These and other studies indicate that the visual system exhibits a **critical period**: an interval of time during which it undergoes rapid development owing to the effects of environmental stimulation. Both before and after the critical period, few effects of experience are found. This means that the effects of visual deprivation during the critical period are particularly debilitating. Normal gains in visual function are not made because of lack of stimulation, and the resulting deficits cannot be overcome because later stimulation has little or no effect. This is true of visual development in humans as well as animals. For example, the effects of **cataracts**—the clouding of the eye's lenses—depends on when they develop and when they are removed. Cataracts that develop after the eighth year of life have almost no lasting effects once they are removed because the critical periods of visual development are over. But cataracts that are present from birth have devastating effects on vision if they are not removed surgically during the first few months of life, a fact that was discovered when effective medical procedures were first developed to remove such cataracts.

Once perfected, cataract removal operations were performed on many adults who had had cataracts from birth, thus giving them vivid visual experiences for the first time. It was hoped that their restored vision would be normal, allowing them to live fully sighted lives. The results of these operations were generally less effective than had been hoped because patients were not able to perceive the environment as normally sighted people do and appeared to be unable to learn to do so. As the literature on visual development now makes plain, normal vision includes important components of maturation and learning that can take place only in the presence of normal stimulation. Patients whose vision was restored as adults sometimes became depressed after the operation, and some even chose to live much of their lives in darkness rather than having to deal with the confusing and chaotic visual experiences that intruded into their lives (Gregory, 1970; Von Senden, 1960).

There is no single critical period in visual development for all visual properties, but there are different critical periods for different properties. The critical period for orientation selectivity, for example, appears to be from one to five weeks of age in cats (Blakemore, Van Sluyters, & Movshon, 1976). That for ocular dominance occurs somewhat later, at five to ten weeks (Daw & Wyatt, 1976). In general, it appears that the critical period for a given type of cortical cells depends on its level in the visual system: Those of lower-level cells occur sooner than those of higher-level cells. This hypothesis fits the current data because orientation selectivity is a property of cells in the input layer of cortex, whereas ocular dominance is characteristic of the output layers (Shatz & Stryker, 1978). This pattern of development makes sense because higher-level cells can develop their response properties only after lower-level cells have developed theirs.

4.2 Psychophysical Channels

Surely one of the most interesting facts about image-based spatial processing is that the functional interpretation of the cells that Hubel and Wiesel discovered over 30 years ago is still very much in dispute. There are some relatively minor disagreements about the precise shapes of the receptive fields and some of the variables that are important in specifying them, but these are completely overshadowed by the controversy that rages over their functional significance: What are these cells *doing*? Thus far, we have considered only one of the contenders in detail—namely, the line and edge detector hypothesis. We briefly mentioned that there is an alternative hypothesis but have not yet explained it. We will now take a closer look at this theory, how it evolved from behavioral research in human visual psychophysics, and how it suggests a different view of spatial processing in area V1.

This second approach to image-based processing arose within a branch of sensory psychology known as **psychophysics**. As explained briefly in Chapter 2 and more thoroughly in Appendix A, psychophysics is the study of quantitative relations between people's conscious experiences (their psyche) and properties of the physical world (physics) using behavioral methods. Calling the methods "behavioral" indicates that psychophysicists, unlike physiologists, do not record electrical events in neurons or directly measure any other aspects of neural activity. Instead, they measure people's performance in specific perceptual tasks and try to infer something about underlying mechanisms from behavioral measurements. For instance, a psychophysicist might be interested in studying how bright a spot has to be within a darker surrounding field for it to produce a just barely perceptible experience. This is called the *threshold* for detecting the spot. A psychophysicist might study how this threshold depends on factors such as the size of the spot, the darkness of the surrounding field, or the length of time the subject has been sitting in the dark before testing. Appendix A explains the standard psychophysical methods for measuring various kinds of thresholds.

From the answers to questions about sensory thresholds and how they are affected by other variables, psychophysicists try to understand the mechanisms underlying people's performance. Since these mechanisms must ultimately be implemented physiologically, there should be a great deal of overlap in the subject matter of psychophysical and physiological approaches to vision. Often there is. For example, in the study of color vision, there has been a satisfying convergence of insights from these two domains that have mutually reinforced each other, as we discovered in Chapter 3. In the study of spatial vision, however, there has been less convergence than one might expect. Indeed, we will see that psychophysical theories of image processing have developed in quite a different direction from physiological theories of line and edge detectors.

For nearly 30 years the psychophysical community has been working within a theoretical framework called the *spatial frequency theory*. It dominates psychophysical theories of spatial vision because it is able to explain a large number of important and surprising results from psychophysical experiments. Unfortunately, it is also a rather complex and technical theory, so we will have to cover a fair amount of background material to understand it. Once we have done so, however, a very different conception will emerge of what the cells that Hubel and Wiesel discovered in striate cortex might be doing.

4.2.1 Spatial Frequency Theory

Like the line and edge detector theory of image processing, the **spatial frequency theory** of image processing is based on an atomistic assumption: that the representation of any image, no matter how complex, is an assemblage of many primitive spatial "atoms." The primitives of spatial frequency theory, however, are quite different from the lines and edges that we considered in the previous section. Rather, they are spatially extended patterns called **sinusoidal gratings**: two-dimensional patterns whose luminance varies according to a sine wave over one spatial dimension and is constant over the perpendicular dimension. Figure 4.2.1 shows examples of various sinusoidal gratings. The graph (top right) shows the luminance profile of the sinusoidal grating. This graph plots the reading of a tiny light meter that traverses the grating perpendicular to the orientation of the stripes. Notice that the light and dark bars look fuzzy or "out of focus." This is because the changes in luminance over space are smooth and gradual instead of sharp, as indicated by the smooth continuous curve of the sinusoidal graph.

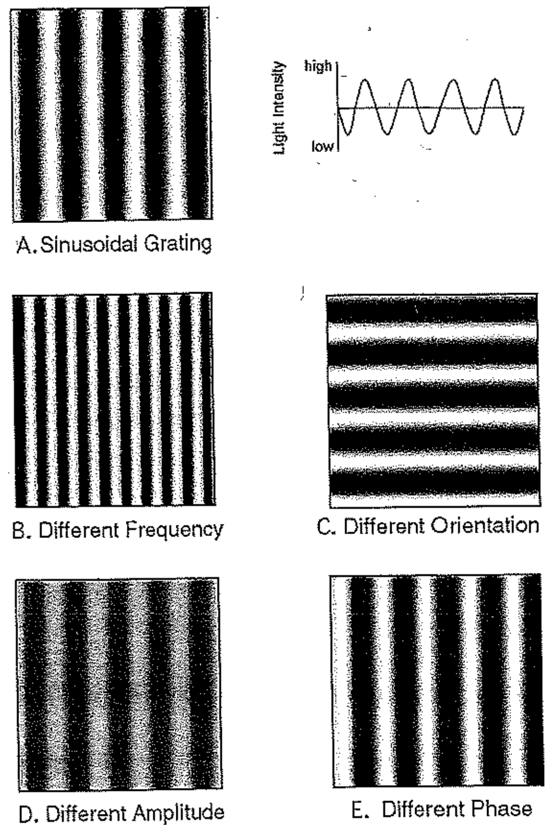


Figure 4.2.1 Sinusoidal gratings. Sinusoidal gratings are shown for (A) a standard grating and four comparison gratings of (B) a lower spatial frequency, (C) a different orientation, (D) a different amplitude, and (E) a different phase. A graph of grating A is shown to its right.

Each primitive sinusoidal grating can be characterized completely by just four parameters: its *spatial frequency*, *orientation*, *amplitude*, and *phase*. Figure 4.2.1 shows how these parameters change the appearance of the grating.

1. The **spatial frequency** of the grating refers to the width of the fuzzy light and dark bars: Low-frequency gratings have thick bars, and high-frequency gratings have thin ones. Spatial frequency is usually specified in terms of the number of light/dark cycles per degree of visual angle, a quantity that varies inversely with stripe width. In Figure 4.2.1, grating B differs from all the others in having a higher spatial frequency.

2. The **orientation** of the grating refers to the angle of its light and dark bars as specified in degrees counterclockwise from vertical. In Figure 4.2.1, grating C differs from all the others in having a horizontal orientation.

3. The **amplitude** (or **contrast**) of the grating refers to the difference in luminance between the lightest and darkest parts, which corresponds to the difference in height between the peaks and the valleys in its luminance profile. Contrast is specified as a percentage of the maximum possible amplitude difference, so 0% contrast is a uniform gray field (since there is zero difference between the lightest and darkest parts), and 100% contrast varies from the brightest white to the darkest black. In Figure 4.2.1, grating D differs from all the others in having a lower amplitude.

4. The **phase** of a grating refers to the position of the sinusoid relative to some reference point. Phase is specified in degrees, such that a grating whose positive-going inflection point is at the reference point is said to have a phase of 0° (called *sine phase*), one whose peak is at the reference point has a phase of 90° (*cosine phase*), one whose negative-going inflection point is at the reference point has a phase of 180° (*anti-sine phase*), and one whose valley is at the reference point has a phase of 270° (*anti-cosine phase*). In Figure 4.2.1, grating E differs from all the others in its phase.

Fourier Analysis. It might seem odd to consider sinusoidal gratings as primitives or atomic elements for spatial vision. After all, we don't consciously experience anything like sinusoidal gratings when we look at naturally occurring scenes. If conscious perception of visual elements were a necessary condition for their having primitive status, the case would be far stronger for bars and edges than for sinusoidal gratings. At least we see bars and edges in natural scenes. There is no reason to suppose that primitive elements in early spatial vision need to be conscious, however. We do not experience tiny points of color that are presumably signaled by the output of the three cone types, for example, yet they are surely the initial set of primitives in the visual system.

There is actually a good theoretical reason for choosing sinusoidal gratings as primitives, but it is a formal mathematical reason rather than an experiential one. The rationale is based on a well-known and widely used mathematical result called *Fourier's theorem*, after the French physicist and mathematician Baron Jean Fourier, who proved it in 1822. As applied to the 2-D image processing problem, **Fourier analysis** is a method, based on Fourier's theorem, by which any two-dimensional luminance image can be analyzed into the sum of a set of sinusoidal gratings that differ in spatial frequency, orientation, amplitude, and phase. Two very simple examples of how sinusoidal gratings can be combined to form more complex images are shown in Figures 4.2.2 and 4.2.3. In Figure 4.2.2 a series of sinusoidal gratings of the same orientation at spatial frequencies of f, $3f$, $5f$, ... are added together in the proper amplitude and phase relationships to obtain a square wave that has sharp edges rather than fuzzy ones. Figure 4.2.3 shows how two such square waves at different orientations can then be added together to produce a plaid pattern.

Fourier analysis is not limited to these simple, regularly repeating patterns, however. It can be applied to complex images of objects, people, and even whole scenes. Although we cannot demonstrate how to construct such complicated images from individual sinusoidal components—far too many gratings would be required—we *can* demonstrate what kind of spatial information is carried by different ranges of spatial frequencies. Figure 4.2.4 shows a picture of Groucho Marx together with two different versions of it that contain only low and high spatial frequencies, respectively. You can see that low spatial frequencies in the middle picture carry the coarse spatial structure of the image (that is, the large black and white areas), whereas the high spatial frequencies in the right picture carry the fine spatial structure (that is, the sharp edges and small details).

The Fourier analysis of an image consists of two parts: the power spectrum and the phase spectrum. The **power spectrum** specifies the amplitude of each

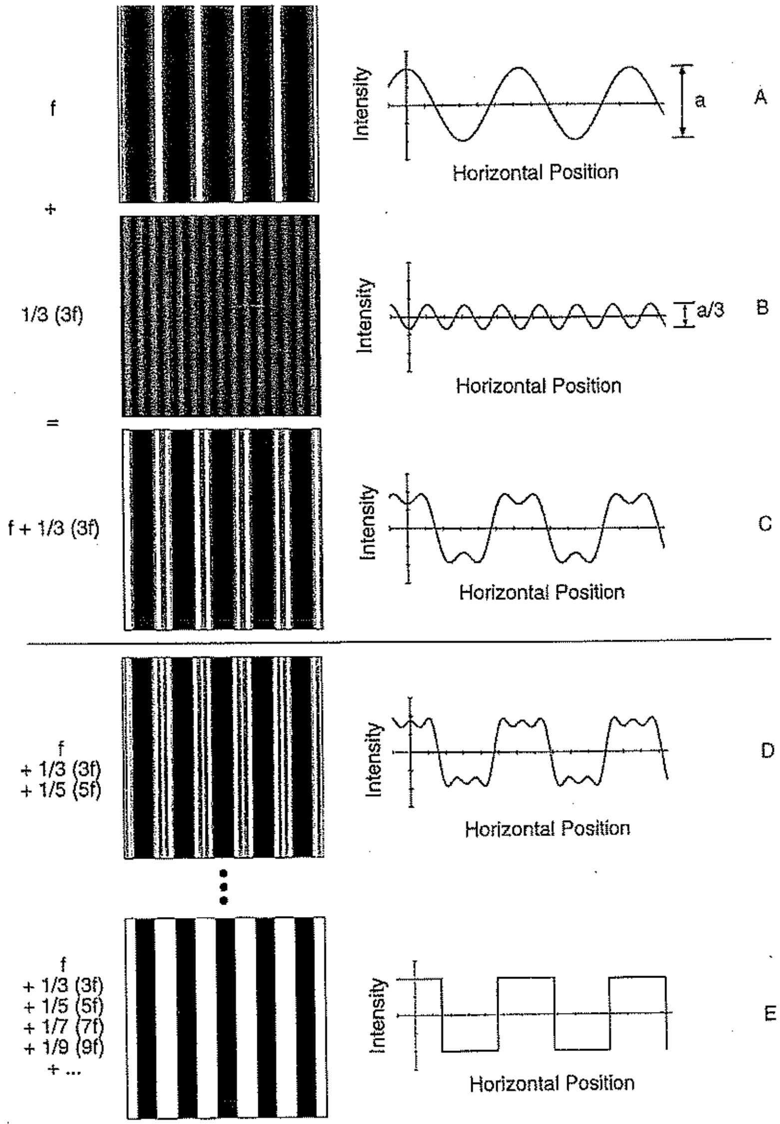


Figure 4.2.2 Constructing a square wave by adding sinusoidal components. (A) A grating at the fundamental frequency (f) of the square wave together with its luminance profile. (B) A grating at the third harmonic ($3f$) with one-third the amplitude. Adding these two gratings results in the grating and luminance profile in part C. Adding the fifth harmonic ($5f$) at one-fifth the amplitude gives the result shown in part D. Adding all the odd harmonics in the proper amplitudes and phases gives the square wave shown in part E.

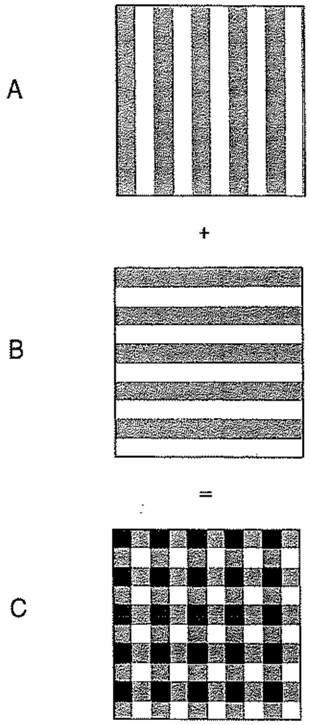


Figure 4.2.3 Constructing a plaid grating by adding square wave gratings at different orientations. The plaid grating at the bottom is formed by adding square wave gratings at vertical and horizontal orientations as shown.

constituent grating at a particular spatial frequency and orientation, whereas the **phase spectrum** specifies the phase of each grating at a particular spatial frequency and orientation. If all of these gratings at the proper phases and amplitudes were added up, they would exactly recreate the original image. Thus, Fourier analysis provides a very general method of decomposing complex images into primitive components, since it has been proven to work for any image. Fourier analysis is also capable of being "inverted" through a process called **Fourier synthesis** so that the original image can be reconstructed from its power and phase spectra. The invertibility of Fourier analysis shows that these spectra contain all the information in the original image.

It is unclear whether the same claims are true for analyzing an image into sets of line and edge primitives or for resynthesizing an image from them, however, because no general theorems like Fourier's have ever been proven that use lines and edges as spatial primitives. Be that as it may, mathematical power and elegance alone are not convincing arguments that the visual system does anything like a Fourier analysis. Empirical evidence must be brought to bear—and so it has been. We will now examine some of these findings.

Spatial Frequency Channels. The spatial frequency theory of image-based vision proposes that early visual processing can be understood in terms of a large number of overlapping **psychophysical channels** at different spatial frequencies and orientations. The concept of a psychophysical channel will require some explaining because it is a fairly technical construct. It can be understood intuitively, however, by analogy with channels with which you are already familiar: the channels in your TV set.

The signals from all TV stations are simultaneously present in the air around us. They are broadcast in the form of electromagnetic energy, which, as you may remember from Chapter 3, lies within a band of wavelengths far outside the visible region of the spectrum (see Color Plate 3.1). Different TV stations broadcast their signals in different subranges so that their signals do not interfere with each other. An important part of what your TV set does is to select the proper subrange of wavelengths from all the others for the particular channel you have tuned. When you tune your TV to channel 4, for example, you are actually tuning it to select just the range of wavelengths on which your local "channel 4" station is broadcasting. Thus, you can think of your TV's channel selector as controlling internal mechanisms that allow it to receive signals selectively from only a small subrange of wavelengths.

The concept of a psychophysical channel is a hypothetical mechanism in the visual system—whose actual physiological substrate is not specified—that is selectively tuned to a limited range of values within some continuum. In the domain of color vision, for example, Helmholtz proposed that there were three channels defined by the wavelength sensitivity curves of three classes of hypothetical elements in the retina (see Figure 3.2.7). In the case of the spatial frequency theory of vision, each channel is defined by the spatial frequency and orientation of the gratings to which it is maximally sensitive.

The spatial frequency approach to image processing asserts that the visual system can be understood as consisting of many overlapping channels that are selectively

Figure 4.2.4 Spatial frequency content of a complex image. The picture of Groucho Marx on the left has been analyzed into its low-frequency information (middle) and high-frequency information (right). Low frequencies carry the global pattern of light and dark; high frequencies carry the local contrast information at the edges of objects. (From Frisby, 1979.)

tuned to different ranges of spatial frequencies and orientations. There is now a great deal of evidence to support this view. One of the landmark papers that launched the spatial frequency theory of vision was published in 1969 by British psychophysicists Colin Blakemore and Fergus Campbell. In it, they reported the results of an experiment that provided striking evidence for the existence of spatial frequency channels in vision. The point of the experiment was to show that when people viewed a sinusoidal grating for a long time, their visual systems adapt selectively to gratings at the presented orientation and frequency but not others, as measured by psychophysical techniques. Because of its historical and conceptual importance to the spatial frequency theory of vision, we will now examine this experiment in some detail.

Contrast Sensitivity Functions. The basic idea behind Blakemore and Campbell's (1969) experiment was to determine the effects of adapting an observer to a particular spatial frequency grating by measuring their sensitivity to such gratings both before and after adaptation. The standard measurement of how sensitive observers are to gratings at different frequencies is called the **contrast sensitivity function (CSF)**. It is determined by finding the lowest contrast at which the observer can just barely detect the difference between a sinusoidal grating and a uniform gray field, that is, the threshold at which a very low-contrast grating stops looking like a uniform gray field and starts to look striped. This threshold is measured for gratings at many different spatial frequencies from low (wide fuzzy stripes) to high (narrow fuzzy stripes).

The fastest and easiest procedure for measuring contrast thresholds is using the method of adjustment. Each subject adjusts a knob that controls the contrast of the grating at a particular spatial frequency on a TV monitor to the point at which he or she can just barely detect its striped appearance. (Other, more complex methods are also available, as explained in Appendix A.) This adjustment procedure is repeated for many gratings at different spatial frequencies. The results of such an experiment can be summarized in a graph in which the contrast at threshold is plotted as a function of spatial frequency, as shown in Figure 4.2.5A. The reciprocal of this graph—made by flipping it upside down—defines the contrast sensitivity function over the spatial frequency continuum, since threshold is high when sensitivity is low and vice versa. The CSF produced by this procedure typically looks like the one shown in Figure 4.2.5B.

You can observe the overall shape of your own CSF by looking at Figure 4.2.6. It shows sinusoidal stripes of increasing spatial frequency along the horizontal axis and of decreasing contrast along the vertical axis. At threshold contrast, your ability to detect the grating disappears, and so the height at which you no longer see the stripes but just a gray background indicates your sensitivity to gratings at the given spatial frequency. If you hold the book about 30 inches from your eyes, the

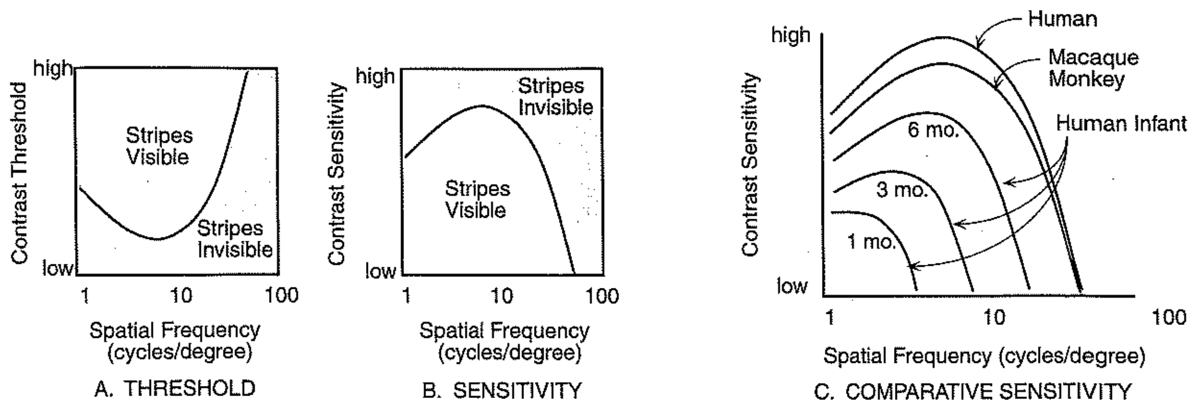


Figure 4.2.5 Contrast sensitivity functions. (A) Minimum contrast at threshold plotted as a function of spatial frequency. (B) Contrast sensitivity plotted as a function of spatial frequency, the inverse of graph A. (C) Contrast sensitivity functions for adult humans, macaque monkeys, and infants at several ages.

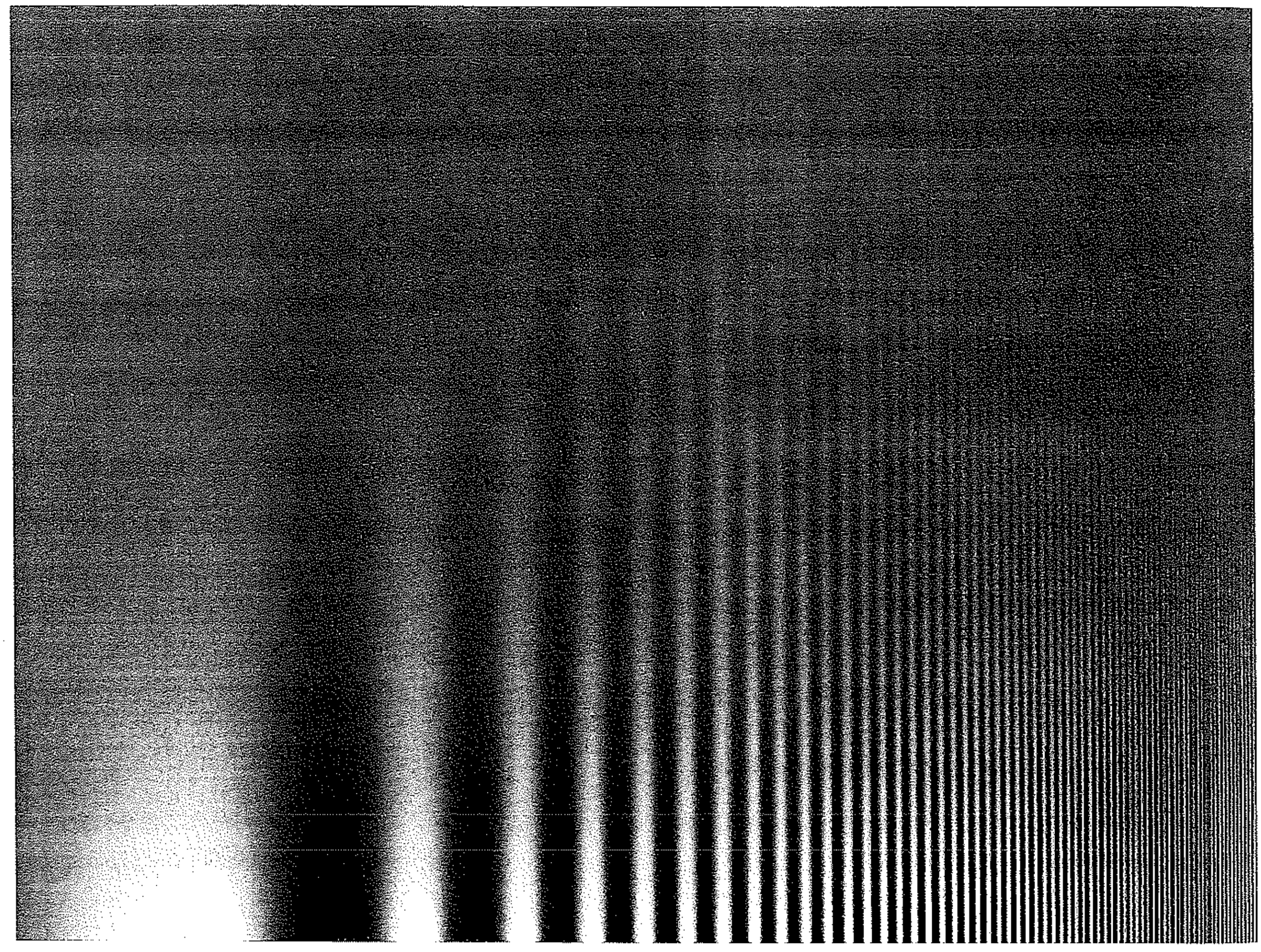


Figure 4.2.6 Demonstration of the shape of the contrast sensitivity function for luminance gratings. Spatial frequency increases continuously from left to right, and contrast increases from top to bottom. The envelope of the striped region should approximate the curve shown in Figure 4.2.5B.

outline of the striped portion of Figure 4.2.6 should look very much like the CSF plotted in Figure 4.2.5B.

The CSF shows that people are most sensitive to intermediate spatial frequencies at about 4–5 cycles per degree of visual angle. Some other CSFs are shown in Figure 4.2.5C for comparison. Notice that babies are much less sensitive at birth, especially at high frequencies (Atkinson, Braddick, & Moar, 1977). As for other species, the macaque monkey's CSF is almost identical to that of humans, a fact that makes macaques an almost ideal animal model for studying the physiology of human spatial vision. If the CSF is measured under low-light (scotopic) conditions in humans, sensitivity to all frequencies drops dramatically, especially at the highest frequencies. This means that at night, when just the rods are operating, human vision lacks the high acuity that it has in daylight. This is primarily because there are no rods in the fovea, the area of greatest visual acuity under photopic (high light) conditions.

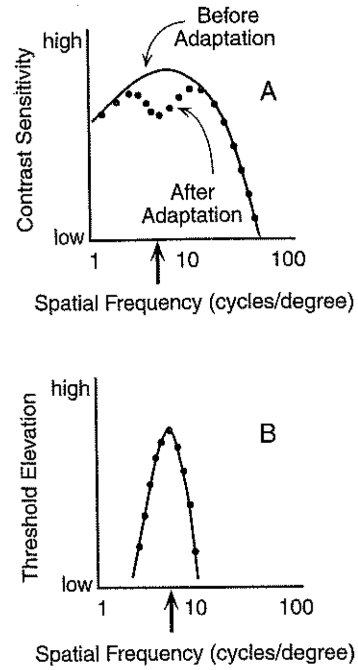


Figure 4.2.7 Contrast sensitivity before and after adaptation to a sinusoidal grating. The solid curve in the top graph shows the normal contrast sensitivity function before adaptation. The data points show it after adaptation to a grating at about 8 cycles per degree (see arrow). The lower graph shows the difference between these two measurements (data points) and a smooth-fitting curve that approximates the reduction in sensitivity (or elevation in threshold).

Selective Adaptation of Channels. Now let us return to Blakemore and Campbell's experiment. After measuring each subject's CSF, they had the subject adapt to a grating of a particular spatial frequency by having him or her scan back and forth over it for a few minutes. Then they remeasured thresholds at each spatial frequency. The extended exposure to the grating caused the subject's visual system to *adapt*, that is, to become less sensitive after the prolonged viewing experience (see Section 1.1.3), but only near the particular spatial frequency and orientation of the adapting grating. The postadaptation CSF, shown as the dotted function in Figure 4.2.7A, indicates just how selective the change in sensitivity is for the spatial frequency of the adapting grating. Test gratings with much lower or higher frequencies were not affected at all by adapting to the grating. The extent of the adaptation can be measured by plotting the difference between the original CSF and the adapted CSF, as shown in the graph in Figure 4.2.7B. This property of selective adaptation is one of the signatures of a psychophysical channel: Each channel adapts to a degree that reflects how sensitive it is to the adapting stimulus. Adaptation therefore results in lowered sensitivity for just a small portion of the spatial frequency continuum rather than equally throughout the whole range.

The results of this experiment can be explained rather simply by a theory based on spatial frequency channels. The theory states that the broad-band CSF that was originally measured actually represents the combined contribution of many overlapping narrow-band channels, each of which is sensitive to a different range of spatial frequencies, as illustrated in the upper graph in Figure 4.2.8. When the adapting grating is presented for an extended period, the channels that are sensitive to that spatial frequency *fatigue*, that is, they "get tired" and respond less vigorously. This fatigue is represented in the lower graph in Figure 4.2.8 as lowered sensitivity in channels near the frequency of the adapting grating. The overall CSF after adaptation thus has a "notch" or "dip" around the adapting grating because, after adaptation, the specific channels that are responsible for perception of the gratings in this frequency range are less sensitive to the same or similar stimuli. It is worth noting that this adaptation effect cannot be

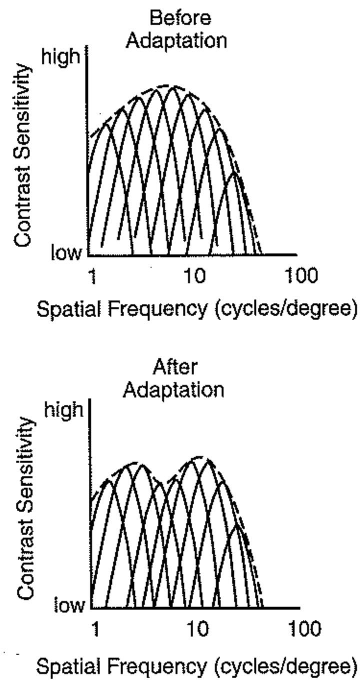


Figure 4.2.8 The multiple spatial frequency channels hypothesis. The contrast sensitivity function (dashed curve in the upper graph) is hypothesized to be the overall envelope of many overlapping spatial frequency channels (solid curves). The dip in contrast sensitivity following adaptation (dashed curve in the lower graph) is hypothesized to be due to selective adaptation by channels near the adapting frequency.

attributed to simple afterimages because observers moved their eyes back and forth over the grating during adaptation, thus thoroughly smearing any afterimage.

Selective adaptation has similar effects on the orientation of gratings (Blakemore & Campbell, 1969; Blakemore & Nachmias, 1971). We demonstrated the aftereffect of such adaptation effects in Chapter 1 (Figure 1.1.3). To measure their effect on thresholds, one must first determine each subject's sensitivity to sinusoidal gratings of a specific spatial frequency at many different orientations. Once this is established, subjects adapt to a single grating at one particular orientation. Sensitivity at each orientation is then redetermined by measuring postadaptation thresholds to gratings of the same spatial frequency but different orientations. The results, shown in Figure 4.2.9 for a relatively low spatial frequency, clearly indicate reduced sensitivity for orientations close to the adapting grating, analogous to the "notch" found in the CSF as a function of spatial frequency.

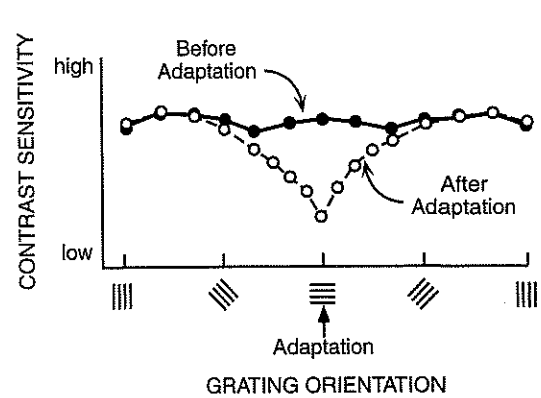


Figure 4.2.9 Selective adaptation to orientation of gratings. Measured sensitivity to sinusoidal gratings is shown as a function of orientation before adaptation (solid circles) and after adapting to a horizontal grating (open circles). (After Bradley, Switkes, & De Valois, 1988.)

Spatial Frequency Aftereffects. Just as gratings of a particular spatial frequency and orientation produce specific adaptation effects, they also produce specific aftereffects. Figure 4.2.10 will allow you to experience the spatial frequency aftereffect for yourself. First, look at the two gratings on the right side and convince yourself that they are identical. Then stare at the gratings on the left for about a minute, moving your eyes back and forth along the horizontal bar in the middle so that the high-frequency grating always stimulates the upper half of your retina and the low-frequency grating always stimulates the lower half. After doing this for a full minute, change your fixation to the bar in the center of the right two gratings. Do they still look the same? If you have adapted for long enough, the upper grating will look as though it has decidedly narrower stripes (that is, higher spatial frequency) than the lower one, even though the two gratings are actually identical. This perception is therefore due to the differential aftereffects of viewing the two adapting gratings.

The standard explanation of this spatial frequency aftereffect is closely related to the one proposed for color afterimages, except that the cells involved are tuned selectively to different spatial frequency bands. Prolonged viewing of the wide grating fatigues the cells that respond selectively to low spatial frequencies in the upper half of the visual field. Similarly, prolonged viewing of the narrow grating fatigues the cells that respond selectively to high spatial frequencies in the lower half of the visual field. The two identical gratings on the right

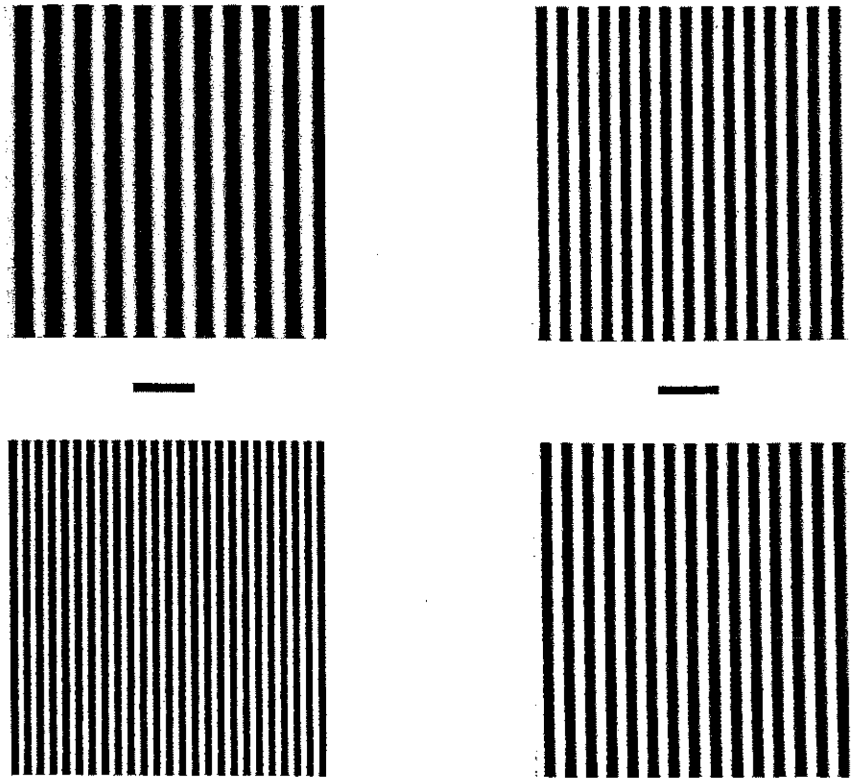


Figure 4.2.10 A demonstration of aftereffects of spatial frequency adaptation. After checking to be sure the two gratings on the right are identical, adapt to the gratings on the left by running your eyes back and forth along the central part for about a minute. Then fixate on the central bar between the gratings on the right and compare their stripe widths. (From De Valois & De Valois, 1988.)

look different because the high-frequency cells are relatively more sensitive on top after adaptation to the low-frequency grating, and the low-frequency cells are relatively more sensitive on bottom after adaptation to the high-frequency grating. These altered sensitivities thus shift the pattern of activity in opposite directions for the upper and lower gratings on the right, producing the experience of narrower stripes in the upper grating and wider stripes in the lower grating.

The reader will recall an analogous demonstration for orientation specific aftereffects—often called *tilt aftereffects*—in Chapter 1 using gratings that differed in orientation (Figure 1.1.3). The explanation in terms of neural fatigue changing the pattern of firing to the identical gratings will also work here, except that the cells in question are ones that are selectively tuned to different orientations rather than to different spatial frequencies.

Thresholds for Sine Wave versus Square Wave Gratings. Further support for the spatial frequency theory of image-based processing has come from many other psychophysical studies of people's detection and discrimination of grating stimuli. The experiments that psychophysicists Norma Graham and Jacob Nachmias (1971) performed to study the difference between detecting gratings made from sine waves versus square waves are particularly elegant. They made several precise, counterintuitive predictions based on spatial frequency theory and found them to be exactly correct.

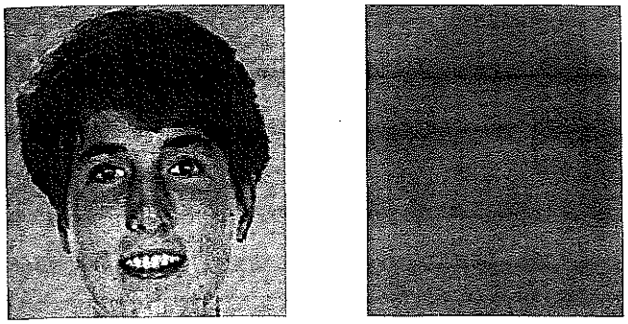
The basis of their predictions was the hypothesis that a square wave grating would be represented in the early visual system not as a unitary stimulus, but as a collection of many sine wave gratings at different spatial frequencies and amplitudes. Specifically, spatial frequency theory asserts that a square wave grating of frequency f with amplitude a is decomposed into a sine wave grating of frequency f with amplitude a, plus another sine wave grating of frequency $3f$ with amplitude $a/3$, plus a third sine wave grating of frequency $5f$ with amplitude $a/5$, and so on (see Figure 4.2.2). The first prediction that Graham and Nachmias tested was that the threshold for detecting the presence of a square wave grating would be exactly the same as that for detecting a sine wave grating with the same spatial frequency as the fundamental (f) of the square wave. The rationale is simply that the threshold for detecting the square wave grating will be crossed whenever any of its sine wave components crosses its own independent threshold. The component at the fundamental frequency will be crossed first because its amplitude is much greater than that of any of the higher harmonics of the square wave ($3f$, $5f$, $7f, \ldots$).

Graham and Nachmias tested this prediction by determining the contrast threshold (i.e., the lowest amplitude) at which a square wave grating can be discriminated from a uniform gray field whose luminance is the same as the average luminance of the grating. (See Appendix A for a description of methods for determining thresholds.) From the subject's point of view, the task was to determine whether the test stimulus has any hint of periodic light-dark striping that distinguishes it from a uniform gray field. Graham and Nachmias found that the threshold for performing this task with a square wave grating was exactly the same as the threshold for a sine wave grating whose spatial frequency was the same as the fundamental frequency of the square wave. This finding is surprising because a square wave grating has a much steeper luminance gradient (i.e., a more rapid change over space from light to dark; see Figure 4.2.2E) than does a sine wave grating (Figure 4.2.2A), and the threshold task can sensibly be considered one of detecting whether the luminance gradient of the test stimulus is greater than zero (the luminance gradient for a uniform field). From this viewpoint, the most obvious prediction is that the square wave grating would have a lower threshold than the sine wave grating because its luminance gradient is steeper. But no such difference was found, just as spatial frequency theory predicted.

Graham and Nachmias also examined the contrast threshold at which subjects could discriminate between a sine wave grating and a square wave grating of the same spatial frequency. In this task, the subject must discriminate between these two different striping patterns rather than just whether or not striping is present. Here again, spatial frequency theory makes a precise, nonintuitive prediction: The contrast threshold for discriminating between a sine wave grating and a square wave grating should be the same as the contrast threshold for discriminating between a uniform field and a sine wave grating whose spatial frequency is the third harmonic ($3f$) of the square wave. This prediction is also based on the hypothesis that the visual system decomposes a square wave grating into a series of sinusoidal components, including its fundamental (f) and all its odd harmonics ($3f$, $5f$, $7f, \ldots$). If so, the difference between a square wave grating and a sine wave grating at its fundamental frequency is only in the presence of the odd harmonics, and for this difference to be detected, one of these harmonics must cross its independent threshold. Because the third harmonic of the square wave grating has the greatest amplitude ($a/3$) of all the odd harmonics, it is the one that should cross its threshold first, and this threshold should be crossed at the same contrast at which the third harmonic can be discriminated from a uniform field.

The results of this experiment again showed remarkably close agreement with the predictions of spatial frequency theory. The fit is all the more remarkable because nobody would have made such predictions from any other existing theory. Many additional experiments have confirmed further tests of spatial frequency theory, making it the dominant psychophysical theory of early spatial vision for the past several decades. It revolutionized the study of visual psychophysics, not only in adult vision, but in infant vision as well.

Development of Spatial Frequency Channels. Psychophysical studies of infant perception have shown that babies see the world quite differently from adults, at least in some respects. These studies are typically conducted by using the preferential looking paradigm discussed in Section 3.2.4 (Fantz, 1958, 1965) with sinusoidal gratings as stimuli. An infant is typically shown a sine wave grating on one side and a homogeneous field on the other side. The two displays are matched for average luminance, so the only difference between them is the degree of modulation into light and dark stripes. If the baby cannot tell the difference, he or she will spend equal amounts of time looking at each. If the grating looks different in any way, the baby will spend more time looking at the grating, presumably be-



A B

Figure 4.2.11 A simulation of adult versus infant perception of a face. The woman's face in part A would look to an infant like the filtered version in part B, in which the high spatial frequencies have been removed. (Courtesy Sheryl Ehrlich.)

cause it is visually more interesting. By varying the contrast between the light and dark stripes and measuring looking times, researchers can measure the infant's contrast sensitivity function (CSF).

The CSFs for infants at several ages are plotted in Figure 4.2.5C (Atkinson et al., 1977). It shows that infants are less sensitive overall to the gratings and that the biggest difference occurs at high spatial frequencies. The perceptual consequences of this fact are illustrated in Figure 4.2.11. Part A shows a picture of a woman's face as it appears to an adult, and part B shows how it (presumably) appears to infants, given the limited range of spatial frequencies to which they are sensitive. The fuzzy contours result from infants' insensitivity to high-frequency information.

4.2.2 Physiology of Spatial Frequency Channels

Psychophysical channels are hypothetical mechanisms inferred from behavioral measures rather than directly observed biological mechanisms of the nervous system. Thus, psychophysical channels are information processing constructs at Marr's algorithmic level of description rather than at his implementational level. If these channels are real, however, they must be implemented somewhere in the visual nervous system. The questions to which we now turn are how and where.

The answers to these questions provide the second theory about the function of the cells that Hubel and Wiesel discovered in striate cortex. There is now substantial evidence that these cells may be performing a **local spatial frequency analysis** of incoming images. The analysis that they perform is "local" because the receptive fields of striate cells are spatially limited to a few degrees of visual angle (or even less in the fovea). This is obviously quite restricted in comparison with the theoretically infinite extent of the sinusoidal gratings on which classic Fourier analysis is based. It is even much more restricted than the large grating stimuli (10° or more) normally used in psychophysical studies. However, a local, piecewise, spatial frequency analysis can be accomplished through many small patches of sinusoidal gratings that "fade out" with distance from the center of the receptive field, as illustrated in Figure 4.2.12. This sort of receptive field structure—called a **Gabor function** (or **wavelet**[2])—is constructed by multiplying a global sinusoidal grating by a bell-shaped Gaussian envelope. The one-dimensional luminance profile of this function is shown in Figure 4.2.12A together with a full two-dimensional display that shows how light intensity varies over space according to a Gabor function.

As described in Section 4.1.2, Russell De Valois, Karen De Valois, and their colleagues have mapped the receptive fields of V1 cells carefully and have found evidence for multiple lobes of excitation and inhibition (see Figure 4.1.8). Such receptive fields clearly look a great deal like profiles of Gabor functions (Figure 4.2.12A). To strengthen the connection between these cells and local spatial frequency theory, De Valois, Albrecht and Thorell (1982) measured the spatial frequency tuning of both simple and complex cells. They found many to be quite sharply tuned to small frequency ranges, as would be expected if they were the biological

[2] Formal distinctions can be drawn between Gabor and wavelet functions (see Field, 1994). If Gabor functions are taken to be all functions derived from a Gaussian modulated sinusoid, then wavelet functions are a particular kind of Gabor function in which the variance of the Gaussian is a constant number of cycles of the sinusoid. That is, wavelets are Gabor functions that are "self-similar" in the sense that they differ only by dilations, translations, and rotations of a single underlying function. We will not make this distinction in the text of this book, however, and refer generically to wavelets as Gabor functions.

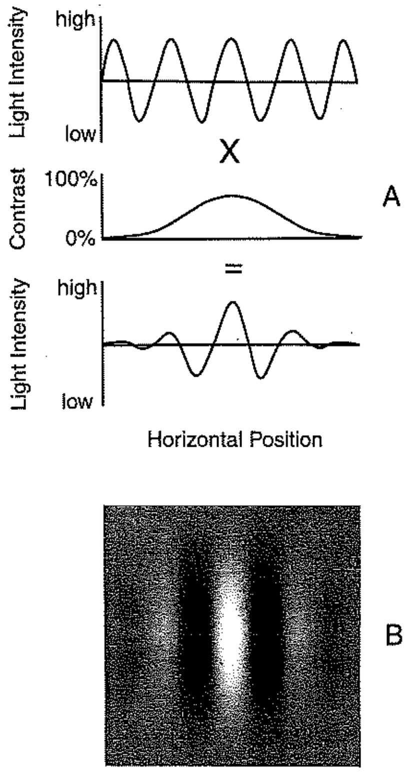


Figure 4.2.12 Gabor functions. A Gabor function is constructed by multiplying a sinusoidal function by a Gaussian function as indicated in part A. The resulting luminance pattern is shown in part B.

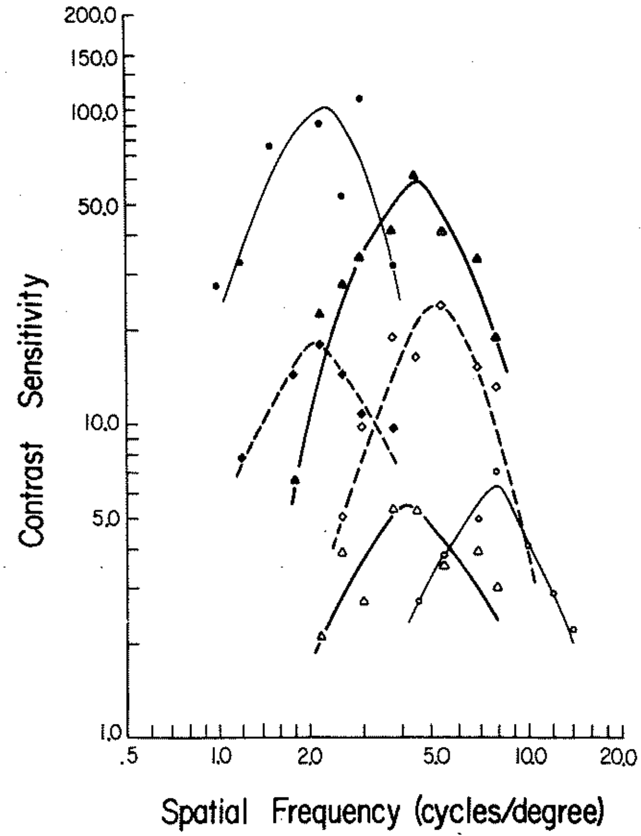


Figure 4.2.13 Contrast sensitivity functions for six cells in macaque striate cortex. Each cell shows fairly sharp tuning in spatial frequency, much as is predicted by the multiple spatial frequency channels hypothesis illustrated in Figure 4.2.8. (From De Valois, Albrecht, & Thorell, 1982.)

implementations of local spatial frequency channels in the brain. Figure 4.2.13 shows a sampling of the frequency tuning characteristics of cells in macaque monkey cortex.

The degree of tuning in cortical cells seems to fall along a continuum; some are very sharply tuned and others quite broadly tuned (De Valois et al., 1982). In general, cells that are tuned to high spatial frequencies have narrower tuning than do cells that are tuned to low spatial frequencies. Simple cells also tend to be more narrowly tuned than complex cells, although the difference is not large. There is a similar continuum in the degree of orientation tuning; some cells respond only to gratings that are very close to their "favorite" orientation, whereas others respond almost equally to gratings in any orientation. As it turns out, the frequency and orientation tuning characteristics of cortical cells are correlated: Cells that are broadly tuned for spatial frequency are also broadly tuned for orientation, and cells that are narrowly tuned for spatial frequency are also narrowly tuned for orientation (De Valois & De Valois, 1988).

Further physiological studies have shown that the cortical layout of cells that are tuned to different spatial frequencies is quite systematic. In particular, they appear to be ordered along a dimension perpendicular to that of orientation selectivity within each hypercolumn (De Valois & De Valois 1988). Spatial frequency and orientation thus define a literal two-dimensional space within hypercolumns. In cats, De Valois and De Valois propose that the spatial frequency dimension is laid out in a Cartesian coordinate system as indicated in Figure 4.2.14A. In monkeys, they believe that the architecture is slightly different, orientation and spatial frequency being arranged as dimensions in polar coordinate space, as illustrated in Figure 4.2.14B. Orientation is represented by direction from the center of a hypercolumn; spatial frequency is represented by distance from the center.

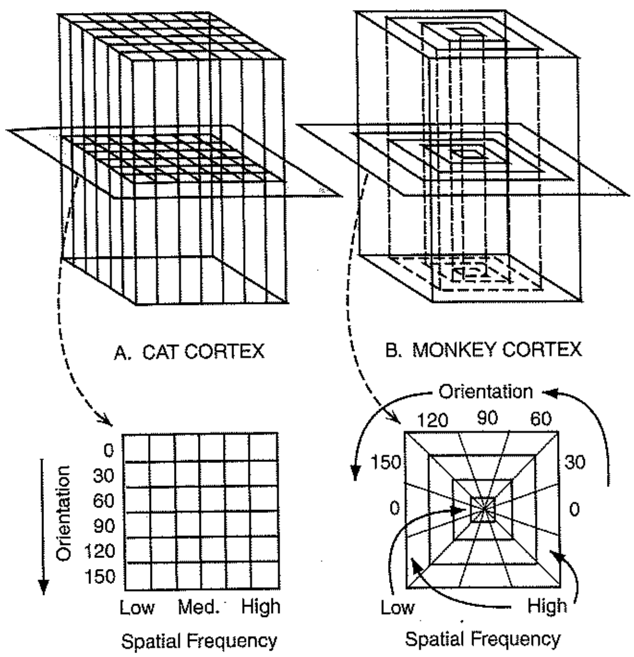


Figure 4.2.14 Models of cortical hypercolumn architecture in cats and monkeys including a spatial frequency dimension. In cats, the spatial frequency dimension is hypothesized to be orthogonal to the orientation dimension in a Cartesian structure, whereas in monkeys it is hypothesized to be radially organized in a polar structure. (After De Valois & De Valois, 1988.)

Although the evidence that simple and complex cells in area V1 may be doing a local spatial frequency analysis of input images is impressive, this conclusion is not universally held. Nevertheless, local spatial frequency theory has led to several interesting and important discoveries about the properties of V1 cells and so must be counted a very serious alternative to the line and edge detector theory suggested by Hubel and Wiesel.

It is interesting to examine the relation between these two theories. They compete because their functional implications are quite different. Spatial frequency theory suggests that these cells are not "detectors" of naturalistic image features, such as lines and edges, but are general purpose analyzers (often called **filters**) that decompose the image into a useful set of primitives that can describe any possible image succinctly. This view does not preclude the existence of line and edge detector cells in the visual system, however. Rather, it simply locates them at a higher level, where the appropriate Gabor filters could be combined to specify the line or edge. Thus, the local spatial frequency theory is potentially compatible with line and edge detector theory but not with the further claim that these detectors are implemented in the cells of area V1. We will return to this controversy later in the chapter, after we have considered the further insights provided by computational approaches to image processing.

4.3 Computational Approaches

Computational theorists have investigated the nature of image processing from a number of different perspectives. The majority have attacked the problem in terms of effective techniques for detecting naturalistic image features such as edges and lines in gray-scale images. Many of the best known practitioners of this so-called traditional approach have worked in the vision group at M.I.T., including David Marr, Tomasso Poggio, Ellen Hildreth, Shimon Ullman, and their colleagues. Much of their research has been aimed at producing a computer implementation of Marr's raw primal sketch that we mentioned briefly at the end of Chapter 2. This group has produced some important results concerning the computational and algorithmic descriptions (in Marr's sense) of edge and line detection problems. This work is closely related to Hubel and Wiesel's conjecture that striate cortical cells are detecting edges and lines.

Despite the relative dominance of this approach, alternative computational views have arisen. One advocates taking a filtering approach to vision, which is based largely on the spatial frequency theory of early vision described in the previous section. Filtering theorists such as Adelson and Bergen (1985), Heeger (1988), Koenderink and Van Doorn (1976a), and Malik (Jones & Malik, 1992; Malik & Perona, 1990) are exploring the computational advantages of using a set of multiorientation, multiscale filters (such as the Gabor functions mentioned in the previous section) as the spatial primitives on which higher level processes operate.

Yet another group of computational theorists is the emerging camp of connectionists who are taking a very different approach to the problem of determining how image processing might work. They are using powerful computational learning techniques (e.g., back propagation, described in Appendix B) that enable neural networks to "program themselves" to perform a well-