From Harris, C. (ed), 1980
Visual Coding and Adaptability
L. Erlbaum Assocs.

Spatial-Frequency Channels in Human Vision: Detecting Edges Without Edge Detectors

Norma Graham

Columbia University

One early approach to the study of vision was to investigate the appearance of small patches of light and to try to explain the appearance of the whole visual field as the juxtaposition of the appearances of many small patches. A theoretical model that can be viewed as a natural descendant of this early approach, made appropriately more rigorous for the case of threshold experiments, is what I will call a single-channel model.

Despite considerable success in accounting for a variety of visual data, it now appears that a single-channel model is an inadequate description of visual perception. Moreover, a currently prevalent view is that trying to describe the appearance of a whole visual field as the juxtaposition of appearances of single points is doomed to failure from the start. Rather, according to this current view, the appearance of things depends on many stages of complicated information processing. The initial stages occur in the retina and further stages extend throughout the highest parts of the central nervous system.

A popular candidate for one of the earliest stages in this chain of visual information processing is a collection of *feature detectors* that simultaneously process different kinds of information in the visual stimulus. Each feature detector is presumed to respond vigorously only when the stimulus situation contains the appropriate "feature"—for example, an "edge detector" would respond only when there is an edge in the appropriate place on the retina.

What I intend to do here is describe the role one kind of psychophysical experiment has played in the rejection of single-channel models of the visual system and in the exploration of feature-detection models. In this kind of experiment, the visibility of compound patterns composed of two or more simpler patterns is compared to the visibility of the simpler patterns alone.

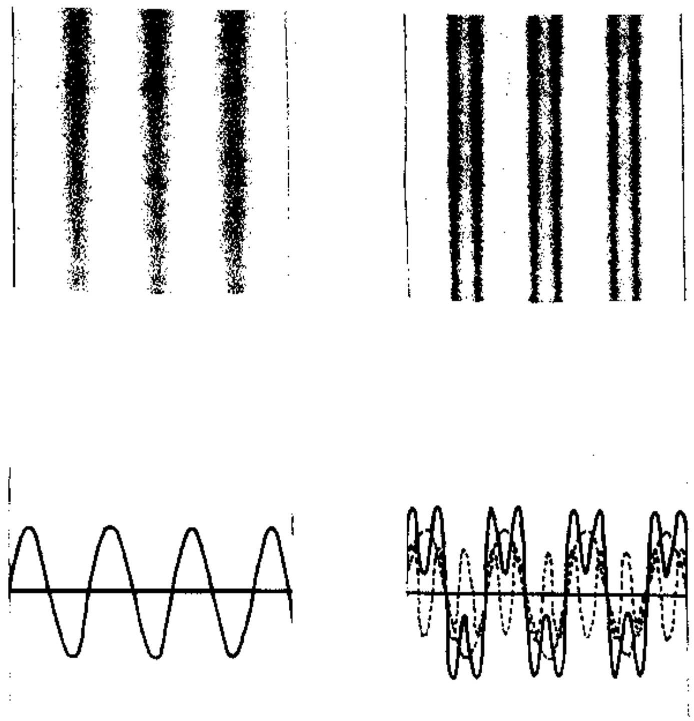


FIG. 1. A simple sine-wave grating containing one spatial frequency is shown on the top left, and a compound grating containing two spatial frequencies is on the top right. Underneath are the intensity profiles of the gratings, showing intensity of the grating at each horizontal location across the pattern. From Graham and Nachmias (1971).

In the first section I will review early experiments on the detection of compound patterns made up only of sinusoidal components (examples of a simple and a compound sine-wave pattern of this type are shown in Fig. 1). These early experiments produced strong evidence against the single-channel model for threshold vision. The findings can instead be interpreted as evidence for the existence of a rather odd type of feature detector—a detector or channel which responds only to patterns containing spatial frequencies within a limited range. Very roughly, being sensitive to a limited range of spatial frequencies means responding best to a particular size of element in the pattern (e.g. width of stripe); a more precise definition of spatial frequency is given below. I'll refer to this kind of channel as a *spatial frequency channel.*

In the second section I will discuss some recent experiments by Shapley and Tolhurst and by Kulikowski and King-Smith. These elegant experiments used patterns made up of sinusoids plus aperiodic stimuli (for example, sinusoids plus lines, sinusoids plus edges) as well as various combinations of aperiodic stimuli. These authors interpreted their results as evidence for the existence of several additional kinds of feature detectors—things like edge detectors and line detectors. The crucial distinction between these new feature detectors and the spatial frequency channels, as will be described later, is that each of these new feature detectors is supposed to respond to a broader range of spatial frequencies than does any spatial frequency channel.

I will argue, however, that these experiments do not actually provide persuasive evidence for the existence of additional feature detectors. On the contrary, my conclusion is that these new findings can probably be explained in terms of the same spatial frequency channels that were inferred from the earlier sinusoid-plus-sinusoid experiments. To reach that conclusion, I will reexamine the newer data in the light of a model that allows for probability summation among spatial frequency channels.

Much of the work referred to here is not mine and I will mention the authors in the appropriate places. Much of the work that is mine has been done in collaboration with Jacob Nachmias of the University of Pennsylvania.

A SINGLE-CHANNEL MODEL

Sine-Wave Gratings

Let's begin by looking at examples of gratings containing one sinusoidal component (the sine-wave pattern in Fig. 1 left) or two sinusoidal components (Fig. 1 right). Below each pattern is a graph that shows how the intensity of the grating varies as you move horizontally across it. For the left pattern, the graph depicts a single sinusoid added to a constant intensity (the mean luminance). For the right pattern, the graphed function is the sum of two sinusoids added to a constant intensity.

For patterns such as these, it is easy to specify and understand what *spatial frequency* is: The spatial frequencies contained in a pattern are the frequencies (cycles per unit distance) of the sinusoids that add up to equal the function relating intensity to distance across the pattern. Thus the left pattern contains only one spatial frequency and the right pattern contains two frequencies, having a ratio of three to one. For the pattern on the left, the spatial frequency is the number of peaks (bright bars) per unit of horizontal distance. We define the *contrast* of a stimulus (a measure of how different the light and dark bars are) as half the distance between the peak and trough intensities divided by the mean intensity.

First, let's describe a typical single-channel model of the visual system and then we can see what such a model predicts for the responses to simple sine-wave gratings, gratings containing only one sinusoidal component. At the same time we can review a few of the basic facts about sinusoidal stimuli.

A Single Channel

Those of you who like physiological analogues can think of a single channel as an array of retinal ganglion cells or lateral geniculate cells or even simple cortical cells. Each cell in the array has the same kind of receptive field (the same shape, the same orientation, the same size, everything the same except the position on the retina). But the receptive fields of different cells in the array, although they overlap, cover different portions of the visual field.

More abstractly, we can consider a single channel to be a two-dimensional array of "weighting functions" (defined below) corresponding point-by-point to the visual stimulus. (My use of the term "channel" is different from some other people's uses. Readers interested in a discussion of this terminology should see page 258.) For the purposes of models like this, the visual stimulus is considered to be two-dimensional as it is on the retina rather than three-dimensional as it is in the world. In fact, we will be dealing only with stimuli that are effectively *one* dimensional: The striped gratings vary in intensity only along the horizontal axis; they maintain the same intensity along any vertical line. Therefore we need consider only a one-dimensional cut across the two dimensions of the single channel. In general, then, the response of a channel is a two-dimensional array corresponding point-by-point to the visual stimulus. But we'll usually be considering a one-dimensional cut across the response: the *response profile.*

The Weighting Function

The magnitude of the response at any point in the single channel's response profile can be specified by a *weighting function.* The weighting function indicates the extent to which light falling at various points on the retina adds to or subtracts from the response at the given point in the single channel. (The weighting function is so named because it describes how the light falling on different points is weighted in determining the response.) In terms of the physiological analogue, the weighting function would be a quantitative description of a cell's receptive field, and the response at a point in the channel would be the output from the cell connected to that receptive field.

One kind of hypothetical weighting function is represented by the small sketches in the top line of Fig. 2. The line as a whole represents a one-

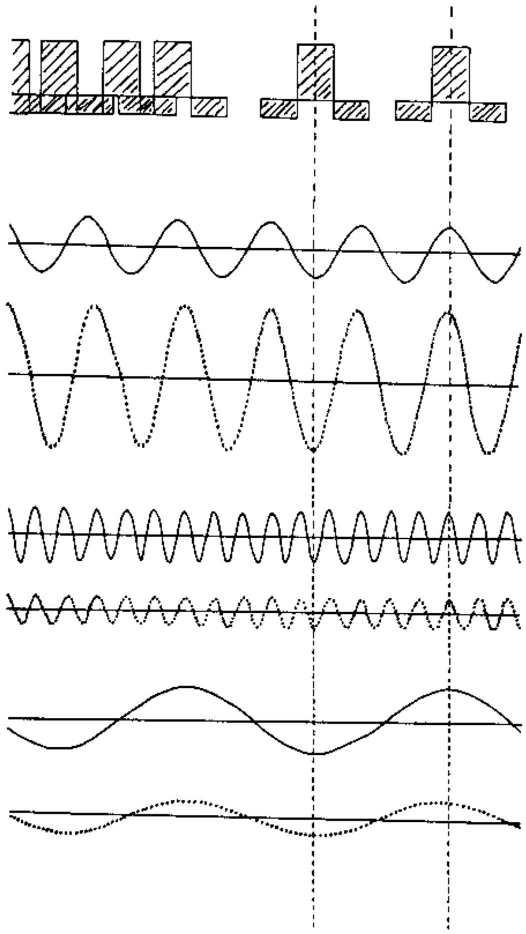


FIG. 2. Diagram of some of the weighting functions that determine the responses of a single channel (top row). The channel's responses (third, fifth, and seventh rows) to three gratings of different spatial frequencies (second, fourth, and sixth rows).

dimensional cut across the single channel. Each of the small sketches of the weighting function indicates that the response at a given point of the channel is increased by light falling anywhere within some small area of the retina and is decreased by light falling anywhere within a surrounding area. (The weighting function shown in the figure has a somewhat artificial "rectangular" distribution for the center excitatory area and for the surrounding inhibitory area. It is easy to substitute other, more plausible configurations.) The weighting functions should really be pictured as being densely distributed all along the top line, with many of them overlapping at any given point, but they have been thinned out for clarity here. Notice that the weighting functions at all points across the channel are assumed to be the *same*; this is an important assumption of the single-channel model.

Response to Different Spatial Frequencies

Your intuition may suggest that this single-channel model should produce big responses for gratings in which the bar-widths match the dimensions of the weighting function, and smaller responses for gratings of other bar-widths. The rest of Fig. 2 shows that such an intuition is approximately correct. Here we see the intensity profiles of three stimuli and the response at each point across the single channel to each of the three. Notice first that, conveniently, the response to any sine-wave stimulus is itself sinusoidal, as long as you are considering linear systems. (We are assuming that the single channel is linear: All it does is add and subtract.)

The second row of the figure shows the intensity profile of a sinusoidal grating of intermediate spatial frequency. Consider the response (third row in Fig. 2) at the middle of the bright bar at the extreme right end of the figure. There is a lot of excitation because a bright bar is illuminating the center of the weighting function (or receptive field on the retina). There is little inhibition because most of the surround of the weighting function is illuminated by dark bars. Little inhibition and a lot of excitation produces a big net response. When the peak response is large (compared to the mean response) we say the channel is responding well to this stimulus pattern.

Now consider the response at the middle of the dark bar. There is a lot of inhibition because most of the negative surround of the weighting function is illuminated by bright bars. The more strongly illuminated the inhibitory surround is, the less the net response. Furthermore, there is little excitation, because the excitatory center is getting little illumination from the dark bar. Thus there is very little response at this point in the channel. So the total difference between the peak response (the response in the middle of the bright bar) and the trough response (the response in the middle of the dark bar) is very large. A large difference between peak and trough also indicates that the channel is responding well to the grating.

However, when you consider a higher spatial frequency grating (such as that for which the intensity profile is shown in the fourth row of Fig. 2), the bars are so closely spaced that *many* bright and dark bars fall within the center of each weighting function. Likewise, many bright and dark bars fall within the surround. Thus the response at any point in the array is approximately the same as that at any other point, because there are always approximately the same number of bright bars as dark bars in both the center and surround. The peak response is small, and so is the difference between peak and trough. That is, the channel does not respond well to this high frequency grating.

Finally, consider what the response (seventh row) to a low spatial frequency grating (sixth row) is like. At the center of the bright bar, there is a lot of excitation (as for the medium spatial frequency) because the center of the weighting function is illuminated by a bright bar. But there is also a lot of inhibition (unlike the case for the medium spatial frequency) because the negative surround of the weighting function is also illuminated by the bright bar. A lot of excitation and a lot of inhibition leads to a smaller response than a lot of excitation and a little inhibition, so the peak response to the low-frequency grating is smaller than to the medium spatial frequency. We say that the channel is not responding well.

Like the peak response, the difference between the peak and *trough* response is also smaller for a low-frequency grating than for a medium frequency. In the middle of the dark bar there is little excitation (like the medium frequency case) but there is also little inhibition (unlike the medium frequency case). Little excitation coupled with little inhibition leads to a trough that is not as deep as for the medium spatial frequency. Therefore the difference between the peak response and trough response is relatively small for the low spatial frequency.

In short, a single-channel model with a center-surround type of weighting function predicts bigger responses to gratings of intermediate frequency than to gratings of lower or higher frequencies.

Psychophysical Data

In predicting the greatest response to intermediate frequencies, a single-channel model does agree well with psychophysical data. A human observer's responsiveness to a grating is often measured by finding the *contrast threshold*, the smallest light-dark contrast that enables the observer to tell there is a grating present rather than a blank field. (In these experiments, the average intensity is held constant while the contrast is varied.) Another measure of the same sort is *contrast sensitivity*, which is defined as the reciprocal of the contrast threshold; the higher the contrast threshold, the lower the contrast sensitivity.

Consider what should happen if a human observer were well described by a single-channel model and if we made the additional assumption that the contrast threshold is the smallest contrast necessary to produce a sufficiently large response somewhere across the single channel. In other words, we assume the contrast threshold is achieved when the contrast is high enough for the channel's peak response to exceed some criterion. Then we would expect the

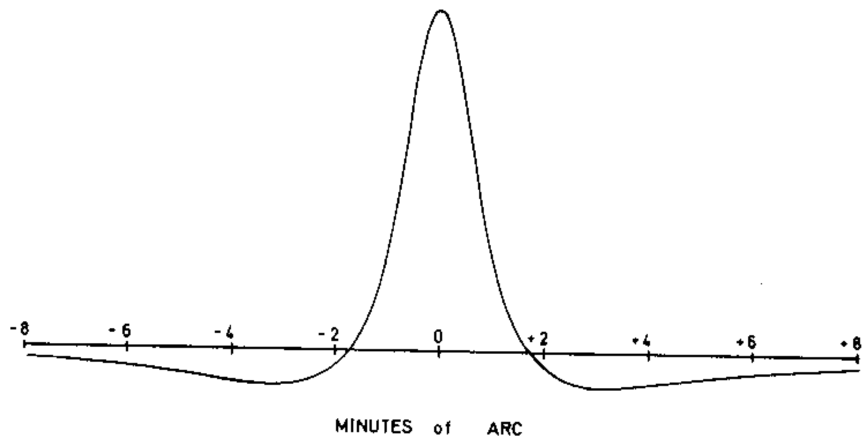


FIG. 3. The weighting function for a single channel model which leads to accurate predictions of the data on human contrast sensitivity for gratings. From John Robson.

contrast threshold to be lowest (contrast sensitivity to be highest) for gratings of intermediate frequencies, since it is at intermediate frequencies that the channel produces the biggest response (for any given amount of stimulus contrast). This is just what does happen: It is well known that human contrast sensitivity is greatest at intermediate frequencies. In fact, you can make the single-channel model agree perfectly with human psychophysical data by simply picking out a weighting function with an appropriate shape. Figure 3 shows a weighting function that works well for human contrast threshold data.

Effect of Changing the Weighting Function

For future reference, it will be useful to consider now how changing the weighting function changes the channel's response to different spatial frequencies. The kind of weighting function we've already looked at, one central excitatory area flanked by surrounding inhibitory areas, always leads to a frequency response function something like the broadest one (B) shown in Fig. 4. It is a rather wide function—the maximum sensitivity is to a spatial frequency of 14 cycles/degree, yet sensitivity is substantial even to spatial frequencies as different as 5 or 30 cycles/degree. With this sort of weighting function, there is a sizable response to spatial frequencies that differ from the best frequency by a factor of two or more. (The function that gives the peak response to various frequencies is the same as what's called "the amplitude characteristic of the Fourier transform." It can be computed easily using the methods of Fourier analysis. It is important to note that although using the methods of

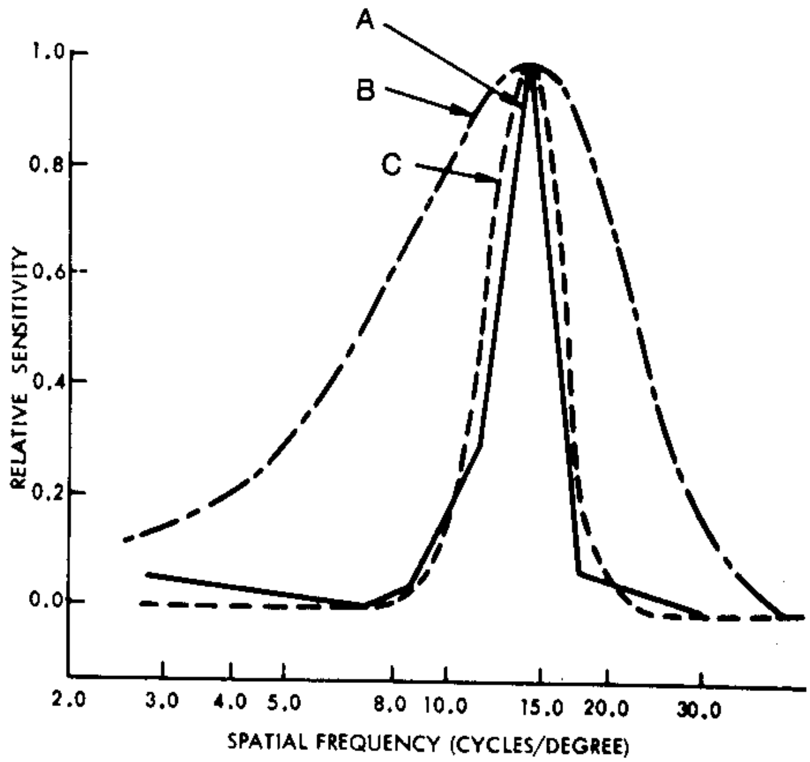


FIG. 4. Theoretical frequency response curves (B and C) for two different channels centered on the same spatial frequency but having different bandwidths. The third curve (A) is a frequency-response curve estimated from data. The vertical axis, marked "relative sensitivity," gives the peak response of the channel to a grating of some fixed criterion contrast (relative to the peak response produced by a 14 cycle/degree grating at the criterion contrast). Or equivalently, since we are considering linear channels, the vertical axis gives the reciprocal of the amount of contrast necessary to produce a peak response that reaches a criterion (relative to the contrast necessary at 14 cycles/degree). From Sachs, Nachmias, and Robson (1971).

Fourier analysis seems to imply analyzing a compound stimulus into its sinusoidal components, it does not in fact imply the use of a multiple-channels model.)

If you change the size of the weighting function (if, for example, you double the widths of both the center and surround but leave the shape unchanged), you will change the channel's best frequency (for example, from 14 cycles/degree to 7 cycles/degree) but you will not change the breadth of the response function. For example, if the original weighting function gave a response to frequencies between 5 and 30 cycles/degree, a ratio of 1 to 6, a weighting function twice as wide will give a response to frequencies between 2.5 and 15 cycles/degree, again a ratio of 1 to 6. This means that when you change only the size of the weighting function, the channel's frequency response function will keep the same shape when plotted against a logarithmic frequency scale as in Fig. 4, but it will be shifted horizontally.

Suppose you wanted to construct a channel that responds only to a very narrow range of spatial frequencies, something more like curves A and C of Fig. 4. These curves depict a sizable sensitivity only to spatial frequencies between 12 and 18 cycles/degree; a 10 or 20 cycles/degree grating gives only a negligible response. What kind of a weighting function would you need in order to produce such a narrow response curve? Figure 5 gives the rather peculiar answer: a multilobed weighting function with *several* evenly spaced excitatory and inhibitory areas.

A channel that is an array of such multilobed weighting functions is sensitive to a much narrower range of spatial frequencies than is the channel of Fig. 2.

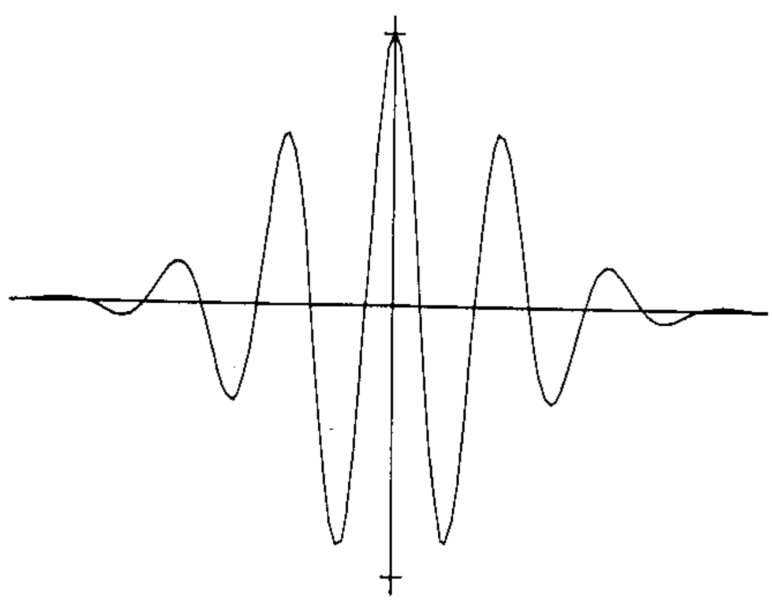


FIG. 5. The weighting function for a channel having a frequency response like that of curve C in Fig. 4. From John Robson.

To see why, first imagine a grating stimulating the channel with multilobed weighting functions, with a bright bar illuminating the central excitatory part of one particular multilobed weighting function. If the grating has the best spatial frequency for this weighting function, bright bars will fall in all the excitatory areas and dark bars will fall in all the inhibitory areas. Therefore there will be a very large response at this point in the channel. The large peak response means that the channel is very sensitive to this stimulus. (Similarly, the channel's response at the point where a *dark* bar falls on the central excitatory part of a multilobed weighting function will be a very small response, so the difference between the peak and trough responses of the channel will be large.)

Now imagine a grating with slightly narrower, more closely spaced bars, again with a bright bar falling on the central excitatory part of the weighting function. You don't need to imagine much of a change in the grating's frequency before dark bars start creeping inward into the outermost excitatory areas, thereby reducing the response. Slightly widening the bars has the same effect, as dark bars creep outward into the outermost excitatory areas. In short, any slight mismatch between the bar-spacing of the grating and the dimensions of the multilobed weighting function will lead to a much reduced response by the channel; thus the channel is sensitive to only a narrow range of spatial frequencies.

Response of a Linear Channel to Other Stimuli

Also for future reference, let's look at another convenient fact about sine waves. Knowing how the channel (or any linear system) responds to sine waves is enough to tell you how the channel responds to any stimulus at all. This apparently magical fact is true because two other facts are true: One—any stimulus at all can be treated as the sum of a number of sinusoidal stimuli; and two—a linear channel's response to a stimulus which is the sum of various component stimuli can be shown to equal the sum of the responses to the various component stimuli. So to calculate the channel's response to any arbitrary stimulus, you just need to know which sine waves the stimulus is the sum of, and then you add up the responses to those sine waves.

Sine Waves Added to Sine Waves

Let's go back to the main discussion. What does a single-channel model of the visual system predict for the response to sine waves? We have shown that a single-channel model can deal very well with thresholds for single sinusoidal gratings if an appropriate weighting function is chosen (Fig. 3). But how well can it do with thresholds for a compound grating composed of two sinusoids added together?

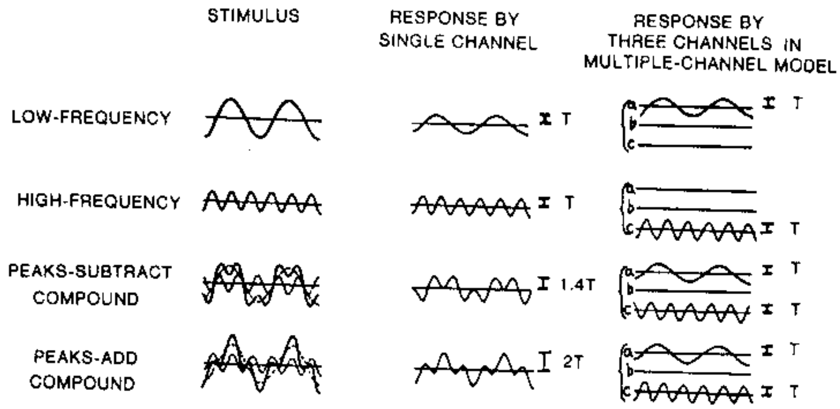


FIG. 6. Four grating patterns (left column) and the responses to them predicted by the single-channel (middle column) and multiple-channels models (right column—three channels are shown). The broken lines indicate the sinusoidal components of the compound gratings. From Graham and Nachmias (1971).

The left column of Fig. 6 presents the intensity profiles of several stimuli and the middle column shows the responses of the single channel to these stimuli. The top stimulus is a simple grating (a sine wave added to a constant luminance), with a contrast selected to put the grating at psychophysical threshold. Remember that we are assuming that "to be at psychophysical threshold" means "to produce a peak response that is as big as a certain criterion." This criterion size is labelled T (for "threshold") in Fig. 6. (In this model of the detection process, the threshold is determined entirely by the peak response. Therefore the spatial ordering of the response magnitudes in the response profile is of no importance. That is, if you took the set of response magnitudes and scrambled them into a new spatial order, you would still predict the same threshold. Another way of saying this is that in the kind of channel models that I am discussing, *all* of the spatial interactions (i.e. all the interactions that depend on distances between points) are a result of the weighting function.)

The second row shows a simple sinusoidal grating of three times the frequency of the first stimulus. Its contrast, too, was chosen to put the grating at threshold; that is, it has been adjusted so that the peak response matches the criterion T.

The third and fourth rows in Fig. 6 show two compound gratings that are combinations of the simple sinusoidal stimuli in the first and second rows. (In both combinations, the two component sine waves have been added together, and added to the same mean luminance as in the simple gratings.) In the compound grating shown in the third row, the two sine waves were positioned so that the darkest point of one coincided with the brightest point of the other (the *peaks-subtract* phase). In the fourth row, the component sine waves were positioned with the *brightest* point of one coinciding with the brightest point of the other (the *peaks-add* phase).

The single channel's response to each of these compound gratings can easily be computed from its responses to the simple gratings. Since the single channel is a linear device and each compound grating is a sum of the simple gratings, the response to the compound grating is just the sum the responses to the simple gratings. (Of course you have to position the simple response functions appropriately, so that they correspond with the positions of the simple sinusoidal components in the compound grating. In this discussion, all patterns are adjusted to have the same mean intensity. This means that, for a linear channel, the mean response across the channel is the same for all patterns and therefore can be ignored in deriving predictions that compare different patterns.) The responses shown in the second column were computed in this way. As you can see, the peak response to each of the compound gratings is now greater than T, the criterion for detection. In fact, it is 1.4 times T for the stimulus in the third row (the peaks-subtract stimulus) and 2.0 times T for the stimulus in the fourth row (the peaks-add stimulus). Therefore, according to this single-channel model, you should be able to reduce the contrast in each component of the peaks-add pattern by a factor of 2.0 and find that the peaks-add pattern would then be at threshold, because the peak response would then just equal T. You should be able to reduce the contrast in each component of the peaks-subtract pattern by a factor of 1.4 and find that the peaks-subtract pattern would be at threshold.

To put it another way, according to the single-channel model, the two compound gratings should be more visible than either simple grating, and further, the peaks-add compound should be more visible than the peaks-subtract. The brightest part of the peaks-add compound is brighter than the brightest part of the peaks-subtract compound, and both are brighter than the brightest parts of the simple gratings.

Evidence Against a Single-Channel Model

What happens when you actually measure the thresholds of humans for such simple and compound gratings? Rather strangely, but definitely, the compound gratings are *not* much more visible than the simple ones—certainly nothing like the predicted factors of 1.4 or 2.0. Moreover, relative position (phase) makes no difference—the peaks-add pattern is no more visible than the peaks-subtract.

This experimental finding is definitive evidence against the version of the single-channel model presented above. Is there some easy way to modify the single-channel model to make it fit this finding? Apparently not. Postulating compressive nonlinearities before or after the single channel doesn't help much. Instead of assuming, as we did above, that the peak response must reach some criterion in order for a pattern to be at psychophysical threshold, you might try some other assumption about the detection process. However, none of the obvious alternatives works, although some alternatives (such as probability summation across space, which will be discussed later in another context) do move the predictions closer to the data—that is, some relatively simple detection processes do lead to single-channel predictions of a difference between simple and compound stimuli that is somewhat less than the factors of 2 and 1.4 predicted by the peak-response criterion (but *not* as much less as found in the experimental data).

Of course, postulating a sufficiently *complicated* detection process as a substitute for the peak detector (which would be like postulating a large number of other stages of processing occurring after the single channel) might predict these data well and also might make an interesting model, but it would be a rather different model from those considered to date.

It should be mentioned that James Thomas and his colleagues at UCLA have done a series of experiments similar to these, involving the detection of disks of different sizes rather than gratings of different spatial frequencies. Their experiments also produced results inconsistent with a single-channel model.

MULTIPLE SPATIAL-FREQUENCY CHANNELS

So now we are left with the problem of explaining the unexpectedly low visibility of two sinusoidal gratings added together. On the basis of preliminary results somewhat similar to these results from adding up sine waves, Fergus Campbell and John Robson advanced a new model in 1968 as an alternative to the single-channel model. They proposed that the important part of the visual system for experiments like these is not a single channel but a collection of *many* channels.

Each of these multiple channels responds only to a relatively narrow range of spatial frequencies. One channel might respond only to low spatial frequencies, another only to high frequencies. The sensitivity of the whole visual system to any pattern is determined by whichever one of the multiple channels is most sensitive to the pattern. In particular, a pattern will be above threshold for the whole visual system whenever it is above threshold for *at least one* of the spatial-frequency channels.

Although one can certainly talk about these channels without specifying any particular physiological mechanism, I find it helpful to think of the multiple channels in more concrete terms. One can consider each channel as an array of receptive fields, or weighting functions, just like the single channel of the single-channel model. Each channel is specialized for a different range of spatial frequencies, so the weighting function (or the receptive field) for each channel has a different size—the channel for low spatial frequencies has a weighting function with much wider excitatory and inhibitory areas than the channel for intermediate spatial frequencies has, and so forth for other channels. This multiple channels model is quite similar to a model proposed by James Thomas, although Thomas's model was not developed to deal with sine-wave grating experiments.

Sine Waves Plus Sine Waves

To see what this multiple-channels model will predict for the threshold of two sine-wave gratings added together, let's look at the right hand column of Fig. 6. The lines labelled A, B, and C represent the responses of three different channels. Channel A is the channel that responds to the low-frequency sine wave in the left column, and it does not respond to the high frequency at all; channel C responds to the high frequency, and not at all to the low frequency; channel B doesn't respond to either one. Since the top stimulus in this figure is assumed to be at threshold and only channel A responds to it, the response in channel A must be at threshold—that is, the peak response by channel A must equal the criterion for threshold, marked T. Similarly for the second pattern: Channel C, in reacting to the second pattern, must give a peak response equal to T.

Now consider the peaks-subtract compound grating shown in the third row. How will channel A respond to it? The compound grating is the sum of the low-frequency and high-frequency sine waves pictured above it (and repeated as dotted lines in the third row). Therefore channel A's response to the compound grating is the sum of its response to the low-frequency component plus its response to the high-frequency component. Since channel A doesn't respond at all to the high-frequency component (its response profile is a flat line), its total response to the sum of the two components looks just like its response to the low-frequency grating alone. We've already said that the low-frequency grating is at threshold for channel A so the compound grating, which gives exactly the same response, must also be just at threshold. Similarly for channel C: Its response to the compound grating looks just like its response to the high-frequency component alone, so the compound grating is just at threshold for channel C. Since the compound grating is just at threshold for each channel individually, it is (according to the multiple-channels model's assumptions) just at threshold for the visual system as a whole. (The analysis differs somewhat if response variability is considered, as is done in the next section.)

So, unlike the single-channel model, and in much better accord with the psychophysical data, the multiple-channels model predicts that the peaks-subtract compound grating should be no more detectable than one of its components. We can go through the same analysis and reach precisely the same conclusion for the compound grating shown in the fourth row, the peaks-add stimulus. It is just at threshold for each channel individually and hence for the visual system as a whole. Therefore this compound grating should be no more detectable than either of its sinusoidal components, and the peaks-add and peaks-subtract gratings should be equally detectable. In other words, relative position or phase of the two components shouldn't matter at all. And that was one of the surprising aspects of the psychophysical data: Peaks-add and peaks-subtract gratings gave the same results.

Even this simple version of the multiple-channels model does quite a good job of predicting a human observer's performance when detecting these kinds of pattern: Compound gratings are (to a first approximation) no more detectable than their most detectable component, and the relative phase between components in a compound grating makes no difference to its detectability. In the next section we will find that when we take response variability into account, the multiple-channels model fits the data even more closely.

Probability Summation Among Multiple Channels

I have been talking as if there were no variability in the visual system, as if a grating with a contrast just below the threshold were invisible every time the subject looked at it and a grating with a slightly higher contrast, just above the threshold, were visible every time. But in fact there is a whole range of contrast levels for which a grating is sometimes visible and sometimes invisible. The "threshold" is arbitrarily defined as that contrast level at which the grating is seen a certain percentage of the time, usually 50%.

As Sachs, Nachmias, and Robson showed, in order to predict the thresholds for compound gratings exactly, one has to take the visual system's variability into account. It turns out that a very simple way of dealing with the variability will do, a way often referred to as "probability summation."

Consider again the response of the multiple channels to the gratings in Fig. 6. Remember that each of the sinusoidal components is individually at threshold; we've picked the appropriate amount of contrast to make that so. Thus on 50% of the trials with the low-frequency component alone, channel A's response is big enough for the observer to see something. Likewise, on 50% of the trials with the high-frequency component alone, channel C's response is big enough for the observer to see something. What happens when the compound grating is presented? According to the multiple-channels model, just the same thing as when its two components are presented separately: On 50% of the trials channel A responds to the compound grating and on 50% of the trials channel C responds. But the trials on which channel A responds are not all trials on which channel C responds (unless the variability in the two channels happens to be perfectly correlated). So on *more* than 50% of the trials *either* channel A or channel C (or both) gives a response big enough to meet the criterion for threshold. Therefore, according to the multiple-channels model, the observer sees something on *more* than 50% of the trials with the compound grating. This means the compound grating is somewhat more detectable than either of its components; how much more depends on the degree of correlation between the channels.

In fact, an assumption of complete independence (no correlation) between channels produces predictions that agree quantitatively with the data. This independence is ordinarily implied by the term probability summation. In general, *probability summation* among channels refers to the increase in the detectability of a pattern that results when two or more uncorrelated channels rather than one respond to the pattern ("summation" because there is an increase in detectability and "probability" because the increase is a direct result of the probabilistic nature of the process). Notice that probability summation can make a compound pattern more detectable than any of its components even if there is no "real" summation among components within any one channel—that is, even if, as far as the response of any one channel is concerned, presenting two (or more) components is no better than presenting one alone.

Thus, within the framework of a multiple-channels model, there are two possible causes of an increased detectability of a compound pattern relative to its components: "probability summation" resulting when channel responses are uncorrelated and more channels respond to the compound pattern than respond to any one of its components alone, and "summation within a channel" resulting when one of the channels responds better to the compound pattern than to any of its components. This distinction between kinds of summation will be very important in the next section.

An Example of Probability Summation

Probability summation can sometimes make a compound grating *substantially* more visible than either of its components. Let me give an example. The data shown in Fig. 7 come from a two-alternative forced-choice experiment that John Robson and I ran, comparing the detectability of compound gratings containing three components (rather than two as in Fig. 2) to the detectability of each of the three components alone. The components' frequencies were in the ratio of 1 to 3 to 9. In the compound gratings, the three components were arranged in either of two phases, and the relative contrasts in the three com-

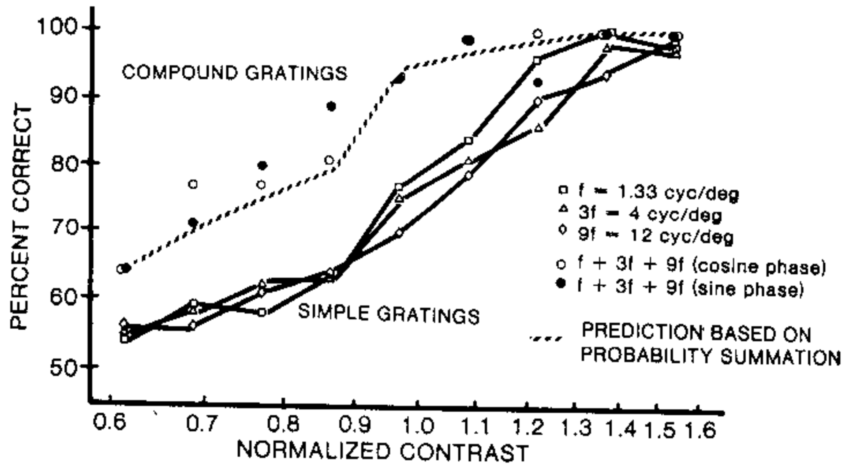


FIG. 7. Percent correct in a two-alternative forced-choice experiment for three simple gratings and two compound gratings containing the three frequencies of the simple gratings but in two different phases. In the compound gratings, the components were arranged so that their peaks coincided (cosine phase) or so that the combination approximated a square-wave grating (sine phase). The broken line is the prediction for the compound gratings based on probability summation among the three simple components.

Each of the curves in the figure has been horizontally translated (which is why the horizontal axis is marked "normalized contrast" instead of "contrast") so that the compound grating represented by any one of the circles is made up of the three components whose detectabilities are given by the three symbols directly below the circle. Notice that the contrasts of the three components in any one of the compound gratings were chosen so that all three would be approximately equally detectable.

The mean luminance of the 7.25° × 4.5° desaturated green (P31) display (29 × 18 centimeters at a distance of 2.28 meters) was 100 millilamberts, and the display was surrounded by a homogeneous white screen of approximately the same mean luminance with outer dimensions of 60 by 60 centimeters. Each trial consisted of two 600 millisecond presentations of a tone, separated by 300 milliseconds; during one tone a grating was presented and during the other the display remained unpatterned. The onset and offset of the grating were gradual, the contrast of the grating during the 600 millisecond period being proportional to $e^{-(t/100)^2}$ where t varied from −300 milliseconds to 300 milliseconds. Trials were initiated by the subject. In any one block of trials, 40 patterns (eight contrast levels in each of the three simple and two compound gratings) were presented once each in a random order. In accordance with a "staircase" rule, contrast levels were sometimes changed between blocks to keep performance at a fairly constant level. Each data point in the figure comes from between 60 and 150 trials. The observer was John Robson, viewing the display binocularly with normal spectacle corrections.

ponents were adjusted so that all three were about equally detectable when presented alone. (In accord with this adjustment of relative contrasts, the horizontal axis in Fig. 7 is labelled "normalized contrast." The normalization was done by dividing the actual contrast by a different factor for each component frequency. Each point on the horizontal axis represents three different frequencies' contrasts in the same ratio in which they appear in the compound gratings. The value of 1.0 was assigned for convenience to that normalized contrast closest to the one producing 75% correct.)

On each trial the observer had to say whether the grating was in the first or second interval. The three lower curves in Fig. 7 show the improvement in seeing each of the three simple gratings as contrast is increased. The circles up above give similar data for a compound grating. Each circle gives the detectability for a compound grating made up of the three simple components whose detectabilities are represented by the three symbols directly below the circle. The dashed line shows the predictions for the compound grating's detectability based on probability summation among the three simple components. As you can see, the fit is quite good, considering the binomial variability inherent in the data.

In a two-alternative forced-choice experiment like this one, a subject will be correct on 50% of the trials if he simply guesses randomly. The 50% guessing rate has to be taken into account when computing the probability-summation predictions. Thus, if a subject is correct on each component 50% of the time (no more than chance), the prediction from probability summation alone is that he will be correct on the complex grating only 50% of the time.

The threshold in forced-choice experiments is typically defined as the contrast needed for 75% correct. For each of the three simple components in Fig. 7, then, the threshold is at a normalized contrast of about 1.0. The threshold for the compound grating is quite a bit lower, at a contrast of about 0.8, lower than the lowest of the thresholds for any of the three components. This substantial difference, a ratio of about 0.8 (0.1 log unit), is fully explained by probability summation among the multiple channels.

It's important to bear in mind that even though the compound gratings in this experiment did have thresholds considerably lower than any of their components' thresholds, the data are still far from consistent with a single-channel model. A single-channel model would predict that when you add up three components in peaks-add phase, you should need only one-third as much contrast to put the resulting compound grating at threshold. In Fig. 7 that would be a contrast of 0.33, off the graph to the left.

Measuring the Bandwidth of a Spatial-Frequency Channel

So far I haven't said anything specific about the *bandwidth* of each of the multiple channels. How wide a range of spatial frequencies does a given channel respond to? All I've said is that the range is considerably narrower than for the single-channel model. In that model, the single channel responds to the entire visible range of spatial frequencies, so it predicts that a compound pattern should be more detectable than any one of its component frequencies, no matter how far apart they are. We know that the channels can't be *that* broadly responsive, because the data show that for widely separated frequencies, a compound pattern is no more detectable than you'd expect from probability summation among independent detectors that each respond to only one component frequency.

There is a way to estimate more precisely the range of responsiveness of an individual channel. Once Sachs, Nachmias, and Robson had shown that (allowing for probability summation among channels) the multiple-channels model accurately predicts the findings for two widely separated frequencies (which were assumed to stimulate completely separate channels), they could use this model to estimate the bandwidth of an individual channel by choosing frequencies that were quite close together.

When two neighboring frequencies were used, they found that the compound pattern was more detectable than probability summation predicts. The "extra" detectability could be attributed to summation within individual channels—to individual channels' having responded to both frequencies. They assumed that only two channels were significantly involved in the detection of any two-component grating—the two, independent channels with center frequencies equal to the two frequencies in the compound grating.

To calculate backwards from the amount of extra detectability for compound patterns to the sensitivity of individual channels, Sachs, Nachmias, and Robson had to use some assumption about the combined effect of neighboring frequencies on an individual channel, that is, about the exact form of the summation within each individual channel. For the patterns used in their experiments, their assumption was equivalent to the following model of a channel (a model that is consistent with everything said about channels so far): Each channel is a linear system exactly like the single channel of Fig. 6's middle column, except that it is sensitive to a narrower range of frequencies; a pattern is at threshold for a channel whenever the peak response across the channel meets a criterion; and the variability in a channel's response (which leads to probability summation among channels) comes from one of two equivalent sources—either the criterion varies from time to time, or the whole response profile of a channel is raised or lowered by a noise signal added to it, which varies from time to time.

In their study, Sachs, Nachmias, and Robson measured the detectability of compound gratings containing two components, one of which always had a frequency of 14 cycles/degree, and so they were able to estimate the frequency response of the channel centered at 14 cycles/degree. You've already seen their estimate of the frequency response of that channel; it is the extremely narrow curve in Fig. 4. Remember that there is nothing mysterious about an extremely narrow frequency-response curve. If you think of a channel as an array of receptive fields or of weighting functions, the frequency-response curve is narrow if the weighting function has not only an excitatory central area and an inhibitory surrounding area, but also auxiliary areas of excitation and inhibition (like Fig. 5).

Studies by Quick and by Sachs, Nachmias, and Robson suggest that, for channels centered at lower spatial frequencies, the estimated bandwidth (on a log frequency axis) may be a good deal broader than for the channel at 14 cycles/degree. If so, then for the lower spatial frequency channels, this estimated bandwidth implies that the weighting functions may be simple center-surround weighting functions.

Discrepant Estimates of Bandwidth

The bandwidth that Sachs, Nachmias, and Robson estimated for the 14 cycles/degree channel is a good deal narrower than the bandwidth usually deduced from a different type of experiment, involving adaptation or masking. The explanation of this difference is not at all clear. It could be that the channels revealed by summation experiments are not the same channels as those revealed by adaptation/masking experiments. Another possibility is that the models currently used to deduce bandwidth from summation and adaptation/masking experiments are inadequate. For instance, perhaps adaptation itself could cause an increase in bandwidth; this possibility was considered and rejected by Lange, Stecher, and Sigel (1973).

Inadequacies in the models that are used (either explicitly or, more often, implicitly) to deduce bandwidth from adaptation/masking experiments are beyond the scope of this discussion. But a possible shortcoming of the model used to deduce bandwidth from summation experiments was alluded to earlier, in the discussion of the single-channel model. There are other plausible assumptions, besides those described above, that we could make about the detection process and about the variability in the channel's responses. Some of these alternative assumptions (one example is discussed in the next section) would lead us to derive a broader bandwidth estimate from summation experiments, an estimate more in line with those from adaptation/masking experiments. However, precise quantitative agreement between such estimates based on the different kinds of experiments remains to be shown, and trying to show it may well reveal more problems. (The assumption that only two channels are involved in the detection of a two-component compound grating may well be another inadequacy of the model used by Sachs, Nachmias, and Robson. But if more than two channels are involved, using the assumption of only two probably makes the estimated bandwidth broader than the actual bandwidth. Thus, changing this assumption could only make the bandwidth estimated from summation experiments even narrower.)

Probability Summation Across the Spatial Extent of a Channel

I will now give an example of a possible detection process other than the simplest form of peak-response detection. The example is particularly appropriate because it involves another instance of probability summation. However, understanding the example is not necessary for understanding the material that follows this section.

In the models described earlier, all of the variability in a channel's responses was assumed to come from one of two equivalent sources—variability in the threshold criterion for each channel or variability in a noise signal that is added to the *whole* response profile of the channel, thereby raising or lowering the profile as a whole. Neither source of variability changes the basic shape of the response profile; points that have equal responses at one time (the peaks, for example) also have equal responses at any other time. In other words, neither of these two sources of variability entails any variation in the relative magnitudes of responses at different points across a channel.

But other sources of variability in a channel's responses are possible and are perhaps even more reasonable. After all, why shouldn't the relative response magnitudes at different points across a channel vary? In the physiological analogue, the responses at different points across a channel are produced by different neurons. If the sensitivity of neurons varies over time, and if the sensitivities of different neurons are not perfectly correlated, there would have to be variation over time in the relative response magnitudes at different points.

Let's try assuming that all the variability in a channel's responses comes from the variability of response magnitudes at individual points across the channel (and not from the two spatially uniform sources of variability mentioned above). At any one moment, the response profile will look more irregular than those in Fig. 6; some "bumps" will be higher than others, for instance. On different trials, the peak response (the highest bump) will occur at different locations—sometimes at a location that doesn't even correspond to a peak in the stimulus.

We can still assume that a pattern is at threshold whenever the peak response reaches a criterion, but now the particular location in the channel that produces the peak response will vary from trial to trial. In this model of the detection process, there will, therefore, be probability summation across the spatial extent of a channel: When there are *more* locations at which a very large response often occurs, there are more chances on any particular trial to get a very large peak response. Therefore, the channel will be more likely to detect the stimulus, and will be more sensitive to the pattern. (In this model of the detection process, as in the simple peak-response model described earlier, the spatial ordering of the response magnitudes is of no importance.)

So, for example, as the number of bars in a grating is increased, a channel's sensitivity to the grating will increase, because more points in the channel will have a chance to detect the grating. (In addition, of course, there will still be probability summation among different channels, because whether or not each channel's peak response exceeds the threshold criterion will still vary from trial to trial.)

As Granger (1973) pointed out, this new model of a detection process, including probability summation across space, predicts that a channel will show *less* summation between components of a compound grating than predicted by the simple peak-detection model. Why this is so can be seen by carefully comparing a channel's responses to simple and to compound gratings. Figure 8 shows the profiles of a channel's responses to a single grating of 12 cycles/degree and to a compound grating of 12 cycles/degree combined with 11 cycles/degree. The contrasts in the gratings were chosen so that the peak response would be the same height in both profiles. If a channel had a simple peak-detection process, therefore, it would be equally sensitive to both patterns.

However, according to the new model, the profiles in Fig. 8 represent only *average* responses. The probability of getting a peak response that meets the criterion on a particular trial does not depend solely on the peak in the average response profile. Rather, it depends on the *number of different locations* across the channel that produce large average responses (and, of course, on how large those responses are). As is clear in Fig. 8, the "beating" between components produces only a few high points in the compound grating's average response profile, whereas there are many high points in the simple grating's profile. Therefore, according to the new model, the channel will be a good deal less likely to detect the compound grating than to detect the simple grating. In short, a model allowing for probability summation across space predicts much less detectability for the compound grating relative to the detectability of the component frequencies (that is, much less summation within a channel) than does the simple peak-detection model.

The upshot is that if we were now to assume that there is probability summation across the spatial extent of a channel, then we would expect compound stimuli to be less detectable than we expected when we were assuming simple peak detection. So when we find experimentally that a compound stimulus is not very much more detectable than its components, we would no longer assume that this means each channel is very insensitive to frequencies other than its optimal one. We would conclude instead that the channel is more sensitive to nonoptimal frequencies than we had previously thought. And saying that a channel is more sensitive to nonoptimal frequencies than we had thought is the same as saying that the channel's bandwidth is broader than we had thought.

Preliminary calculations suggest that, when probability summation across space is considered, the data from Sachs, Nachmias, and Robson's experiment may be consistent with an estimated bandwidth almost as large as that of the broadest curve shown in Fig. 4. This bandwidth is substantially larger than

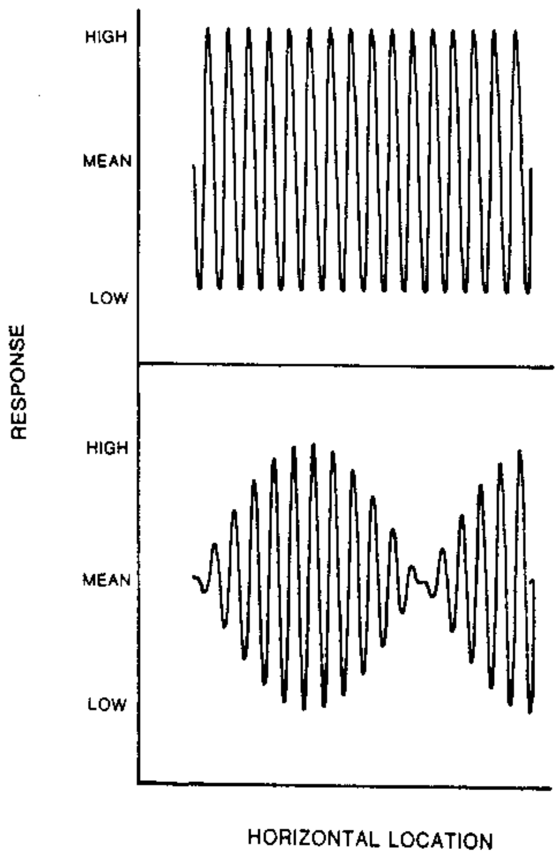


FIG. 8. The top graph is a channel's response profile to a simple grating of 12 cycles/degree, and the bottom graph is the response profile to a compound grating containing components of 11 cycles/degree and 12 cycles/degree. The contrasts of the gratings were chosen so that the two response profiles would have the same peak height. In particular, the contrasts in the two components of the compound grating were chosen so that the responses to each component alone would have the same peak height. Hence, the contrast in the 12 cycles/degree simple grating was twice the contrast of the 12 cycles/degree component in the compound grating.

that originally estimated and, if correct, the weighting function for the 14 cycles/degree channel might be of the center-surround type. However, it is an open question whether or not probability summation across space should be included in the model of a spatial frequency channel. And it is worth emphasizing that this broader bandwidth still is not nearly wide enough to make the multiple-channel model into a single-channel model.

Summary

What I have said up to this point can be summarized as follows: To explain the measured thresholds of compound gratings composed of several sinusoidal components, it is not sufficient to assume a single-channel model. It *is* sufficient to assume probability summation among multiple channels each of which responds to only a relatively narrow range of spatial frequencies. Exactly how narrow a range depends on the particular model of a channel's detection process that is assumed. The range is narrower, for example, when simple peak detection is assumed than when probability summation across spatial extent is considered. When I refer to a channel as "narrow" in what follows, I mean only that the range of frequencies to which the channel responds is narrow enough to require at least several such channels to span the range of frequencies to which a human observer is sensitive.

EDGE DETECTORS AND OTHER NEW FEATURE DETECTORS

In the rest of the chapter I would like to discuss some recent interesting experiments by Shapley and Tolhurst and by Kulikowski and King-Smith. These experiments were interpreted by their authors as evidence for other kinds of feature detectors in addition to spatial-frequency channels of the sort discussed above. First I'll review the interpretations given by the authors and then I'll go on to look at an alternative explanation that assumes there are no feature detectors other than the spatial-frequency channels.

Sine Waves Plus Broadband Stimuli

Rather than adding sine waves only to sine waves as in the experiments described above, Shapley and Tolhurst and Kulikowski and King-Smith added sine waves to broadband "test stimuli," such as edges and lines. (A *broadband stimulus* is a stimulus that can be considered to be the sum of a large number of sine-wave components, with a fairly large range of different frequencies. All nonrepetitive or aperiodic stimuli are broadband. An edge stimulus is a bright homogeneous field next to a dark homogeneous field. A line stimulus is a bright stripe superimposed on a dark field. The intensity profiles of an edge and of a line, as well as of the other aperiodic stimuli that were used, are shown in the insets of Fig. 11).

They did their experiments in the following way: They set the contrast in the sine-wave grating at some level below threshold. They then had the subject adjust the contrast in the superimposed test stimulus (in the edge, for example) until the subject could just barely see that there was a pattern present instead of a blank field. (The mean intensity of the pattern was held constant while the contrast was being adjusted.) This procedure was repeated with several subthreshold values (including zero) of sine-wave contrast.

The data they obtained were plotted as in Fig. 9: The test-stimulus contrast needed to make the compound pattern visible was plotted for each level of contrast in the subthreshold sine wave. Their actual data looked much like the fictitious data in Fig. 9. The points fell on a straight line, and the line intersected the horizontal axis at a contrast far above the threshold for the sine-wave grating alone (here called 1.0).

Frequency Responses of the New Detectors

The investigators interpreted these results within the framework of the following model: A large number of different feature detectors exist, each of these detectors is a linear system, and a stimulus is always detected by the detector that has the lowest threshold for that stimulus. Notice that there is no provision for probability summation in their model—that is, it never happens that the relative sensitivities of feature detectors fluctuate so that a stimulus is sometimes detected by one feature detector and sometimes by another.

Using their model, they could easily interpret data like that of Fig. 9. The data were gathered using added sine waves with low contrasts, including zero. All of the data points fall on a straight line as would be true if only a single linear feature detector were acting. Therefore they assumed that a single feature detector did determine all the data points. That detector would be the one with the lowest threshold for the test stimulus alone (i.e. the detector that determines the point on the plot where the sine-wave contrast is zero) and so will be called the "test-stimulus detector."

Then they could infer the test-stimulus detector's sensitivity to sine waves of various frequencies by using data like that in Fig. 9. (Why not directly measure the detector's sensitivity for a sine wave by presenting a sine wave to the observer? You can't, according to this kind of model, because when you present a sine wave by itself, its threshold is determined by whatever detector is most sensitive to the sine wave rather than by the test-stimulus detector.) The way to infer the test-stimulus detector's sensitivity is to see where the straight line through data like that in Fig. 9 cuts the horizontal axis: That intercept should tell what contrast in the sine wave would produce a threshold

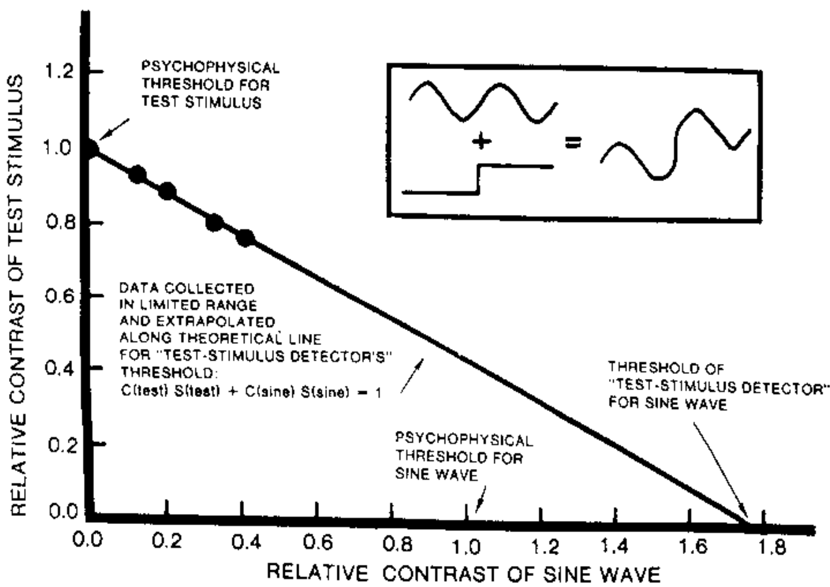


FIG. 9. Illustration of the method used by Shapley and Tolhurst and by Kulikowski and King-Smith to measure the "sensitivity of a test-stimulus detector to sine waves." Inset at upper right shows the intensity profile for one kind of stimulus they used—a combination of an edge and a sine wave. Data points are fictitious points typical of their actual data, showing, for each amount of contrast in the sine wave, how much contrast in the edge is necessary to make the compound pattern just visible to the observer. The straight line drawn through the data points is assumed to represent the responses of a linear "test-stimulus detector" whose behavior is described by the equation given in the figure, where C(test) and C(sine) are the contrasts in the test stimulus and sine, respectively, and S(test) and S(sine) are the sensitivities of the test-stimulus detector for the test stimulus and sine, respectively. (Sensitivity is, as usual, the reciprocal of threshold.)

response by the test-stimulus detector when there is no contrast at all in the test stimulus (0.0 on the vertical axis). So the reciprocal of this intercept is the test-stimulus detector's sensitivity to that sine wave.

For each test stimulus, Kulikowski and King-Smith and Shapley and Tolhurst used sine waves of a number of different spatial frequencies, finding the intercept for each one, as shown for the fictitious data in Fig. 10 (left). Then, by plotting the reciprocals of the values of those intercepts against the spatial frequency, they produced a *frequency-sensitivity curve* like that in Fig. 10 (right). This curve shows, for each frequency of sine wave, the inferred "sen-

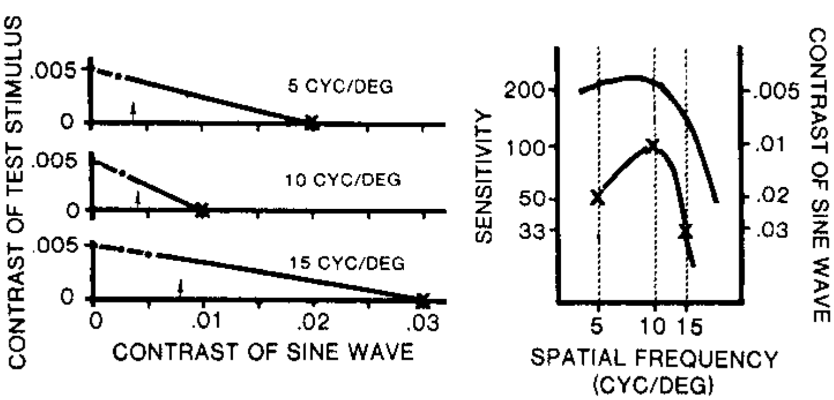


FIG. 10. The three plots on the left are the fictitious results of adding sine waves of three different frequencies (5, 10, and 15 cycles/degree) to a test stimulus. See Fig. 9 for more details about the method. The plot on the right shows the "frequency-sensitivity curve of the test-stimulus detector" that is derived from the plots on the left (and from similar plots for other frequencies). For each spatial frequency, the curve gives the reciprocal of the intercept from a plot like those on the left. (The righthand axis shows the actual value of the intercept.)

sitivity of the test-stimulus detector" for that sine wave. Notice that the steeper the line on one of the plots in Fig. 10 (left), the greater the corresponding sensitivity on the frequency-sensitivity curve in Fig. 10 (right). I will sometimes refer, therefore, to the inferred sensitivity of the test-stimulus detector for a sine wave as "the effectiveness of a sine wave in reducing the threshold for the test stimulus," as a reminder of what was actually measured.

Figure 11 shows the actual frequency-sensitivity curves from Kulikowski and King-Smith's data for six different test stimuli (one of which was a sine-wave grating as in the earlier experiments). The lower curve in each panel shows the data for the sensitivity of the test-stimulus detector. The upper curve is the psychophysical contrast-sensitivity function. Shapley and Tolhurst's curve for an edge, not shown here, is similar to Kulikowski and King-Smith's. As you can see in Fig. 11, there are at least five different curves for the six different test stimuli. (The curves for the two lines shown in the upper right and lower left panels are very similar and might be considered identical.) By these investigators' interpretation, this indicates the existence of five different detectors.

Weighting Functions of the New Detectors

Now, if these detectors are indeed linear systems (as was suggested by the straightness of the data plotted as in Fig. 9), the curves in Fig. 11 are just the

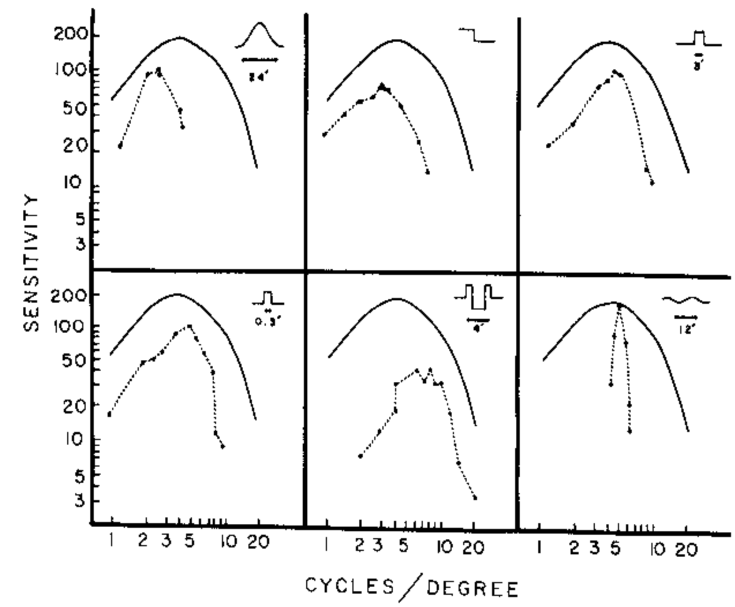


FIG. 11. Data from Kulikowski and King-Smith's experiments adding sine waves to six different test stimuli. The top curve in each panel is shown for reference and is the usual psychophysical contrast sensitivity function; it gives the reciprocal of the threshold contrast for a simple sinusoidal grating as a function of the spatial frequency of the grating. The lower curves connect the data points for the sensitivity of six "test-stimulus detectors" calculated by the method illustrated in Figs. 9 and 10. The test stimuli were a blurry bar (upper left), an edge (upper middle), a 3-minute wide line (upper right), a 0.3-minute line (lower left), a triphasic light-dark pattern (lower middle) and a sine-wave (lower right).

frequency responses of linear systems. So the curves in Fig. 11 can easily be transformed mathematically to reveal the spatial weighting functions that characterize the various detectors. (Just as the spatial weighting function can be transformed to give the frequency response by taking the Fourier transform, you can take the inverse Fourier transform of the frequency response (if you are willing to assume something about the phase characteristics of the system) in order to get the spatial weighting function.)

The results of this transformation are shown in Fig. 12. The weighting function of the edge detector consists of one excitatory region next to an inhibitory region; the grating detector has a multilobed weighting function much like that shown earlier for the original interpretation of the Sachs, Nachmias, and Rob-

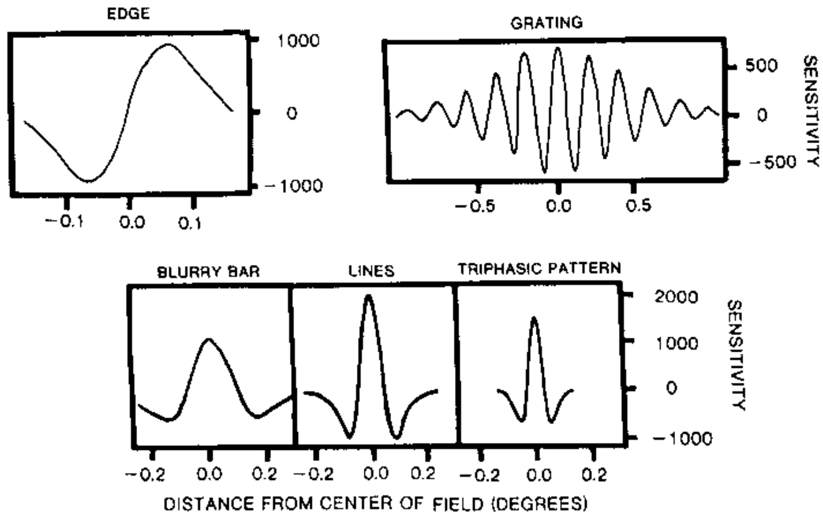


FIG. 12. The weighting functions characteristic of the detectors for which the frequency responses are shown in Fig. 11. There are only five weighting functions shown because the data for the two widths of line (upper right and lower left in Fig. 11) were so similar. These weighting functions were calculated from the data in Fig. 11 by using the assumptions that the detectors were linear and that certain kinds of symmetry would be found in the weighting functions. Notice that the scales on the axes are different for the different stimuli.

son data (Fig. 5); and the other detectors have various versions of a center-surround weighting function.

These weighting functions, strictly speaking, describe the detection of the test stimulus when the test stimulus occupies a particular location in the visual field (since the test stimulus was presented only in one location in these experiments). Presumably, however, similar responses could be evoked over a wide area of the visual field. Then you could describe the detection of an edge, for example, by an array of the appropriate weighting functions (each weighting function being an excitatory region next to an inhibitory region) spread over a large part of the visual field. You might refer to this whole array as the "edge detector" or you might refer to each weighting function as an "edge detector." Either usage will make sense in everything that follows. In any case, in line with the previous definition of a "channel," you might well refer to the whole array of weighting functions as a channel.

The channel deduced from the case where the test stimulus was a grating is akin to a spatial-frequency channel. (It might not be exactly the same as a spatial-frequency channel, because the logic used in deducing it, which ignored probability summation, might not be quite right.) The channels deduced from cases where the test stimuli were nonrepetitive (and therefore broadband) are distinguished from the spatial-frequency channels by their responsiveness to a much broader band of spatial frequencies.

Other Predictions

If the detectors are linear, one can make two kinds of testable predictions from the above data. The expected results for various other combination stimuli, like edges and lines added together, are predictable from the weighting functions (or equivalently from the sine-wave sensitivity profiles). And the actual threshold of the visual system for a test stimulus alone is also predictable. The investigators made these predictions and checked some of them against data; the predictions fit the data quite well.

BROADBAND FEATURE DETECTORS OR PROBABILITY SUMMATION AMONG SPATIAL-FREQUENCY CHANNELS?

Shapley and Tolhurst's and Kulikowski and King-Smith's experiments and interpretations make rather a pretty story, explaining quantitatively a variety of interesting results. And in some ways, edge detectors and line detectors are more appealing to common sense than are the spatial-frequency channels. We've all seen and drawn lines and edges, whereas sinusoidal gratings are a laboratory curiosity.

Too Many Channels?

However, two things bothered us about these studies. First, almost any time a new test stimulus is used a new feature detector—a new channel—is found. How many channels are we going to end up with? Somehow it seems wrong to end up with an infinite number of them. In some sense of the word, of course, there *must* be a different channel for each stimulus that we can *name* differently; if we can recognize two things as different, then somewhere in the nervous system the responses to these two things must be different; the information about the two things must be channeled differently at some point. But that is not the kind of channel I thought we were studying and it is not the kind of channel people talk about studying—we talk as if we were dealing with a limited number of feature detectors which form an early stage in visual information processing.

Probability Summation

The second bothersome thought was: "What would happen if you tried to take probability summation into account?", or'in other words, "Where have all the spatial-frequency channels gone?" Shapley and Tolhurst and Kulikowski and King-Smith's logic assumes that there is no probability summation, that whatever detector has the lowest threshold for the pattern will always detect the pattern (i.e. there will be no variability from trial to trial in the relative magnitude of the responses from various detectors and thus it will never happen that on one trial one detector has the biggest response and on another trial another detector does). Using this assumption, they could indeed rule out the possibility that any of the spatial-frequency channels detects the broadband test stimulus. According to their logic, if a spatial-frequency channel *ever* detects the broadband test stimulus (or the combination of test stimulus with a very low contrast sine wave), then it must *always* detect the broadband test stimulus. And the kind of curve plotted in Fig. 11 would look like the narrow frequency response of a spatial-frequency channel, just like the data from a sine-plus-sine experiment. Obviously, the broadband-stimulus curves don't look narrow, so it does seem that the spatial-frequency detectors play no role in detecting broadband stimuli.

The assumption that there is no probability summation was certainly a reasonable one to begin with, especially since it led to such good quantitative predictions. But we know that probability summation does occur in the sine-plus-sine experiments, so a model that ignores probability summation (like Shapley & Tolhurst's or Kulikowski & King-Smith's) cannot explain those experiments completely.

Threshold for an Edge

Once you acknowledge the presence of probability summation, it becomes much more difficult to decide whether various test stimuli are being detected by new kinds of feature detectors or by conglomerates of spatial-frequency channels. For example, you might try to make a straightforward calculation of what probability summation among spatial-frequency channels would predict about the threshold for an edge. To perform such a calculation you'd need to know how many channels there are (which we do not know), what their frequency response is (which we do not know, except at 14 cycles/degree), and what the lower part of their psychometric function looks like. With such a large area of ignorance in which to make auxiliary assumptions it is, as you might expect, possible to construct a model, involving only probability summation among spatial-frequency channels, that does accurately predict the threshold for an edge. The problem, though, is that when so many auxiliary assumptions are used to predict so little data, you cannot claim the successful prediction as a clear validation of the model.

Sine Waves Plus Broadband Stimuli

So it seemed worthwhile to see whether we could devise a testable model with fewer auxiliary assumptions, which could be checked against some other kind of data. In particular, we asked what probability summation among spatial-frequency channels would predict about the experiments using broadband test stimuli combined with sine waves. If these predictions had disagreed with Shapley and Tolhurst and Kulikowski and King-Smith's data (no matter what kind of auxiliary assumptions were made), we would at last have had a conclusive demonstration of the insufficiency of spatial-frequency channels. As it turned out, the predictions agreed with all their data using only a few reasonable auxiliary assumptions. This agreement gives considerable support to the view that the only kind of channels involved in the detection of threshold stimuli are spatial-frequency channels.

Qualitative Predictions from a Probability-Summation Model

Let's consider at a qualitative level (before going on to some quantitative predictions) what you might expect to happen when a sine wave is added to a broadband test stimulus, if you think the only relevant part of the visual system is a set of spatial-frequency channels and if you allow for probability summation.

The broadband test stimulus activates a certain subset of the spatial-frequency channels: the subset that responds to the spatial frequencies of which the test stimulus is composed. What happens when you add a sine wave to the test stimulus? If the sine wave's spatial frequency is *not* contained in the test pattern, the sine wave will not affect the channels that are responding to the test stimulus. So at low contrasts the added sine wave will have no effect at all on the threshold for a combination of itself and the test stimulus. The sine wave will not contribute anything to the detection of the test-plus-sine combination until the sine wave's contrast is high enough to strongly activate its *own* spatial-frequency channel. And then the sine wave will contribute only because its own channel is probability summating with the channels that are responding to the test stimulus, not because it is increasing the response of any of the channels that the test stimulus is activating.

However, you would expect something quite different to happen if you add a sine wave of a frequency that *is* a substantial component of the test stimulus. As soon as you add any of the sine wave at all, no matter how low its contrast, you increase the likelihood of detection. When the sine wave is added, the response of the channel tuned to the sine wave's frequency goes up from a low level (due to the test stimulus alone) to a higher level. So the threshold for detecting the test-plus-sine combination is lower than the threshold for the test stimulus alone.

Of course, if the test pattern were a sine wave with the same frequency as the sine wave you are adding, the threshold would be even lower. In that case you would be adding a sine wave to a sine wave of the same frequency, so only one channel would be involved in detecting either the test stimulus or the test-plus-sine combination. For a broadband test stimulus, though, containing many frequencies, a large number of different spatial-frequency channels is involved in detecting the test stimulus. Adding a sine wave increases the response of only one (or a few) of them, so there is only a small effect on the threshold.

More generally, if only spatial frequency channels are involved, the effectiveness of adding a sine wave to a test stimulus should depend directly on how much of that sine wave's spatial frequency is present in the test stimulus—the greater the relative amount of the spatial frequency present, the greater the effectiveness of adding it. Notice that, qualitatively, this is what does happen in the experiments (Fig. 11). When a blurry bar (which contains only low spatial frequencies) was used, only sine waves with low spatial frequencies were effective; when a triphasic light-dark-light pattern (which contains only high spatial frequencies) was used, only sine waves with high spatial frequencies were effective.

Predictions for Simple Test Stimuli

Can we calculate *quantitative* predictions from a model with probability summation among spatial-frequency channels and no other feature detectors? For the kind of experiment in which a sine wave is added to a test stimulus, it is rather easy to calculate predictions, *if* you choose a certain kind of test stimulus. Figure 13 shows the results of some calculations for four specially selected test stimuli. For these four stimuli, the calculations are easy because we don't have to worry about how many channels there are and what their bandwidths are.

One of the four test stimuli was a sine wave grating. The other three test stimuli consisted of two, three, or five sine waves added together. The sine waves in any one test stimulus were very different in frequency, so each sine wave would affect a different channel. The contrasts in the sine waves were adjusted so that each of the two, three, or five channels involved was presumably responding at the same level. (It was assumed, for convenience, that a sine wave affects only one channel.) To carry out the calculations, it is necessary to assume *some* form for a channel's psychometric function. Purely for convenience, the psychometric function used was a log-linear function that spanned a range of seven log units on the log contrast axis. Neither of these assumptions used for convenience is crucial.

Each line in Fig. 13 represents the predictions for adding a sine wave of a frequency contained in the test stimulus to one of the test stimuli. (The label

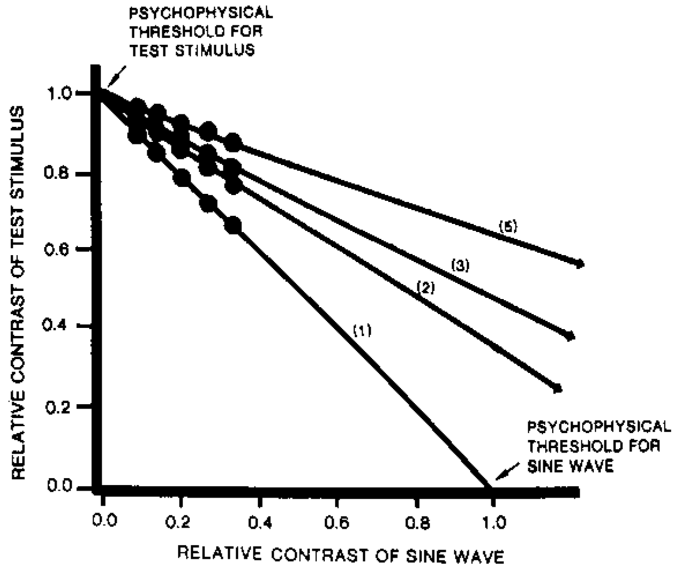


FIG. 13. Quantitative predictions for the thresholds of combinations of sine waves and certain test stimuli, assuming that there are only spatial-frequency channels with probability summation among them. The four hypothetical test stimuli consist of one, two, three, or five sine-wave components (as indicated by the numbers next to the lines). See text for further details.

next to each line identifies the test stimulus, by specifying the number of frequencies it contains.) The symbols show, for various amounts of contrast in the added sine wave, how much test-stimulus contrast is necessary to put the test-plus-sine combination at threshold.

The plots in Fig. 13, depicting predictions from a model of probability summation among spatial-frequency channels, look just like plots of data from a Shapley and Tolhurst or Kulikowski and King-Smith experiment (see Fig. 9). The points fit quite well onto straight lines (although in fact the linearity is only approximate). When the lines are extended, they hit the horizontal axis far beyond the threshold for the single sine wave. The approximate linearity simply shows there are many ways of getting a straight line.

The positions of the intercepts in Fig. 13 make sense according to the qualitative argument given earlier. To repeat briefly: When the intercept is at a higher contrast than the threshold contrast for a sine wave alone, that means that the sine wave is less effective in reducing the threshold for a broadband stimulus than it is in reducing its own threshold (i.e. when the test stimulus is also a sine wave, with the same frequency). The reason, according to a model of probability summation among spatial-frequency channels, is that the test stimulus is detected by several spatial-frequency channels (two, three, or five channels for the test stimuli used here) whereas a sine wave is detected by only one channel. Thus, when you add a sine wave to a broadband test stimulus, you assist only one of *several* channels that each contribute to detection at one time or another (producing probability summation). When you add a sine wave to a single sine wave of the same frequency you affect the one channel that is completely responsible for the detection. So, reasonably enough, the sine wave helps more in the latter case.

Figure 13 displays another property of the predictions from the model of probability summation among spatial-frequency channels: The larger the number of channels involved in detecting a test stimulus, the less it helps to add a sine wave to that stimulus (that is, the farther out the intercept of the data line with the horizontal axis). The reason for this is an extension of the argument above: When more channels contribute to detecting the test stimulus, any one channel contributes less, so the less the effect of adding a sine wave which affects only one channel.

Assumptions for a Quantitative Probability-Summation Model

The remaining sections of this chapter are for those readers who would like a more detailed derivation of quantitative predictions for the test stimuli actually used by Shapley and Tolhurst and by Kulikowski and King-Smith, instead of predictions that are qualitative or restricted to a special kind of test stimuli, as discussed so far. The derivation will show that, using only a small set of reasonable assumptions, we can calculate very good fits to the previously obtained data.

We are going to assume that the data result from probability summation among a set of spatial-frequency channels without any other, specialized, broadband detectors at work. The calculations plotted in Fig. 13 show that probability summation among spatial-frequency channels does predict the *general* type of results found experimentally when a sine wave is added to *certain* broadband stimuli. But we haven't yet demonstrated that such a probability-summation model can predict quantitatively the results for various other test stimuli as you vary the spatial frequency of the sine wave. To do so requires either some assumptions about the number of channels, their bandwidth, and their psychometric functions, or a general assumption that avoids those problems. I've chosen to make such a general assumption here because, although it produces only an approximation of the predictions of a complete model, the general assumption conveys a better idea of why the model makes the predictions it does. And anyway the approximation is not too bad.

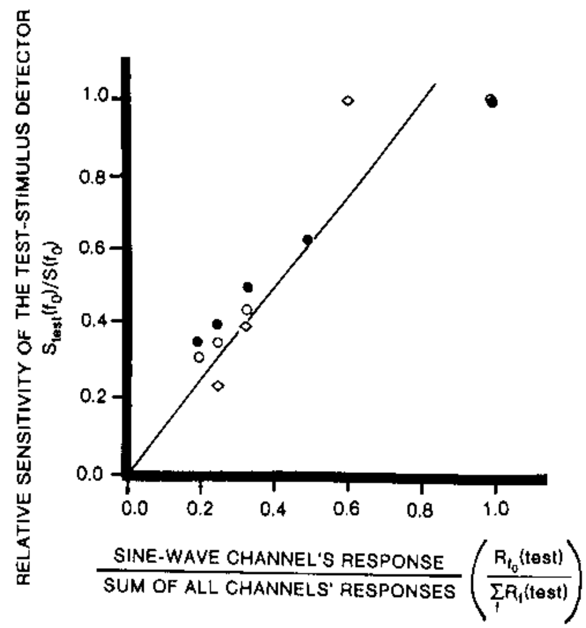


FIG. 14. Some results from sample calculations assuming only spatial-frequency channels with probability summation among them. The horizontal axis gives the proportion of the sum of the responses of all channels that is being contributed by the channel responding to the sine wave. The vertical axis shows the effectiveness of adding the sine wave to the test stimulus relative to the effectiveness of adding the sine wave to itself (or, in other words, the sensitivity of the test-stimulus detector to a sine wave divided by the sensitivity of the visual system to that sine wave). See the lower left section of Fig. 15 for definition of the symbols used on the axes.

To explain the motivation for the particular assumption I used (Assumption 1 in Fig. 15), the predictions from Fig. 13 are plotted in a different way in Fig. 14. The horizontal axis of Fig. 14 gives the contribution by the one channel that the added sine wave activates, as a fraction of the sum of the average magnitudes of all channels' responses to the test stimulus. (The average magnitude of a channel's response is simply the average peak in the channel's response profile, because, in terms of the models presented in the first half of this chapter, the peak response determines whether a channel detects a stimulus.) When the test stimulus has two sinusoidal components adjusted to affect two channels equally, and the added sine wave affects one of those two channels, the quantity on the horizontal axis is 1/2. When the test stimulus has three components adjusted to affect three channels equally, the quantity is

1/3, and so on. (For definitions of the symbols on Fig. 14's axes, see the bottom of Fig. 15.)

Figure 14's vertical axis gives the effectiveness of adding a sine wave to the test stimulus, relative to the effectiveness of adding the sine wave to itself. In other words, the vertical axis gives the "sensitivity of the test-stimulus detector" to a sine wave, divided by the sensitivity of the visual system to the sine wave. The solid points in Fig. 14 were determined from the intercepts of the lines drawn in Fig. 13. (The other points come from other kinds of sample calculations. The open circles come from calculations like those of Fig. 13 but using a log-linear psychometric function spanning 5 decibels rather than 7 decibels on the log contrast axis. The diamonds come from calculations using a test stimulus composed of two sine waves far apart in frequency where the contrasts were *not* adjusted to produce equal responding in the two affected channels but to produce several different ratios of responding; a 7 decibel psychometric function was used.)

Assumption 1 in Fig. 15 is a more general form of the relation suggested by the straight line in Fig. 14. The fact that the points in Fig. 14 fall roughly along a straight line suggests that the relative effectiveness of the added sine wave is approximately proportional to the fraction of the total response of all the channels that is contributed by the channel that the added sine wave activates. I am assuming that the proportionality shown in Fig. 14 for a few sample calculations is true for all cases of probability summation among multiple channels. (See Fig. 15 for a formal statement of this assumption.) Such an assumption has the great advantage of circumventing the problems of how many channels there are, their exact psychometric functions, etc. It is not completely accurate, because the relative effectiveness depends not only on what proportion of the total response a given channel contributes but also on the *distribution* of the responses across the other channels. And even in the case of test stimuli composed of equally balanced components, the relative effectiveness is not strictly a linear function of the fraction of total response contributed by the added sine wave's channel. However, this assumption is quite accurate enough for an investigation of whether probability summation among spatial-frequency channels can predict the kind of results found when sine waves are added to broadband stimuli.

The other assumption used here, Assumption 2 in Fig. 15, is that the average magnitude of the response of any channel to the test stimulus (more precisely, the average peak response) is proportional to the maximal sensitivity of that channel (taken to equal the contrast sensitivity of the visual system for that channel's center frequency) multiplied by the amount of that channel's center frequency which is contained in the test stimulus (the magnitude of the test stimulus spectrum at that frequency). Because this assumption is based entirely on what happens at one spatial frequency (the channel's center frequency), it is necessarily an approximation for any channel that is not extreme-

Assumption 1:

$$\frac{S_{test}(f_0)}{S(f_0)} = A \cdot \frac{R_{f_0}(test)}{\sum_f R_f(test)}$$

(justified by sample calculations)

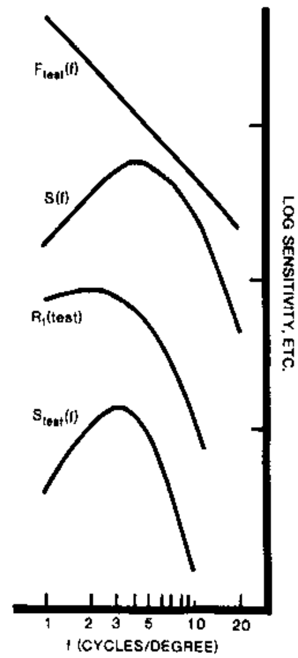
Assumption 2:

$$R_f(test) = B \cdot F_{test}(f) \cdot S(f)$$

(narrow channels with equal widths on linear frequency axis)

Conclusion:

$$S_{test}(f_0) = \frac{A \cdot B}{\sum_f R_f(test)} \cdot F_{test}(f_0) \cdot [S(f_0)]^2$$



$S_{test}(f_0)$ = sensitivity of "test-stimulus detector" to frequency f_0 as derived by Shapley-Tolhurst-Kulikowski-King-Smith method
= effectiveness of a sine wave of frequency f_0 in reducing threshold of test stimulus

$S(f_0)$ = psychophysical contrast-sensitivity function
= sensitivity of channel centered at frequency f_0 to sine wave of frequency f_0

$R_{f_0}(test)$ = average magnitude of the response of channel centered at frequency f_0 to the test stimulus

$F_{test}(f_0)$ = amount of frequency f_0 contained in test stimulus
= spectrum of test stimulus

A and B are constants, the values of which are not known

FIG. 15. (Left) The assumptions and conclusions of a model of probability summation among multiple spatial-frequency channels for predicting the results of experiments in which sine waves are added to broad-band test stimuli.

(Right) Illustration of the calculations involved in the probability summation model. The top curve shows the spectrum of an edge—the function giving the amount of each spatial frequency present in an edge. The second curve is the psychophysical contrast-sensitivity function (from Kulikowski & King-Smith)—the function giving the reciprocal of the contrast threshold for simple sinusoidal gratings. The third curve is the product of the first two and, according to Assumption 2 of the probability-summation model, is proportional to the responses that an edge produces in channels centered at different spatial frequencies. The fourth curve is the product of the second and third and is proportional to the "sensitivity of the edge detector" as predicted by the probability summation model.

ly narrow (any channel with greater than zero bandwidth). But it is a very good approximation for the spatial-frequency channels since, as can easily be shown by direct calculation from the models described in the first part of this chapter, the approximation is good even for quite wide channels.

The relation of the constant of proportionality, B, to the bandwidth of the channel does depend on the particular detection process that is assumed. To predict a given value for the constant of proportionality, B, you need to assume a larger channel bandwidth when using the detection-process model that allows for probability summation across space than when using simple peak detection (in which case bandwidth will actually equal B if you define the spectrum of test stimuli carefully). The additional assumption, embodied in Assumption 2, that the same constant of proportionality B holds for every channel across the whole frequency range is tantamount to the assumption that every channel has the same bandwidth, measured on a linear frequency axis (thus on a log frequency axis the bandwidth is broader for a low-spatial-frequency channel than for a high). Quick's study suggests that this assumption is quite reasonable.

Assumption 2's specification of the average magnitudes of the spatial-frequency channels' responses to an edge is illustrated by the top three lines in the right half of Fig. 15. The topmost line is the spectrum of an edge—how much of each spatial frequency is present in the edge. The second, curved line is the psychophysical contrast-sensitivity function (from Kulikowski & King-Smith's study), which also tells us the peak sensitivities of the channels centered at various spatial frequencies. The third line is the product of multiplying the functions in the first two. By Assumption 2, the value of this product at a given spatial frequency is proportional to the average magnitude of the response to an edge by the channel centered at that spatial frequency.

Quantitative Predictions from a Probability-Summation Model

Putting Assumption 1 together with Assumption 2, we can easily derive a quantitative prediction (see Fig. 15): The "sensitivity of the test-stimulus detector" (what I've been calling "the effectiveness of a sine wave in reducing the threshold for the test stimulus") should be proportional to the spectrum of the test stimulus multiplied by the square of the contrast-sensitivity function. Or in other words, the sensitivity of the test-stimulus detector to a given frequency is predicted to be proportional to the average response magnitude of the channel centered at that frequency multiplied by the psychophysical contrast sensitivity for that frequency.

The predicted "sensitivity of the edge detector" is given by the bottom curve in Fig. 15, which is the product of the second and third curves. Similar predictions can easily be made for the other test stimuli used in Shapley and Tolhurst's and Kulikowski and King-Smith's experiments.

The constant of proportionality for these predictions of test-stimulus detector sensitivity, as can be seen in Fig. 15, is equal to the product of the two unknown constants of proportionality from Assumptions 1 and 2 (the constants do not depend on which test stimulus you are considering) divided by the sum of all the channels' average responses to the test stimulus (this sum does, of course, depend on which test stimulus you are considering). Thus, if you cannot estimate the sum of all the channels' average responses, you are left with a different constant of proportionality for each test stimulus. In that case, you'd have to fit the predicted sensitivities (e.g. bottom curve, Fig. 15) to the actual data (Fig. 11) separately for *each* test stimulus, by finding the constant of proportionality that produces the best fit. (In practice, you plot both the predicted sensitivities and the actual data on log-sensitivity axes and shift the predictions vertically to get the best possible fit to the data.)

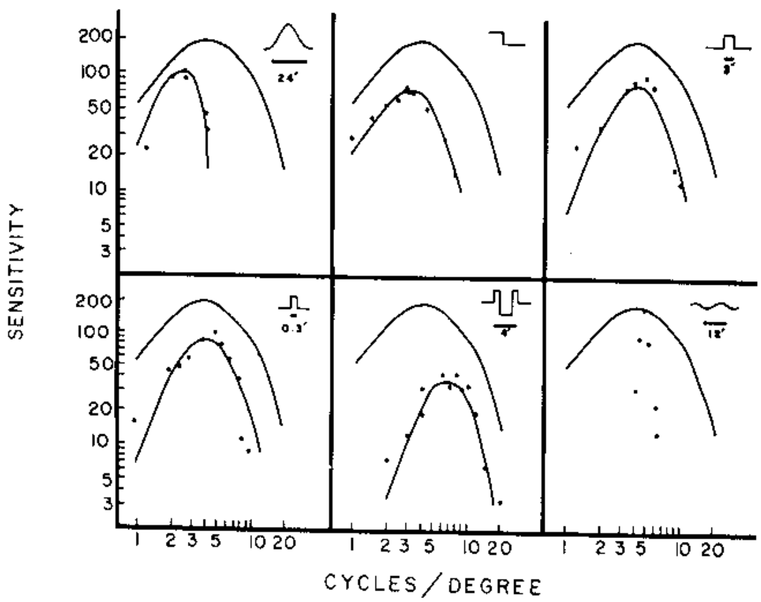


FIG. 16. Comparisons between predictions from the probability-summation model and data from Kulikowski and King-Smith's experiments. The solid points are the "sensitivities of the test-stimulus detector" to sine waves of different frequencies, as calculated from the experimental data. (The same data points were shown in Fig. 11.) The lower curve gives the predictions from the probability-summation-among-spatial-frequency-channels model. (The upper curve is the usual psychophysical contrast-sensitivity function.)

(A) Data and predictions for experiment using a blurry bar. (B) Edge. (C) 3.0-minute wide line. (D) 0.3-minute line. (E) Triphasic light-dark-light pattern. No predictions are made for the sine wave grating since it is not a broadband stimulus.

Figure 16 shows how well the data collected by Kulikowski and King-Smith are fit by predictions based on probability summation among spatial-frequency channels. In each section, the top curve is the visual system's overall contrast-sensitivity function, shown for reference. The points are the experimental data. The bottom curve is the prediction from probability summation among spatial-frequency channels. (There are no predictions for the sine-wave grating since it is not a broadband stimulus.) As you can see, the predictions are a very good fit to the data. They are not perfect, of course, but neither are the data. (By making an additional assumption, you can fit the data almost as well with just one free parameter. See the note on page 260-261.)

Remaining Problems

As mentioned earlier, Shapley and Tolhurst and Kulikowski and King-Smith showed that a few examples of two other kinds of data were quantitatively consistent with their results from the test-stimulus-plus-sine experiment: the results from test-stimulus-plus-another-test-stimulus experiments (such as edge-plus-line) and the thresholds for each test stimulus alone. It can easily be argued, on the basis of some sample calculations, that the consistency found between the data from the test-stimulus-plus-sine experiments and the data from the test-stimulus-plus-test-stimulus experiments would be expected from a model of probability summation among spatial-frequency channels. What cannot be quantitatively predicted on the basis of probability summation among spatial-frequency channels using the approach presented here are the actual threshold contrast values for various test stimuli alone.

If instead of the assumptions used here, an explicit model of probability summation among channels is constructed, then the thresholds for test stimuli can be predicted and, in the process, estimates of channel density and bandwidth are obtained (Graham, 1977). This estimate of channel bandwidth is in good agreement with the estimate from the sine-plus-sine experiments like those of Sachs, Nachmias, and Robson and of Quick. Both kinds of estimates depend in similar ways on the exact model assumed for the channel; that is, on whether or not probability summation across space is included in the model.

A possible shortcoming of the probability summation model is that, in the long run, even one free parameter to fit data is one too many—its value may prove to be inconsistent with some other kind of data.

Finally, even if the collection of spatial-frequency channels *could* detect broadband stimuli, it might not actually do so. An observer might, for example, ignore the spatial-frequency channels when engaging in tasks for which some other part of the visual system seemed more appropriate. Or there might be inhibition among spatial-frequency channels, so that when many of them are responding, none responds very well.

Conclusion

At the moment, despite some remaining questions, I feel there is no way to rule out the hypothesis that only relatively narrowband spatial-frequency channels, with probability summation, are involved in the detection of the various kinds of stimuli used in these experiments. That hypothesis means that data of the kind shown in Fig. 16, for example, from an experiment in which sine waves of various frequencies were added to a broadband test stimulus, might be better viewed as the result of probability summation among several relatively narrowband channels than as the frequency response of a single "test-stimulus detector."

As was explained earlier, however, these "relatively narrowband" spatial-frequency channels may not be as narrowly tuned as was originally deduced from the sine-plus-sine experiments. If probability summation across space does occur, the bandwidth of these channels may turn out to be much broader than we at first thought.

To put this last point another way, suppose that we knew that an early stage of the visual system consists of a set of spatial-frequency channels. Suppose that each channel is an array of identical receptive fields (identical in all characteristics except location in the visual field), but the characteristics of the receptive fields (size, in particular) vary from channel to channel. Suppose further that there is probability summation across space (across different locations) and across channels (across different kinds of receptive field). What, then, would be the bandwidths of these channels (what would be the Fourier transforms of the weighting functions associated with the various channels)? The answer is not yet known. But this much can be said. The bandwidths will probably be greater than those deduced from sine-plus-sine experiments when probability summation across space was ignored (although probability summation across channels was considered). But the bandwidths will probably be narrower than those deduced for test-stimulus detectors from broadband-test-plus-sine experiments when probability summation across channels was ignored (probability summation across space was not very important because all the test stimuli occupied fixed locations in the visual field). And such an intermediate bandwidth is what I mean by "relatively" narrowband.

Perhaps we are wrong to consider the results of detection-summation experiments like those described here as telling us anything about a limited set of parallel feature detectors (spatial-frequency channels) early in the chain of visual information processing. Perhaps we would do better to consider all the data as resulting from much more complex processes. But if we are going to consider these experiments as revealing the existence of parallel feature detectors, it seems possible to explain all the psychophysical data described above quite simply: You do *not* need to conclude that there are broadband detectors, like edge detectors and line detectors, in addition to the relatively narrowband spatial-frequency channels. Relatively narrowband channels alone, with probability summation among them, would be sufficient. The exact characteristics of these channels, however, remain to be determined.

FURTHER COMMENTS FOR INTERESTED READERS

Terminology

Channels. Rather than using a purely arbitrary word to name a concept, a person often chooses a word suggestive of the concept. Unfortunately, what the word suggests to one person may not be what it suggests to another. I have used the word *channel* to mean a two-dimensional array or, in terms of the physiological analogue, a collection of receptive fields that are identical except in position. I have heard at least two objections to this use of the word.

To some people, a "channel" should be something that produces only a single number as its output. These people might prefer to call each single neuron a channel. Or they might consider a channel to consist of the kind of array that I call a channel *plus* a "detector" whose output is either the height of peak response or perhaps a 0 or 1 depending on whether the peak response exceeds the criterion level. (The terminology I have used is consistent with that used in audition. The input for an auditory channel is an acoustic waveform whose amplitude varies with time. The output is also a waveform, filtered, whose amplitude varies with time. The output is a single number only for a single instant. When we draw the analogy between audition and vision, visual space takes the place of auditory time. So it is consistent with the usage in audition to consider both the input and output of a visual channel to be a waveform whose amplitude varies with spatial location.)

To some other people, a "channel" should be something that produces a distinctive perceptual effect—that is, the outputs of different channels should be kept quite separate and should make qualitatively different contributions to perception. Although my definition of "channels" does not exclude their having qualitatively different effects, the definition in no way requires it. And I do not want to require it, at least not in this context, since it is irrelevant for the kinds of experiments described here.

Detectors. There are several possible uses of terms like "line detector." If the bandwidth of the spatial frequency channels turns out to be wide enough, the weighting functions associated with the channels will be of the simple center-surround type (if symmetric). Each spatial frequency channel might then be called an array of "line detectors." These "line detectors," however, would *not* be the detectors deduced by Kulikowski and King-Smith from the experiments using combinations of lines and sine waves—that is, their weighting functions would not be the same, except by accident.

There is another possible use of the term "line detector." If it does turn out that the detection of lines is done by a subset of the spatial-frequency channels, that subset might be called a "line detector."

Modifications of the Model

Some slight modifications of the multiple-channels models described in this chapter will have little or no effect on these models' predictions. Three examples follow.

One peak detector or several? It was assumed here that each channel (each array of neurons having the same kind of receptive fields) has its own peak detector—that is, it was assumed that a pattern is above threshold for the visual system when it is above threshold for at least one channel, and that it is above threshold for a particular channel whenever the peak response in that channel (the response of the neuron that gives a bigger response than any other neuron in that particular channel) is above some criterion. However, without changing the predictions at all, one could instead assume that there is only one peak detector associated with the whole collection of channels—that is, it could be assumed that a pattern is above threshold whenever the peak response in the whole collection of channels (the response of the neuron that gives a bigger response than any other neuron in *any* channel) is above the criterion.

Peak or peak-trough detection? We could assume that the threshold is based on peak-trough detection (the difference between the highest and lowest points in a channel's response profile) instead of on peak detection (the difference between the highest point and the average across the profile). For very narrowband channels, the change in assumption would not affect predictions at all. For channels with slightly wider bandwidths (like the current idea of a spatial-frequency channel), the change will affect the predictions somewhat. (Assuming peak detection instead of peak-trough detection improves the approximation contained in Assumption 2 of the model summarized in Fig. 15.)

Retinal inhomogeneity. A third modification that might make little difference in the predictions of the multiple-channels model is based on retinal inhomogeneity. All sizes of receptive fields may not be present at all places in the retina. The small receptive fields subserving high-spatial-frequency channels may be located within and near the fovea, whereas the broader receptive fields subserving low-spatial-frequency channels may be located more peripherally. Whether or not retinal inhomogeneity makes a substantial difference in the predictions for aperiodic stimuli depends on several factors, including the bandwidth of the channels and the exact distribution of receptive field sizes. Retinal inhomogeneity remains a potentially important factor that has not been adequately explored.

Effect of Limited Extent of Gratings

Embodied in Assumption 2 of the model summarized in Fig. 15 is the assumption that the peak sensitivity of a channel is correctly estimated by the visual system's contrast sensitivity for that channel's center frequency. This assumption is, unfortunately, introducing an approximation that may vary systematically with spatial-frequency. The peak sensitivity of a channel in the model is the sensitivity for a grating containing only a single sinusoidal component. But in order to have only a single component, the grating would have to be infinite in extent. The contrast sensitivity of the visual system measured in an experiment is of course based on sinusoidal gratings that are limited in extent. In fact, changing the extent of gratings is known to change the shape of the contrast-sensitivity function. In other words, changing the extent changes the estimates of peak sensitivities by different factors for different channels.

There are good reasons, which should be incorporated into a more complete model of multiple channels, why varying the extent of gratings might have this experimentally observed effect. For one thing, a sine-wave grating that is limited in extent contains a band of frequencies in addition to the nominal frequency and thus stimulates a number of channels. For another, if the extent of a sine-wave grating is small enough, and the weighting functions are multilobed, the whole grating will be narrower than the weighting functions of the most sensitive channels, and then the peak response in these channels will be smaller than the peak in the channels' responses to an infinite grating. Further, the detection process in channels might depend on the extent of a grating in a way that would produce the observed effect of varying extent. (However, if the detection process is either simple peak detection or peak detection with probability summation across space, it would not depend on extent.) Finally, as mentioned above, the retina is not completely homogeneous, so the response may be different depending on how much of the retina is stimulated by the grating.

Reducing the Number of Free Parameters

If you fit the data for each test stimulus separately as was done in Fig. 16, you are using as many free parameters as there are test stimuli. (The original investigators used as many free parameters as there were data points.) It turns out that you can fit the data almost as well, and yet reduce the number of free parameters to *one*, by making an additional assumption. This assumption allows you to estimate a quantity that is proportional to the sum of all the channels' average responses to the test stimulus, regardless of what the test stimulus is. What you assume is that the channels are evenly spaced along the linear spatial-frequency axis (there is no evidence on this one way or another). Then the sum of the average responses of *all* the channels will be approximately proportional to the sum of the average responses of a *subset* of channels which have center frequencies evenly spaced along the linear spatial-frequency axis. (In my calculations I used a spacing of 1 cycle/degree.) You are then left with only one free constant (A times B divided by the constant of proportionality used in estimating the sum of the average responses), and that single constant can be adjusted to fit the data for all the test stimuli simultaneously.

The only difference between the predictions that are obtained if one free parameter is allowed and the predictions shown in Fig. 16 is that the vertical position of the predicted curve for the blurry bar is higher relative to the vertical positions of the predicted curves for other stimuli. This change in relative vertical position produces somewhat less impressive, although not bad, fits between data and predictions.

REFERENCES

(This list contains only those papers directly described in this chapter plus several recent papers to give the reader an entry into the current technical literature.)

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

Graham, N. Visual detection of aperiodic spatial stimuli by probability summation among narrowband channels. *Vision Research*, 1977, *17*, 637-652.

Graham, N., & Nachmias, J. Detecting of grating patterns containing two spatial frequencies: A comparison of single and multiple channel models. *Vision Research*, 1971, *11*, 251-261.

Graham, N., Robson, J. G., & Nachmias, J. Visual detection of aperiodic spatial stimuli by probability summation among narrowband channels. *Vision Research*, 1978, *18*, 815-825.

Granger, E. An alternative model for grating detection. Paper presented at the meetings of the Association for Research in Vision and Ophthalmology, Sarasota, Florida, May, 1973.

Kulikowski, J. J., & King-Smith, P. E. Spatial arrangement of line, edge and grating detectors revealed by subthreshold summation. *Vision Research*, 1973, *13*, 1455-1478.

Lange, R. V., Sigel, C., & Stecher, S. Adapted and unadapted spatial frequency channels in human vision. *Vision Research*, 1973, *13*, 2139-2143.

Quick, R. F., Mullins, W. W., & Reichert, T. A. Spatial summation effects on two-component grating thresholds. *Journal of the Optical Society of America*, 1978, *68*, 116-121.

Sachs, M. B., Nachmias, J., & Robson, J. G. Spatial frequency channels in human vision. *Journal of the Optical Society of America*, 1971, *61*, 1176-1186.

Shapley, R. M., & Tolhurst, D. J. Edge detectors in human vision. *Journal of Physiology,* 1973, *229,* 165-183.

Thomas, J. P. Model of the function of receptive fields in human vision. *Psychological Review,* 1970, *77,* 121-134.

Wilson, H., & Bergen, J. A four mechanism model for spatial vision. *Investigative Ophthalmology and Visual Science, Supplement,* 1977 (April), 46. (Abstract)