

(Ciba Foundation Symposium
1994, p 309-329)

Non-linearities in texture segregation

Norma Graham

Department of Psychology, Columbia University, New York, NY 10027, USA

Abstract. The existence of complex (non-Fourier, second-order) channels is suggested by some characteristics of segregation perceived between regions distinguished by visual texture. These complex channels consist of two linear-filtering stages separated by a rectification-type non-linearity. We have investigated (i) the spatial frequency selectivity and orientation selectivity of their first-stage filters; (ii) the relationship between the preferred values of orientation and spatial frequency at the first and second filters; (iii) spatial pooling and its implications for the non-linearity at the middle of the complex channel; and (iv) the dynamics of complex and simple linear channels. An intensive non-linearity is also necessary to explain perceived region segregation. This intensive non-linearity might arise from an *early local non-linearity* preceding the channels (perhaps retinal light adaptation) or from *normalization* among the channels themselves (perhaps due to intracortical inhibition). Deciding between these two candidates has been more difficult than we had hoped. It appears that: (i) this intensive non-linearity operates for both simple and complex channels; (ii) the effects on it of changing mean luminance or spatial scale may be accounted for by a sensitivity parameter; (iii) it can be dramatically compressive even at contrasts less than 25% for high mean luminances and large scales; and (iv) at even lower contrasts there is an accelerating non-linearity that acts before the second filter of the complex channels.

1994 *Higher-order processing in the visual system*. Wiley, Chichester (Ciba Foundation Symposium 184) p 309-329

Patterns like those shown in Figs 1 and 2 have regions composed of the same elements (in the same proportions) but differently arranged. I will refer to these as 'element-arrangement textures' although they are hardly typical examples of natural-language 'textures'. Using these patterns, Jacob Beck, Anne Sutter and I tested a simple model composed of simple (linear, first-order, Fourier) spatial frequency and orientation-selective channels and found it wanting (Sutter et al 1989, Graham 1991, Graham et al 1992a). Many of our results could be explained by the presence of simple channels, but other results suggested an important role for at least two kinds of non-linearities—a spatial one and an intensive one.

The spatial non-linearity may be modelled by invoking complex channels (non-linear, second-order, non-Fourier processes). Each complex channel consists of two linear filtering stages separated by a rectification-type non-linearity (a

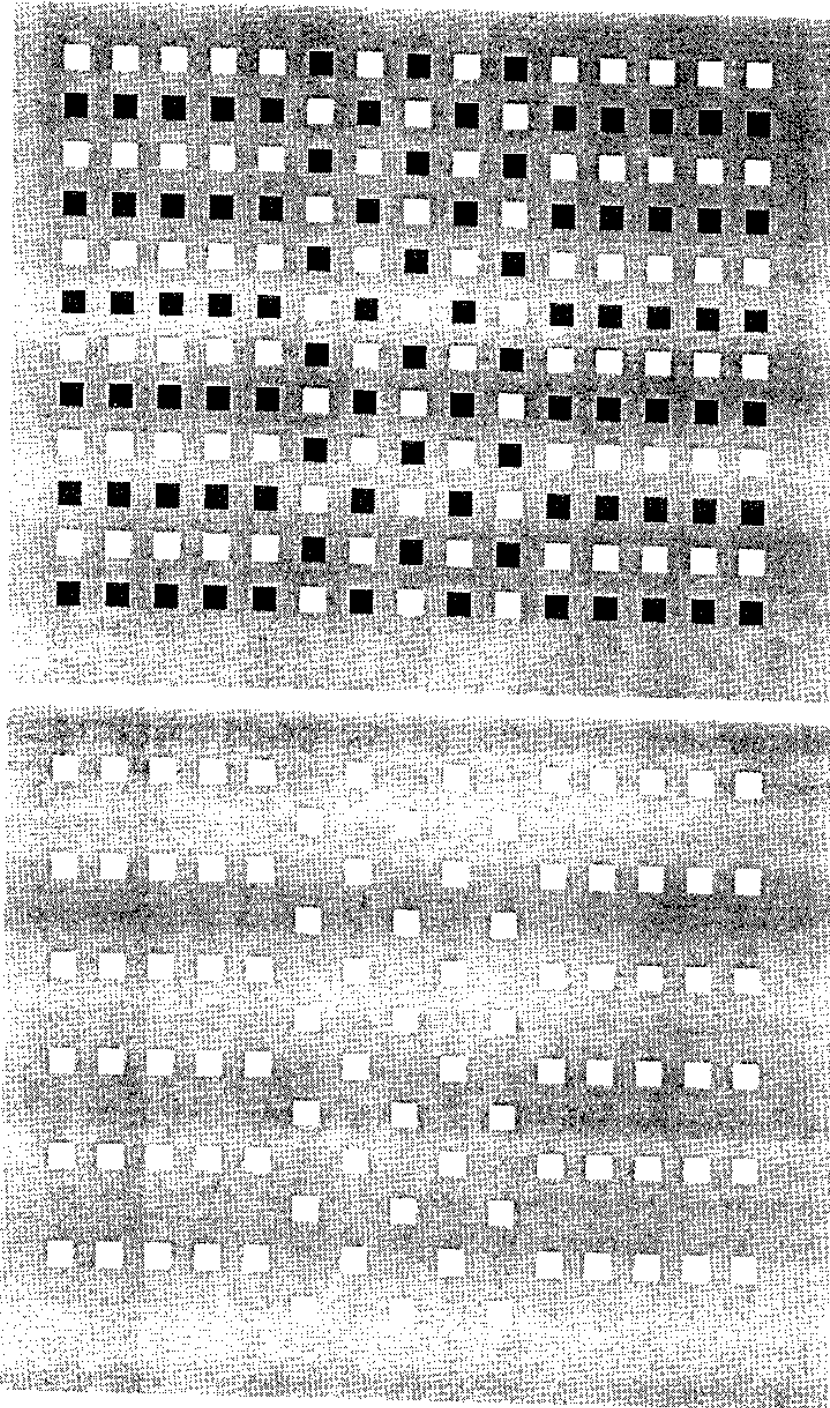


FIG. 1. Two examples of element-arrangement patterns where the elements are squares. In the bottom panel only one type of element is visible (i.e. elements of the other type have zero contrast). Reproduction will have distorted the stimuli in these figures somewhat.

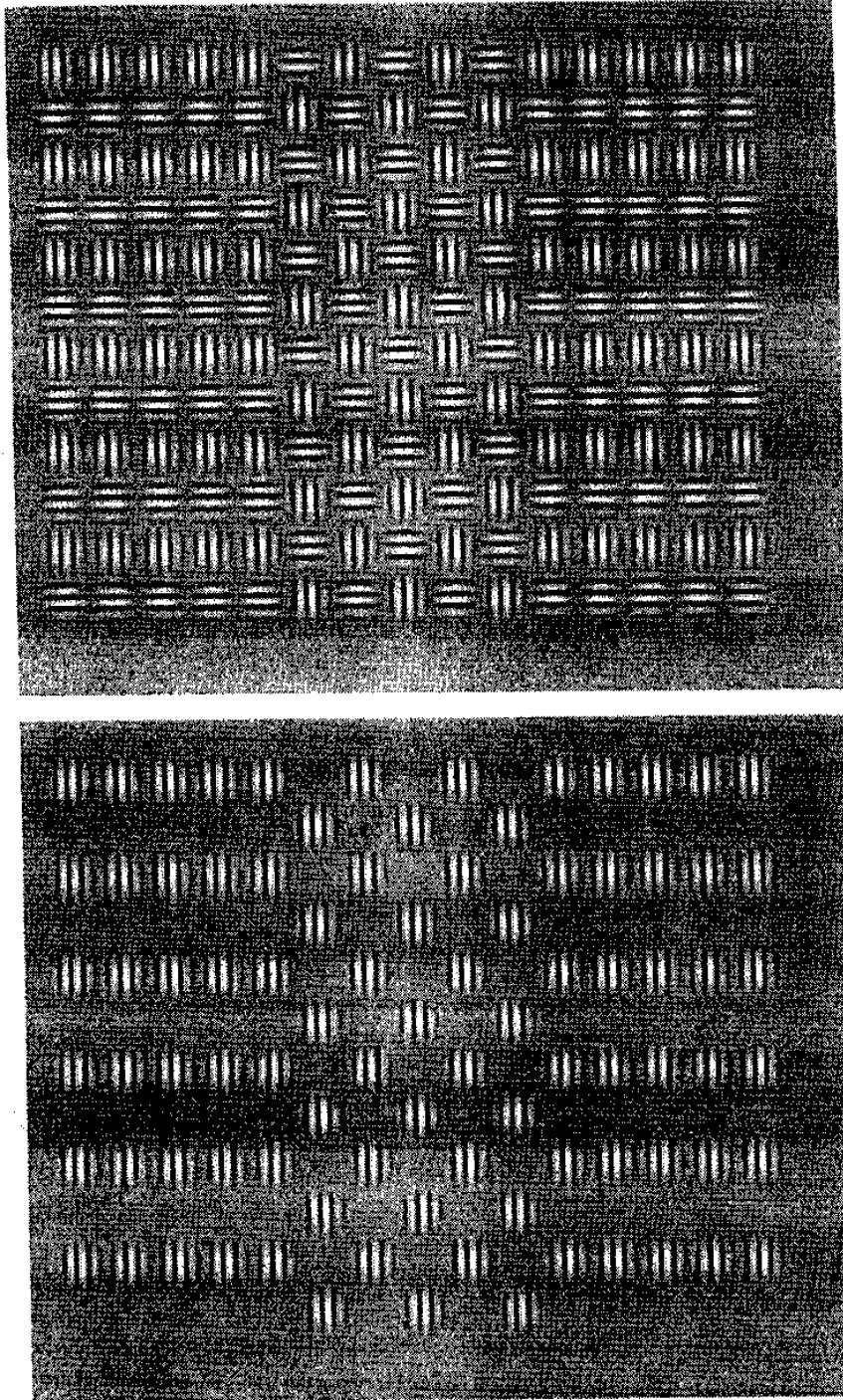


FIG. 2. Two examples of element-arrangement patterns where the elements are Gabor patches. In the bottom panel only one type of element is visible (i.e. elements of the other type have zero contrast). Reproduction will have distorted the stimuli in these figures somewhat.

structure like that hypothesized for a complex cortical cell). Two possible complex channels are illustrated in Fig. 3; they are distinguished by the kind of receptive field characterizing their first stage.

The intensive non-linearity might be modelled either as a relatively local non-linearity preceding the spatial frequency and orientation-selective channels, or as an inhibitory interaction (a normalization network) among the channels themselves. This intensive non-linearity is dramatically compressive even at contrasts below 25%.

Subsequent to the channels themselves, some further processing of the outputs at different positions in different channels must occur to produce the perceptual segregation (or, operationally, the observer's rating of the extent to which the regions segregate immediately and effortlessly). Although these processes are extremely interesting (see, e.g. Nothdurft 1994, this volume), we make no attempt in this work to model these processes. We do need to characterize the information retained in these processes in order to calculate a prediction for the observer's segregation rating from our assumptions about the channels. Thus we have considered a variety of rules for the pooling across spatial positions and across different channels that must enter into this computation. As it turns out, the particular rule makes very little difference to the conclusions here (see Sutter et al 1989, Graham et al 1992a for more details).

In more recent studies, we have tried to characterize further the spatial and intensive non-linearities involved in the perceptual segregation of element-arrangement texture. Our progress to date is briefly summarized here.

About the complex channels in texture segregation

First-filter bandwidths

The property of the complex channel we have investigated in greatest detail is the bandwidth of the first-stage filters on both the spatial frequency and orientation dimensions (Graham et al 1992b, 1993). If, for example, the physiological substrate for these first-stage filters were lateral geniculate nucleus (LGN) neurons, the first stage should show little orientation selectivity (since LGN receptive fields are approximately concentric, as shown in the top panel of Fig. 3). If, however, they were simple cortical cells, one might expect substantial orientation selectivity (as shown in the bottom panel of Fig. 3). To measure these bandwidths, we used patterns composed of Gabor-patch elements (e.g. Fig. 2), varying the difference between the orientations or spatial frequencies of the two types of elements. For a fixed contrast of one element type, the contrast in the other was varied to find the maximal amount of interference. While performance varied dramatically from person to person, the first-stage filters of the complex channels were always selective for both spatial frequency and orientation. In magnitude, the estimated bandwidths of the complex

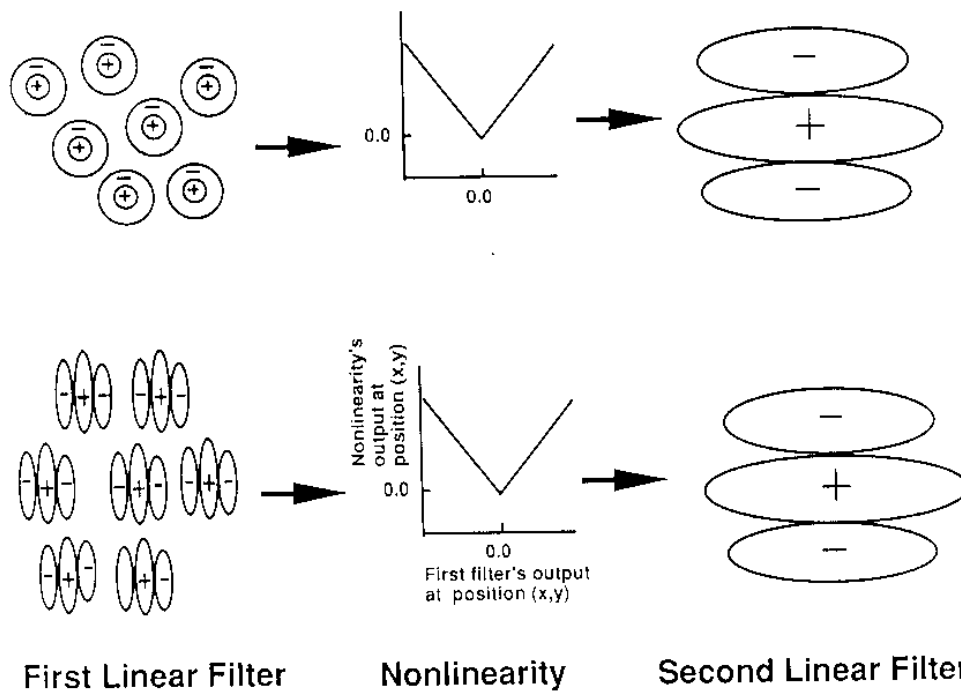


FIG. 3. Diagrams of complex channels with a non-orientation-selective first stage (top panel) and with an orientation-selective first stage (bottom panel).

channels' first-stage filters were always substantially wider than those for simple channels (by at least a factor of two) but narrower than the bandwidth of LGN cells.

Three comments are worth making here. First, to repeat, individuals differ substantially on this task. This is reminiscent of the individual differences recently described by Cannon & Fullenkamp (1993) in the effect of lateral interactions on perceived contrast of grating patches surrounded by other gratings. I never saw individual differences of this magnitude or significance in the near-threshold literature (Graham 1989, 1992).

Second, in studies measuring competition between motion paths, Werkhoven et al (1992, 1993) reported a very different result, namely that there was no spatial frequency or orientation selectivity in the second-order processes serving this motion perception.

Third, the quantitative estimate of the bandwidth depends substantially on the assumed form of intensive non-linearity (although this dependence was not as great as the individual differences). Thus, to estimate the bandwidth more definitively for any individual, we would need to know more than we do about that non-linearity.

Relationship between preferred values at the first and second filters

What is the relationship between the value of spatial frequency or orientation preferred by the first filter and that preferred by the second? For example, do both filters tend to have the same orientation or do they tend to have perpendicular orientations (as in the bottom panel of Fig. 3 and as used to explain phenomena at texture boundaries, e.g. Wilson & Richards 1992)? We used patterns where only one element type was visible and we varied the relationship between the orientation of the grating-patch elements and that of the stripes in the striped region so that the local and global orientations were either identical or perpendicular (the latter is shown in Fig. 2 [bottom]). Both cases segregated easily. Thus we assume that complex channels of both configurations (and probably many others) must exist. In the one observer we studied extensively, there was a slight bias for the identical case over the perpendicular case.

Similarly, many different relationships between the preferred spatial frequencies of the first and second filters must exist. A weak bias in that relationship has been reported previously (using somewhat different patterns) in favour of the configuration where the preferred frequency of the first filter is three or four octaves higher than that of the second (Sutter et al 1991).

The nature of the non-linearity at the middle of the complex channel

There are many questions one might ask about the rectification non-linearity in the middle of the complex channels. The question we know most about is this: is the embedded non-linearity a rectification of the absolute-value type (made up of straight-line segments as drawn in Fig. 3) or is it better modelled as an accelerating power law with an exponent of two or more (e.g. squaring or half-squaring)? We have done a number of spatial-pooling experiments using grating-patch elements of different sizes and contrasts to determine the extent to which contrast differences can compensate for area differences (Graham et al 1992c), analogous to the experiments using square and other aperiodic elements in Sutter et al (1989).

The results of these experiments cannot be explained by complex channels if the embedded non-linearity is an absolute-value type rectification (unless the second filter is no longer assumed to be a linear filter) but may be explained if the middle non-linearity is a power law with an exponent of at least two. (Some uncertainty results from the effects of spectral spread for which we have not yet done sufficient calculations.) Thus, at this point, I would bet in favour of a power law (e.g. squaring or half-squaring, or, in the case of texture, both, as described by Sperling et al 1994, this volume).

Dynamics of processing

We have used a speed-accuracy trade-off paradigm to investigate dynamics by comparing several discrimination tasks done with element-arrangement textures made either of Gabor patches (stimulating complex channels) or of Gaussian blobs where the two element types are of opposite contrast (stimulating predominantly simple channels). The initial results suggest that the dynamics of these complex channels are slower than those of the simple channels (Sutter & Graham 1992).

About the intensive non-linearity in perceived texture segregation

Because the intensive non-linearity is of interest in itself and because knowledge of it is frequently required to draw conclusions about the complex channels, we have undertaken a number of experiments intended to characterize its properties. Most of these experiments used *constant-difference series of stimuli*.

In such a series, the two element types in any given stimulus have the same spatial characteristics but differ in amount and/or sign of contrast (i.e. both element types may be squares that are lighter or darker than the background by various amounts, or both may be grating patches that are 180° out of phase and/or of different contrast). All the stimuli in a given series have elements of the same spatial characteristics (e.g. all squares) and the *difference* between the luminances of the two element types is held constant in a series (hence the name). But the absolute luminances of the two element types vary together from one stimulus to another in the series. In such a series, therefore, patterns exist where both element types have the *same sign of contrast* (e.g. square elements where both types are darker or both are lighter than the background), where one element type has zero contrast so there is *one element type only* apparent (e.g. either dark or, as in Fig. 1, light) and where the two element types are of *opposite sign of contrast* (e.g. dark squares and light squares as in Fig. 1).

These constant difference series are interesting because, according to models containing only simple (linear, first-order) channels, perceived segregation of all patterns in a constant difference series is almost identical. (Briefly, perceived segregation is predicted to be close to identical because the energy at the fundamental frequency is close to identical. Further explanation can be found in Graham 1991 and Graham et al 1992a.) The empirically measured segregation, however, is far from being identical across members of a series but tends to be highest for the one-element-only patterns.

Adding complex channels to the simple channels reconciles the models to the data for opposite-sign-of-contrast patterns (relative to one-element-only patterns), but the problem with the same-sign-of-contrast patterns remains. Empirically, the farther the luminances in the same-sign-of-contrast patterns get from the background luminance, the worse the segregation gets. It is this

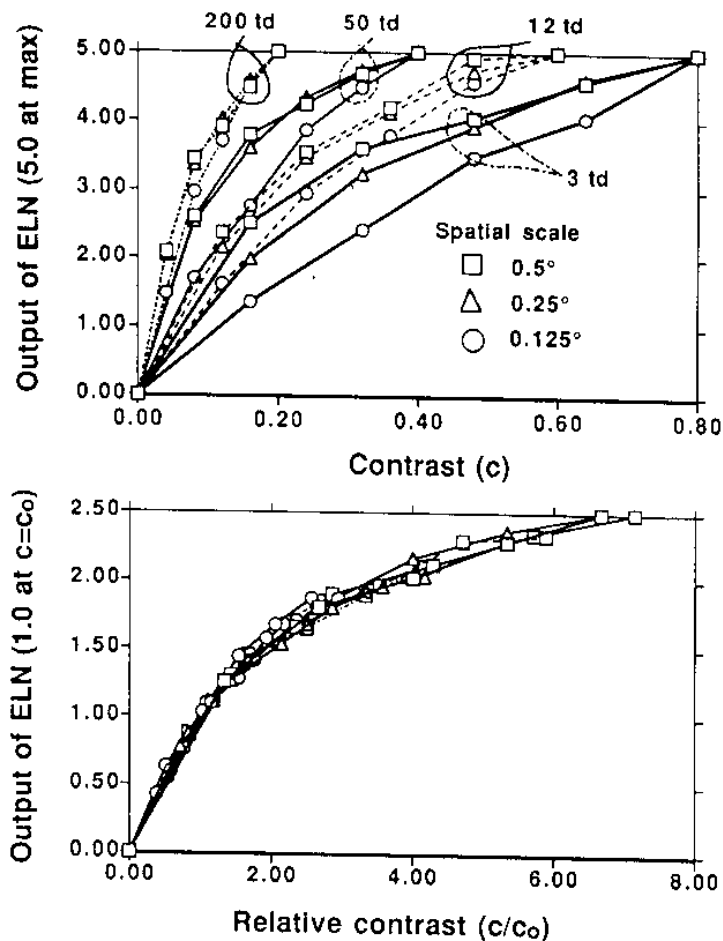


FIG. 4. The best-fitting early local non-linearities (ELNs) from an experiment using element-arrangement textures with square elements in 12 conditions (three spatial scales \times four background luminances) with one observer. The highest background luminance was approximately 200 trolands, td (18 foot-Lamberts seen through an artificial pupil of 2 mm diameter); the other three background luminances were 0.6, 1.2, and 1.8 log units below this level. Spatial scale was varied by proportionately scaling down both the size of the square elements and the inter-element spacing while keeping the same number of rows and columns in the pattern (as if one viewed the patterns of Fig. 1 from three different viewing distances). For the results shown here, both the width of the square and the inter-square spacing were the same and were 0.5° (16 pixels), 0.25° (8 pixels), or 0.125° (4 pixels), depending on condition. The three different scales were mixed within a session while mean luminance was changed between sessions. Other methodological details as in Graham et al (1993); the observer *mh* in that study is the observer here.

The top panel plots the output of the non-linearity (scaled to equal 5.0 at the highest contrast used in each condition) against Weber contrast ($\Delta L/L$). If the strong version of the early local non-linearity hypothesis were true, the three functions for the different scales at a given mean luminance would be identical. In fact, those for the smallest scale (circles) are always less compressive than for the largest scale (squares).

The bottom panel plots the output (scaled to equal 1.0 at the segregation-threshold contrast c_0) against relative contrast (contrast divided by segregation-threshold contrast

result that can be explained by a compressive non-linearity, which can be either relatively local and occurring before the channels (e.g. retinal light adaptation) or relatively global and occurring at the level of the channels in the form of a normalization network (e.g. inhibitory interaction among V1 neurons). The intuition behind these explanations is described in more detail in the original papers (Graham 1991, Graham et al 1992a). For the results of our original studies, models embodying either of these explanations fit the empirical results well.

For simple and complex channels

We have consistently found that the results are similar for same-sign-of-contrast stimuli affecting predominantly simple channels (e.g. square elements or blobs) or affecting predominantly complex channels (e.g. grating-patch elements), implying that the compressive non-linearity is active for both kinds of channels.

Increments versus decrements, approximate symmetry

When using square or blob elements, the results for same-sign-of-contrast patterns where both element types are decrements (darker than the background) are quite similar to the results when both element types are increments. Thus, for simplicity, we assume here that the compressive intensive non-linearity acts identically on increments and decrements (i.e. the degree of compression depends on the absolute value of the difference between the element luminance and the background, or the compressive function is odd-symmetric when plotted against deviation from the background luminance). In fact close examination of our recent data shows that there is a small asymmetry with stimuli composed of decrements being more easily segregated than those with increments of the same magnitude.

Effect of spatial scale and mean luminance

To try to disentangle the two suggested mechanisms for the compressive non-linearity, we did several experiments using constant-difference series with solid-square elements at several different spatial scales. With the observer whose results are shown in Fig. 4, we used three different scales and four different background luminances (for details see figure legend). We fit these results both using models assuming a normalization network and using models assuming an early, local non-linearity (with a modification of the procedure used by Graham et al 1992a). As with our earlier studies, when restricting our attention to a single spatial

c_0). The segregation-threshold contrast c_0 is the contrast of the one-element-only pattern at the given spatial scale and background luminance that yields a mid-range texture rating. Now the 12 functions juxtapose as expected if the weaker version of the early local non-linearity hypothesis is true.

scale and a single mean luminance, both kinds of models fit well and fit equally well (e.g. explaining 97 or 98% of the variance in each of the 12 conditions represented in Fig. 4).

How early could the early local non-linearity be?

The inferred intensive non-linearities at different scales, however, were *not* identical as a function of contrast. This is illustrated in the top panel of Fig. 4 for the model assuming an early local non-linearity. The output of the best-fitting early local non-linearity for each of the 12 conditions is plotted as a function of physical contrast. Since the output for each condition is only specified up to a multiplicative factor, it was set equal to 5.0 at the highest contrast used in each condition to allow easy comparison of curve shapes. (We used the same contrasts for the three spatial scales at a single background luminance but used different contrasts for each of the four background luminances. Thus the 12 curves in Fig. 4 top panel fall into four sets of three.) The three functions at a given background luminance have quite different shapes: the function for the smallest scale (circular symbols) is the most linear and that for the largest scale (square symbols) is the most compressive.

This difference between the early local non-linearity functions at different scales rejects one very strong version of the early local non-linearity hypothesis: namely, that the compressive non-linearity acts directly, point-by-point, on the luminance values. To put it another way, which makes it clear just how over-strong this version is, these results reject the hypothesis that the compression occurs before optical blurring. (Over-strong as this may be, some of us slip into a similar hypothesis when we forget that black-white stimuli presented to the observer become multiple grey-level stimuli on the retina with the distribution of grey levels depending on the spatial characteristics of the stimuli.)

However, the results from varying spatial scale cannot reject a weaker version of the early local hypothesis. In this version, the compression still occurs relatively locally but after some process which sets overall sensitivity to the spatial frequencies and orientations composing the pattern. To see the explanatory power of this hypothesis, consider the bottom panel of Fig. 4. Here the inferred best-fitting early local non-linearities from the top panel are replotted against physical contrast divided by a 'segregation-threshold contrast'. This is the contrast which, in a one-element-only pattern at the given background luminance and spatial scale, produced a mid-scale rating of perceived segregation. Also, the vertical scale in the bottom panel of Fig. 4 was adjusted to equal 1.0 at the segregation threshold for each of the 12 conditions to be consistent with the assumption that the sensitivity setting occurs before the early, local non-linearity. Note that the curves now juxtapose extremely well for different spatial scales and also for different mean luminances.

In short, although the compressive non-linearity cannot be so local and so early that it operates directly on the luminance, it could be local and operate before the channels but after a sensitivity-setting stage. If so, the degree to which it compresses its input is the same at all background luminances in the 1.8 log unit range that we studied. Further, as a semilog plot would show, over the range of contrasts studied in Fig. 4 (but not at lower contrasts, see below) the shape of the inferred early local non-linearity is very close to logarithmic.

How local could the early, local non-linearity be?

There is some evidence that the compressive non-linearity cannot be too local, or alternatively that the same compressive non-linearity cannot act on the same elements in different tasks. In one study, we compared region segregation judgements of element arrangement textures (like those we have been describing) with population-segregation judgements (how well the elements of one type stood out from among elements of the other type when the elements were randomly mixed rather than arranged in stripes and chequerboard). If the compressive non-linearity that acts in region-segregation judgements of element arrangement textures were local enough to act on single elements, one might expect it to show up in the population-segregation judgements as well. But these two kinds of judgements were very different for same-sign-of-contrast patterns (Beck et al 1991).

The normalization model

Are the results at different scales and different mean luminances consistent with predictions based on inter-channel interaction, as modelled by a normalization network? The answer seems to be yes (although it is less clear what that model predicts). Normalization networks normalize everything relative to the total responses in some part of the system. By adjusting the contrasts at the two scales to equalize the extent of segregation at the two scales, you have probably also adjusted total response in the relevant subsystem to be much the same. Then, all aspects of performance should be the same. They do seem to be (as is reflected in the bottom panel of Fig. 4, although those functions were derived by fitting the other model).

Comparison with other compressive non-linearities

The comparison of several non-linearities in Fig. 5 provides some perspective on the nature of this non-linearity. The top panel of Fig. 5 shows two non-linearities from models for psychophysical results usually thought to reveal 'light-adaptation' processes. The bottom panel of Fig. 5 shows four functions that represent physiological results. The function inferred from Fig. 4 (at a contrast appropriate to large scale and high mean luminance) is repeated in both panels.

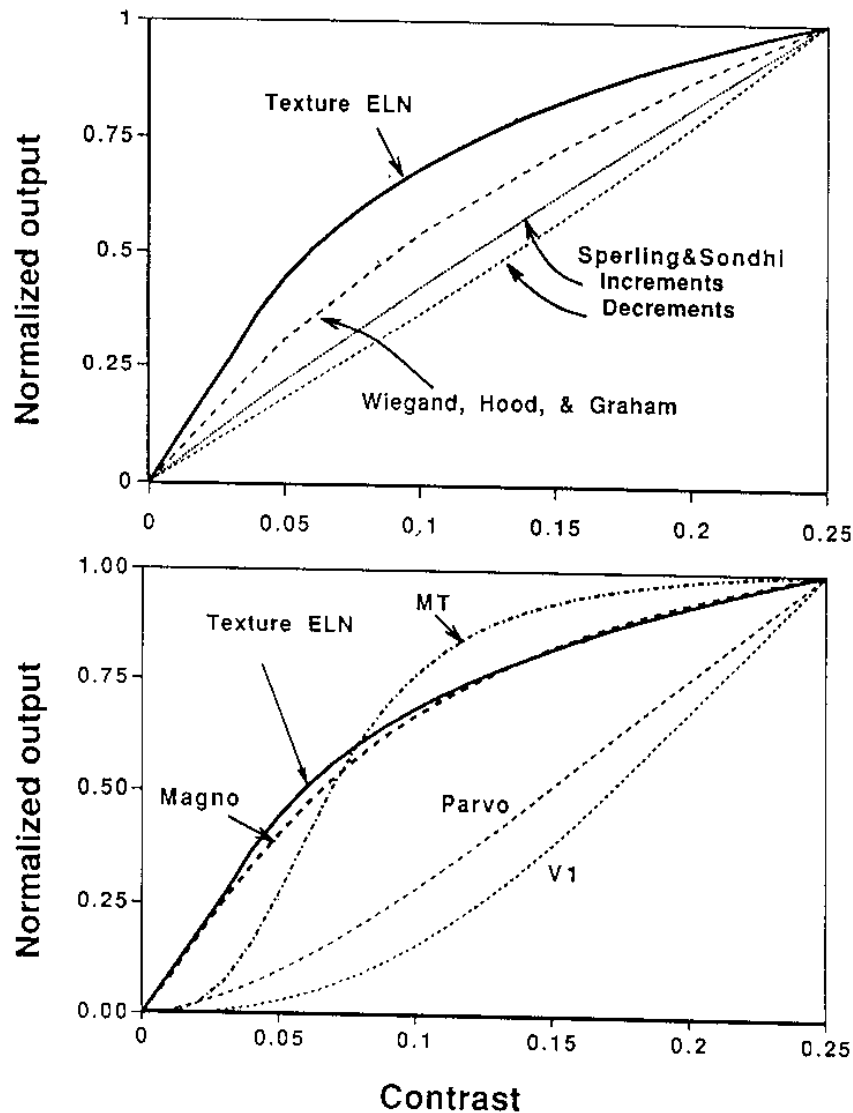


FIG. 5. The best-fitting early local non-linearity (ELN) from the study of Fig. 4 plotted against Weber contrast (at the appropriate level for the largest spatial scale at the highest mean luminance) and compared to the compressive non-linearities from other sources. All the functions are vertically scaled to be equal to 1.0 at a contrast of 25%.

The top panel shows two non-linearities from models predicting psychophysical results usually thought to reveal the dynamics of 'light-adaptation' processes: the model of Sperling & Sondhi (1968) and that of Wiegand et al (1992) (also see Graham & Hood 1992, Wiegand 1993). The peaks in the responses to increments (both models) and decrements (the Sperling & Sondhi Model) from a background luminance of 600 trolands are plotted. The bottom panel shows four functions that represent physiological results for single neurons from four visual areas. The functions were computed from the median values of the parameters given by Sclar et al (1990).

Clearly, the functions from light-adaptation psychophysics and from parvocellular LGN and from V1 cells are not as compressive as the function inferred from texture-segregation results, but those from magnocellular LGN and area MT are. Handle these comparisons with care.

Acceleration at very low contrasts

We have now done several experiments using constant-difference series with very small differences (N. Graham & A. Sutter, unpublished results). Thus, the patterns sometimes contained elements that were of very low contrast and might be barely if at all detectable. In fitting the models to these data, two clear results emerged: the best-fitting early local non-linearity was strongly accelerating at the very low contrasts and, probably as a direct result of this, the best-fitting model assuming early-local non-linearity fitted substantially better than that assuming normalization.

Thus, even if normalization turns out to be the better explanation of the compressive effects, it would need to be augmented by something more like an early, local acceleration at low contrasts. The relationship of this acceleration to that discussed above in connection with the non-linearity at the middle of the complex channels is not yet clear to me.

A parting comment

Putative 'higher-order processes' (here meaning processes that paste together information on a relatively local scale into information on a more global scale) have been suggested as components in a number of perceptual tasks—perceived segregation of texture regions, deciding between competing motion paths, finding a path through static multi-element patterns, and others. All these tasks, while sharing some component processes, may also be strongly dependent on unshared processes. To put it another way, the 'higher-order processes' by themselves cannot explain the whole journey from stimulus to the observer's reported perception. As the characteristics of the higher-order processes involved in each task become known in greater detail, it should become easier to compare the higher-order processes involved in different tasks and then to use the differences among tasks to suggest further hypotheses about visual mechanisms.

Acknowledgement

This work was partially supported by National Eye Institute grant 1 RO1 EY08459.

References

- Beck J, Graham N, Sutter A 1991 Lightness differences and the perceived segregation of regions and populations. *Percept Psychophys* 49:257–269

- Cannon MW, Fullenkamp SC 1993 Spatial interactions in apparent contrast: individual differences in enhancement and suppression effects. *Vision Res* 33: 1685-1695
- Graham N 1989 *Visual pattern analyzers*. Oxford University Press, New York
- Graham N 1991 Complex channels, early local nonlinearities, and normalization in texture segregation. In: Landy ML, Movshon JA (eds) *Computational models of visual processing*. MIT Press, Cambridge, MA
- Graham N 1992 Breaking the visual stimulus into parts. *Curr Dir Psychol Sci* 1:55-61
- Graham N, Hood D 1992 Modeling the dynamics of light adaptation: the merging of two traditions. *Vision Res* 32:1373-1393
- Graham N, Beck J, Sutter A 1992a Nonlinear processes in spatial-frequency channel models of perceived texture segregation: effects of sign and amount of contrast. *Vision Res* 32:719-743
- Graham N, Sutter A, Venkatesan C, Humaran M 1992b Nonlinear processes in perceived region segregation: orientation selectivity of complex channels. *Ophthalmic & Physiol Opt* 12:142-146
- Graham N, Sutter A, Venkatesan C 1992c Characteristics of complex channels (second-order processes) in texture segregation. *Optical Society of America Technical Digest Series* 23:36(abstr)
- Graham N, Sutter A, Venkatesan C 1993 Spatial-frequency-selectivity and orientation-selectivity of simple and complex channels in region segregation. *Vision Res* 33:1893-1911
- Nothdurft H-C 1994 On common properties of visual segmentation. In: *Higher-order processing in the visual system*. Wiley, Chichester (Ciba Found Symp 184) p 245-268
- Sclar G, Maunsell JHR, Lennie P 1990 Coding of image contrast in central visual pathways of the macaque monkey. *Vision Res* 30:1-10
- Sperling G, Sondhi MM 1968 Model for visual luminance discrimination and flicker detection. *J Opt Soc Am* 58:1133-1145
- Sperling G, Chubb C, Solomon JA, Lu Z-L 1994 Full-wave and half-wave processes in second-order motion and texture. In: *Higher-order processing in the visual system*. Wiley, Chichester (Ciba Found Symp 184) p 287-308
- Sutter A, Graham N 1992 Investigating the dynamics of texture segregation processing using the speed-accuracy-tradeoff method. *Optical Society of America Technical Digest Series* 23:159(abstr)
- Sutter A, Beck J, Graham N 1989 Contrast and spatial variables in texture segregation: testing a simple spatial-frequency channels model. *Percept Psychophys* 46:312-332
- Sutter A, Chubb C, Sperling G 1991 Further measurements of the spatial frequency selectivity of second-order texture mechanisms. *Invest Ophthalmol & Visual Sci (suppl)* 32:1039(abstr)
- Werkhoven P, Sperling G, Chubb C 1992 The dimensionality of texture-defined motion. *Invest Ophthalmol & Visual Sci (suppl)* 33:1049(abstr)
- Werkhoven P, Sperling G, Chubb C 1993 The dimensionality of texture-defined motion: a single channel theory. *Vision Res* 33:463-485
- Wiegand T 1993 A psychophysically-based computational model of the dynamics of light adaptation. PhD thesis, Columbia University, New York, USA
- Wiegand T, Hood D, Graham N 1992 Light adaptation dynamics: some psychophysical data and some improvements to a model. *Invest Ophthalmol & Visual Sci (suppl)* 33:2825(abstr)
- Wilson HR, Richards WA 1992 Curvature and separation discrimination at texture boundaries. *J Opt Soc Am A Opt Image Sci* 9:1653-1662

DISCUSSION

Wilson: It's clear that you need some form of compression somewhere. You considered either a front-end pointwise non-linearity or the possibility of using pooling and normalization after all the filtering rectification. You chose normalization after rectification (which I'm sure will work), but would pointwise (one-per-filter) simple compression after the rectification also work?

Graham: I don't think that a pointwise non-linearity at the back end (after the filtering) will work. The filter output at the fundamental frequency is the same for all the same-sign-of-contrast patterns in a series, so you can't compress it and hope to produce any differences—you would be compressing them all from the same point.

Lennie: I've no idea whether the grouping effects you've described are actually confined to the intensity domain. You can imagine chromatic analogues of all the experiments you've described—it's not clear to me that the results would be different.

Graham: It's not clear to me, either. Jacob Beck has done some work (unpublished results) with coloured versions of the sorts of patterns I showed, but it's very hard to say one way or the other at this point.

Movshon: Relating a transducer function deduced from a psychophysical experiment to a transducer function measured in a neuron is probably perilous for many of these junctions, specifically for the reason that the psychophysics transducer function has absorbed the variability. If the noise is constant at all neural response magnitudes, then you might make the comparison a little less perilously. But we know for all the classes of cortical neurons (the bottom part of Fig. 5) that variability changes very sharply with response. It does not for the geniculate cells. If you want a comparison, you should plot a Z transform signal from the cortical cells. This will give a very different answer.

Lennie: Are you saying that the variance of the response of an LGN cell doesn't change with amplitude?

Shapley: Keith Purpura, Rudi Kaplan and I have said that this is additive (Purpura et al 1989), but that's in a stabilized image situation. If you do it, as these experiments were done, not with stabilized images, then the effective noise might actually grow with contrast.

Lennie: That's what I would have expected.

Movshon: Why?

Shapley: Because you might get variance introduced by eye movements as a consequence of the dithering of the eyes around the borders.

Movshon: On the other hand, your eye movements might be more precise in the higher-contrast cases.

Shapley: You need to take it into account.

Georgeson: If you plot your non-linear response function on log-log axes, what sort of slope or exponent do you get?

Graham: Log-log isn't as useful as semilog. Except for very low contrasts (lower than in the experiments I showed here), the function is extremely close to logarithmic. Plotting it on semilog paper gives an almost straight line.

I would like to say something about the comparison between two sets of results. One is a set of results for motion perception that were published by Werkhoven et al (1993) and which Charlie Chubb mentioned in his talk (Chubb et al 1994, this volume). These results show *no* selectivity for spatial frequency (and, perhaps, for orientation—Werkhoven et al 1992) at the first stage of the non-Fourier channels for motion. Or, as they phrase it, you need only a single motion channel because the same channel is sensitive to all (visible) spatial frequencies. On the other hand, our results in a texture region-segregation task (Graham et al 1993), showed a great deal of selectivity for both spatial frequency and orientation. Thus one needs multiple complex channels (with different orientation and spatial frequency selectivities) for the texture case. In other words, the bottom panel of Fig. 3 is correct for the texture case.

I have been worrying about what the difference could be. Is it really a case of different channels for texture than for motion? Perhaps. But some other things need to be considered first. For one thing, there is the problem of differences among observers.

The Werkhoven et al (1993) data, which Charlie Chubb didn't show although he mentioned their results, are convincing and clear. But the data are from only two observers (and rather experienced observers at that). Our published results, from a few more observers, showed rather larger differences (Graham et al 1993). Perhaps observers just differ in bandwidth (but have the same bandwidth in both the motion and texture situations) and Werkhoven et al (1993) happened to get two of the broadest-bandwidth observers.

The recent Cannon & Fullenkamp (1993) paper is interesting in this regard. As they explain in their introduction, they had used three subjects in an earlier study and found only suppression, not enhancement (in a task measuring perceived contrast under conditions of spatial interactions). But they were worried because Ejima & Takahashi (1985) had reported some enhancement under similar conditions. So they tried to track down any changes in condition. For a new study they ran 10 subjects. Of these 10, eight showed enhancement, with some of the eight showing a great deal! The other two observers showed only suppression—these were the authors MC and SF. Thus, there is a reason to worry about possible individual differences in the texture versus motion comparison. On the other hand, we have now run about 10 subjects in our texture task and never found one with bandwidths nearly as broad as those for the motion task!

This summer I got to wondering whether the difference between the motion and texture results was something intrinsic to the tasks involved: a region-segregation task in our texture experiments versus a path-competition task in the motion experiments of Werkhoven et al (1993). So I made a static analogue

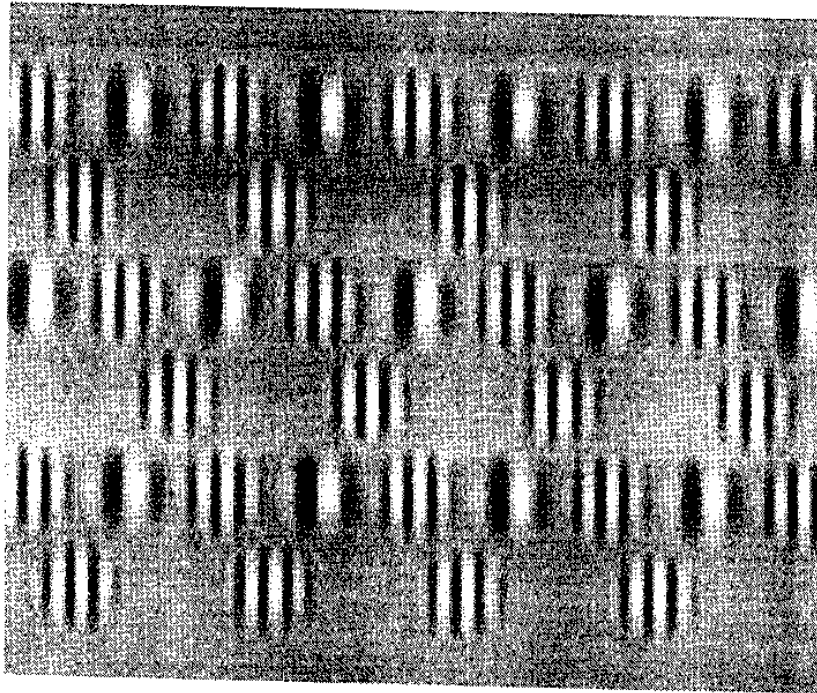


FIG. 1. (*Graham*) This is formally analogous to the Werkhoven, Sperling & Chubb's scheme shown in Fig. 1 of the paper by Chubb et al (1994, this volume).

(shown in Fig. 1 [*Graham*]) of the kind of path-competition stimuli used by them. If you compare it to Fig. 1 in Charlie Chubb's paper (Chubb et al 1994, this volume), you will see the analogy. 'A' and 'B' in his figure are the high- and low-spatial-frequency patches here. (Remember that in the motion case, each row represents a frame; also, each row was wrapped around an annulus so that motion was circular either anticlockwise or clockwise.) In our static analogue, we used patterns just like those shown in Fig. 1 (*Graham*). (In the one shown in this figure, the homogeneous path is down towards the right and contains only high-spatial-frequency patches; the heterogeneous path is down towards the left, containing alternating low- and high-spatial-frequency patches.) We simply asked the observers to judge to what extent the homogeneous path or the heterogeneous path dominates in the perception.

The really astonishing result Werkhoven et al (1993) found for their motion experiments using patches of sinusoidal gratings is this: if you adjust the contrast in the two spatial frequencies (no matter how far apart they are) so that the *total energy* on the heterogeneous path is greater than that on the homogeneous, the heterogeneous path is what people see all the time. In short, their motion perception is integrating across enormous spatial frequency differences!

Chubb: Yes, in particular, a texture-defined motion display which pits a homogeneous motion path against a heterogeneous path composed of alternating high-frequency texture patches and low-frequency patches (equal in contrast to the high-frequency patches) will consistently elicit motion in the direction of the heterogeneous path.

Graham: But in the static analogue, as you can probably see for yourself (Fig. 1 [*Graham*]), the result isn't even close to that for motion! A person may see both the homogeneous path and the heterogeneous path, but nobody I asked said that the heterogeneous path dominated (except for at very close spatial frequencies).

We repeated exactly the Werkhoven et al (1993) experiment using these static analogue patterns with one of the subjects run earlier on the texture region-segregation task (CV in Graham et al 1993). As expected from the informal observations, her static path-competition results didn't look like those for motion. She showed selectivity for spatial frequency (that is, the heterogeneous path could only dominate when the spatial frequencies were very close). In fact, her inferred bandwidth on this texture path-competition task was just like her bandwidth on the texture region-segregation task. So, I don't think that the difference between the results of Werkhoven et al (1993) for motion and our results for texture is only a consequence of some difference between region-segregation and path-competition tasks (although I still think some differences between those tasks could exist).

Shapley: One thing that strikes me as different is that this is a task where the temporal modulation is low frequency—that is, it is one second or two seconds temporal exposure—and the whole essence of the motion task is that it's temporal modulation of some moderate temporal frequency. If you reversed the contrast of your textures at moderate temporal response, perhaps the difference would disappear.

Graham: I don't know—do you think that would work? Is this a temporal frequency problem as opposed to a motion texture problem?

Sperling: Whether temporal modulation improves or impairs spatial region segregation is an interesting question. However, when you create a dynamic stimulus, you not only affect the discriminability of differently textured spatial regions, you also produce motion paths according to the spatiotemporal modulation pattern. In terms of channels, I think the difference between performance in motion tasks and in texture-segregation tasks can be summed up simply in this way: there's lots of evidence of separate frequency channels in texture vision; however, Werkhoven et al (1993) find that second-order motion perception seems to be served primarily by a single, low-pass frequency channel.

Graham: But this gets back to something that I didn't understand about the discussion following Hugh Wilson's paper (Wilson 1994, this volume), which was about the orientation bandwidth.

Wilson: At the time we developed the model, there were no data one way or the other concerning non-Fourier orientation bandwidth and so we chose oriented filters. The first evidence on non-Fourier bandwidths is very recent (Werkhoven et al 1993).

Burr: We measured orientation bandwidths in two ways: with a masking paradigm (Anderson et al 1991) and by summation (Anderson & Burr 1991), with similar results from the two techniques. The bandwidths vary considerably with spatial frequency, from around 45° at 10 cycles/degree to 80° at 0.1 cycles/degree (full width at half height).

Graham: Are we able to conclude from a combination of Anderson & Burr (1991) and Werkhoven et al (1993) that Hugh Wilson's wiring diagram was wrong? In Hugh's wiring diagram, the same filters that go up the Fourier path form the first stage in the non-Fourier path. If so, however, are those filters oriented when Anderson & Burr (1991) study them, and not when Werkhoven et al (1993) study them?

Shapley: They're not that oriented, that's the issue.

Burr: They're fairly broad.

Shapley: What is the actual orientation tuning of the front end for this channel? It might be that they would agree.

Sperling: In the first paper (Werkhoven et al 1993), we didn't study orientation tuning at all. In the second paper (Werkhoven et al 1994), we did find a small amount of orientation tuning but not enough to measure a bandwidth for orientation.

Burr: Do you get different results with the rotating stimulus and the displacement stimulus?

Chubb: We haven't done experiments with the horizontal displays.

Shapley: If you are actually measuring the orientation tuning for motion, it was quite broad in those experiments (Anderson & Burr 1991).

Burr: Yes, quite broad: around 60° (full width at half height) for the spatial frequencies used here. The other point to remember is that half height is quite an arbitrary descriptor; the filters will respond to high-contrast stimuli of most orientations.

Wilson: For full bandwidth at half height, I used 45° in the model I discussed (Wilson 1994, this volume): I referred to $\pm 22.5^\circ$.

Morgan: There are other indications that individual differences in high-order filtering might turn out to be quite profound and interesting. An example concerns the ability of observers to locate the positions of the centroids of dot clusters. The dots are randomly thrown into a notional circle, and the task is to locate the centroid. We know this is a second-level task because they do just as well with balanced clusters which would be invisible to first-order filters. Thresholds decrease (or performance improves) as a function of the square root of the number of elements, as you would expect, and relative efficiency is pretty high.

I had one subject who totally baffled me, she had extremely low efficiency on this task. She did not benefit from having extra dots, so the efficiency fell drastically as dots were added. For most subjects it is about constant up to 20 dots and then starts to fall off. I could detect no other problems with her vision. Her thresholds for Vernier alignment of two dots (or spatial interval between two dots) were in the normal range—what she couldn't do was to integrate positional information across dots (Morgan & Glennerster 1991).

Graham: George Sperling told us that in the motion case, five of his 15 observers see the motion carried by half-wave stimuli, the other 10 don't (Sperling et al 1994, this volume).

Sagi: Some of the inter-subject variability may arise from them having different thresholds. These second-stage filters may start to operate only above some contrast level and this critical contrast may vary across subjects. In the type of experiments Cannon & Fullenkamp (1991) were running, one may get enhancement or suppression depending on how far the stimulus contrast is from the threshold. This is certainly true for contrast-discrimination experiments where at low base (pedestal) contrasts you get just-noticeable difference values that are lower than absolute threshold (enhancement) and at higher base contrasts you get just-noticeable difference values that are larger than the absolute threshold (suppression).

Graham: I think Cannon & Fullenkamp's (1993) fourth experiment was actually measuring contrast thresholds and saying that individual differences were not just a matter of sensitivity differences.

There is something else I have wondered about in the Cannon & Fullenkamp studies. The first two observers (who seem to have been themselves) were probably the most practised observers—most practised by years, I would think, not just by 10–20 sessions. Is that perhaps why they were different from the next eight? Another study where there was a big individual difference, Gurnsey & Browse (1992), seemed to be a difference between practised and unpractised observers. When a difference showed up between a naive observer and a practised observer, they ran one relatively naive observer for 10 or more sessions on a set of the patterns. That observer showed improvement on at least one of the patterns but still couldn't do one or more patterns as well as the very highly practised observer (the author RG). The take-home message here seems to be that there is learning and the learning can go on for a very long time. Karni & Sagi (1991) and Fiorentini & Berardi (1981) are two other studies that make the point about learning even more clearly and then begin to study it in some detail. So maybe differences in amount of effective practice on a task contribute to differences among observers on tasks like those we are discussing—tasks using suprathreshold stimuli incidentally.

References

- Anderson SJ, Burr DC 1991 Spatial summation properties of directionally selective mechanisms in human vision. *J Opt Soc Am A Opt Image Sci* 8:1330-1339
- Anderson SJ, Burr DC, Morrone MC 1991 The two-dimensional spatial and spatial frequency properties of motion-sensitive mechanisms in human vision. *J Opt Soc Am A Opt Image Sci* 8:1340-1351
- Cannon MW, Fullenkamp SC 1991 Spatial interactions in apparent contrast: inhibitory effects among grating patterns of different spatial frequencies, spatial patterns and orientation. *Vision Res* 31:1985-1998
- Cannon MW, Fullenkamp SC 1993 Spatial interactions in apparent contrast: individual differences in enhancement and suppression effects. *Vision Res* 33:1685-1695
- Chubb C, McGowan J, Sperling G, Werkhoven P 1994 Non-Fourier motion analysis. In: Higher-order processing in the visual system. Wiley, Chichester (Ciba Found Symp 184) p 193-210
- Ejima Y, Takahashi S 1985 Apparent contrast of a sinusoidal grating in the simultaneous presence of peripheral gratings. *Vision Res* 25:1223-1232
- Fiorentini A, Berardi N 1981 Learning in grating waveform discrimination: specificity for orientation and spatial frequency. *Vision Res* 21:1149-1158
- Graham N, Sutter A, Venkatesan C 1993 Spatial-frequency-selectivity and orientation-selectivity of simple and complex channels in region segregation. *Vision Res* 33:1893-1911
- Gurnsey R, Browse RA 1992 Micropattern properties and presentation conditions influencing visual texture discrimination. *Percept & Psychophys* 41:239-252
- Karni A, Sagi D 1991 Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proc Natl Acad Sci USA* 88:4966-4970
- Morgan MJ, Glennerster A 1991 Efficiency of locating centres of dot-clusters by human observers. *Vision Res* 31:2075-2083
- Purpura K, Kaplan E, Shapley R 1989 Fluctuations in spontaneous discharge and visual responses in M and P pathways of the macaque monkey. *Invest Ophthalmol & Visual Sci (suppl)* 30:297(abstr)
- Sperling G, Chubb C, Solomon J, Lu Z-L 1994 Full-wave and half-wave processes in second-order motion and texture. In: Higher-order processing in the visual system. Wiley, Chichester (Ciba Found Symp 184) p 287-308
- Werkhoven P, Sperling G, Chubb C 1992 The dimensionality of texture-defined motion. *Invest Ophthalmol & Visual Sci (suppl)* 33:1049(abstr)
- Werkhoven P, Sperling G, Chubb C 1993 The dimensionality of texture-defined motion: a single channel theory. *Vision Res* 33:463-485
- Werkhoven P, Sperling G, Chubb C 1994 Perception of apparent motion between dissimilar gratings: spatio-temporal properties. *Vision Res*, in press
- Wilson HR 1994 Role of second-order motion signals in coherence and transparency. In: Higher-order processing in the visual system. Wiley, Chichester (Ciba Found Symp 184) p 227-244