360.

Klein, S. A. & Levi, D. M. (1985). Hyperacuity thresholds of 1 sec: Theoretical predictions and empirical validation. *Journal of the Optical Society of America A, 2*, 1170–1190.

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

Levitt, H. (1971). Transformed up–down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*, 467–477.

Marr, D. & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B, 207*, 187–217.

Morrone, M. C. & Burr, D. C. (1988). Feature detection in human vision: A phase-dependent energy model. *Proceedings of the Royal Society of London B, 235*, 221–245.

Nachmias, J. & Sansbury, R. V. (1974). Grating contrast: Discrimination may be better than detection. *Vision Research, 14*, 1039–1042.

Solomon, J. A., Sperling, G. & Chubb, C. (1993). The lateral inhibition of perceived contrast is indifferent to on-center/off-center segregation, but specific to orientation. *Vision Research, 33*, 2671–2683.

Watson, A. B. (1983). Detection and recognition of simple spatial forms. In Bradick, O. J. & Slade, A. C. (Eds), *Physical and biological processing of images* (pp. 100–114). Berlin: Springer.

Watson, A. B. & Pelli, D. G. (1983). QUEST: A Baysian adaptive psychometric method. *Perception & Psychophysics, 33*, 113–120.

Watt, R. J. & Morgan, M. J. (1985). A theory of the primitive spatial code in human vision. *Vision Research, 25*, 1661–1674.

Wilson, H. R. (1986). Responses of spatial mechanisms can explain hyperacuity. *Vision Research, 26*, 453–469.

Wilson, H. R. & Gelb, D. J. (1984). Modified line-element theory for spatial-frequency and width discrimination. *Journal of the Optical Society of America A, 1*, 124–131.

Wilson, H. R. & Humanski, R. (1993). Spatial frequency adaptation and contrast gain control. *Vision Research, 33*, 1133–1149.

Wilson, H. R. & Richards, W. A. (1989). Mechanisms of contour curvature discrimination. *Journal of the Optical Society of America A, 6*, 873–882.

Wilson, H. R., McFarlane, D. K. & Phillips, G. C. (1983). Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research, 23*, 873–882.

Wilson, H. R., Levi, D., Maffei, L., Rovamo, J. & DeValois, R. (1990). The perception of form: Retina to striate cortex. In Spillmann, L. & Werner, J. S. (Eds), *Visual perception: The neurophysiological foundations* (pp. 231–272). New York: Academic Press.

Yager, D. & Kramer, P. (1991). A model for perceived spatial frequency and spatial frequency discrimination. *Vision Research, 31*, 1067–1072.

Acknowledgements—We thank Yuanchao Gu and Andrew Luebker for their technical assistance. We also thank Hugh Wilson for providing us with the source code for his simulation. We thank J. S. Mansfield for comments on this manuscript. A portion of this research was first presented at the annual meeting of the Association for Research in Vision and Ophthalmology in May 1990. This research was supported by National Institutes of Health grant Ey-02857 to Gordon E. Legge.