# Elements of Early Vision for Computer Graphics

**James A. Ferwerda**
*Program of Computer Graphics, Cornell University*

**O**ver the past decade, visually based techniques in computer graphics have blossomed. Important advances in perceptually driven rendering, realistic image display, high-fidelity visualization, and appearance-preserving geometric simplification have all been realized by applying knowledge of the limitations and capabilities of human visual processing. Much of this work is grounded in the physiology and psychophysics of early vision, which focuses on how visual mechanisms transduce and code the patterns of light arriving at the eye. This tutorial surveys some of the fundamental findings in the study of early vision including basic visual anatomy and physiology, optical properties of the eye, light sensitivity and visual adaptation, and spatial vision.

**This tutorial surveys some of the key features of early visual processing that have informed the development of perceptually based graphics algorithms.**
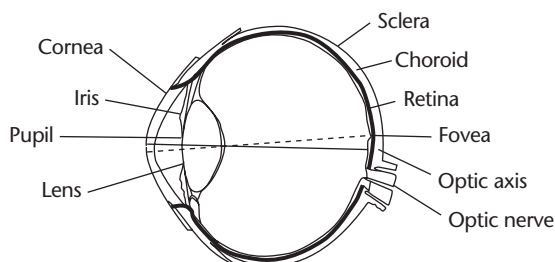
## Visual anatomy and physiology

Understanding human vision begins with the study of basic visual anatomy and physiology. It's important to study the "hardware" of the visual system because this can give insights into the kinds of information that can be coded by visual mechanisms.

### The eye

The visual system begins at the eye. Figure 1 shows a cross section through a schematic human eyeball. The *anterior* section of the eyeball contains the eye's optical system whose major structures are the *cornea*, *lens*, and *iris*. The cornea provides about two-thirds of the eye's refractive power, but the lens provides fine focal control for targets at distances from 20 feet down to about 4 inches.[1] The iris sits in front of the lens and has a central aperture known as the *pupil* that admits light to the eye's central cavity. The space between the cornea and lens is filled with a fluid known as the *aqueous humor*. The central cavity of the eyeball is filled with a gelatinous fluid known as the *vitreous*.

The eye's *posterior* section has three layers. The *sclera* is a tough outer covering that protects the interior from damage and helps maintain the eye's roughly spherical shape. The *choroid* is a middle layer that provides the blood supply to the eye's cellular structures. The *retina* is the interior layer that contains photoreceptor cells and their associated neural tissues.

### The retina

The retina is composed of two major classes of photoreceptor cells known as the *rods* and *cones* because of the shapes of their outer segments. Each retina has somewhere between 100 to 120 million rods and 7 to 8 million cones. The rods are extremely sensitive to light and provide achromatic vision at low (*scotopic*) illumination levels. The cones are less sensitive than the rods but provide color vision at high (*photopic*) levels. The photosensitive segments of the rods and cones lie closest to the choroid layer. This means that light striking the retina must first pass through several layers of neural tissue before reaching the photoreceptors. Only in a small 1.5-mm diameter area near the optic axis called the *fovea* are the photoreceptive surfaces directly exposed to light.

The rod and cone systems are sensitive to light wavelengths from approximately 400 to 700 nanometers (nm). The rods have their peak sensitivity at about 498 nm. Three types of cones have bandpass spectral response characteristics. The short wavelength or "blue" cones have their peak response at 420 nm, the medium wavelength or "green" cones peak at 534 nm, and the long wavelength or "red" cones peak at 564 nm. Significant



**1** Structures of the human eye. (Adapted from Atkinson.[2])

Cornea
Iris
Pupil
Lens
Sclera
Choroid
Retina
Fovea
Optic axis
Optic nerve

overlap exists between the response ranges of the different classes of cones, which means that spectrally broadband stimuli will simultaneously activate multiple cone types.

The rods and cones aren't distributed equally over the retinal surface. The fovea has the densest packing of cones but is nearly devoid of rods. Cone density falls off in a nearly exponential manner with increasing eccentricity and asymptotes to a constant low level at about 20 degrees into the retinal periphery. In contrast, rod density increases from near zero in the fovea to a maximum at an eccentricity of 20 degrees. Rod density drops further into the periphery and both rods and cones reach their minimum density levels at 75 to 80 degrees away from the fovea.

### Retinal receptive fields

The rods and cones synapse on a network of neurons in the retina's outer and inner *plexiform layers*. Figure 2 shows a schematic cross section through the plexiform layers of a rhesus monkey. The cells in the plexiform layers connect groups of rods and cones to *ganglion cells* whose neural fibers form the *optic nerve*. The spatially localized group of photoreceptors that serve a particular ganglion cell is called the cell's *receptive field*.

The receptive fields of ganglion cells are the basic units of visual coding. Electrophysiological studies of cats have shown that many receptive fields have an *antagonistic center/surround organization*.[3] The activation produced by stimulation in the center of a receptive field tends to be suppressed by stimulation in the annular surround. Uniform stimulation over the whole receptive field typically produces only a weak response.

Researchers have identified two classes of ganglion cell receptive fields. *On-center cells* increase their firing rate in response to increments of light in the centers of their fields, and *off-center cells* increase their firing rate in response to light decrements. The antagonistic organization of receptive fields means that early on in the visual system, information about the absolute intensity of light is mostly lost and primarily contrast is signaled to later stages of visual processing. This has significant implications for theories of surface lightness and illumination perception.

Ganglion cells can also be classified by the pattern and duration of their responses to changes in light in their fields.[4] *X cells* show a sustained response to increments or decrements in the centers of their fields. *Y cells* show a brief transient change in response and then return to their base firing rate.

Approximately half of all retinal ganglion cells have receptive fields that show spectral as well as spatial opponency.[5] The red–green opponent cells take their primary input from long and medium wavelength cones. Yellow–blue opponent cells take their input from all three cone types, with opposition between the sum



1. Cone receptor
2. Rod receptor

3. Flat bipolar cell
4. Midget bipolar cell
5. Rod bipolar cell
6. Amacrine cell
7. Horizontal cell

8. Midget ganglion cell
9. Diffuse ganglion cell
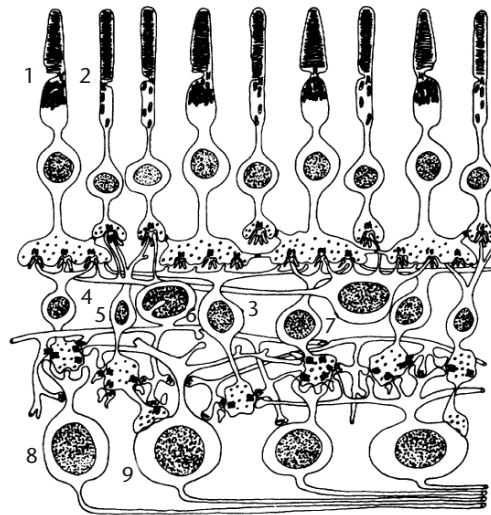
**2** Cross section of the primate retina. (Adapted from Atkinson.[2] Reprinted by permission of John Wiley and Sons, Inc.)



Retina

Optic nerve

Optic chiasm

Lateral geniculate nucleus

Visual cortex
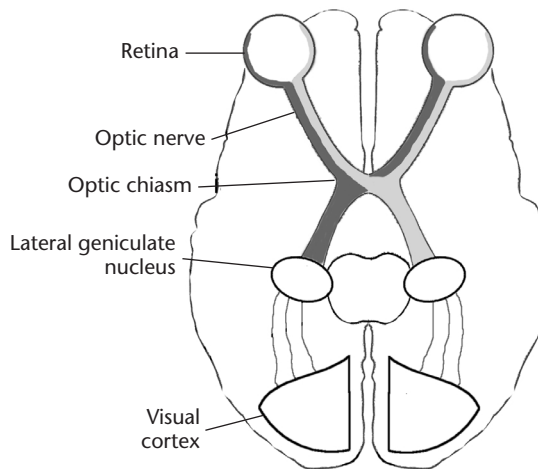
**3** Pathways in the visual system. (Adapted from Sekuler and Blake.[7])

of the long and medium wavelength cones and the short wavelength cones. The discovery of cells with spectrally opponent properties has been used to support physiologically based theories of color perception.[6]

### Visual pathways

Figure 3 shows the major neural pathways in the visual system. The long axons of the retinal ganglion cells form the optic nerve, which contains about one million fibers (of which 100,000 serve the fovea). The optic nerve bundle exits the eyeball at approximately 17 degrees to the nasal side of the optic axis. There are no photoreceptors in this area commonly known as the *blind spot*.

The fibers of the optic nerve project to the *optic chiasm*. At this junction, fibers from the nasal portions of each retina cross over to the opposite side of the head. These crossing fibers join with fibers from the temporal portions of the opposite retina and project to the *lateral geniculate nuclei* (LGN) in each hemisphere.

The six layers of the LGN receive specialized input from the optic nerve fibers of each eye. Two *magnocellular* layers take primary input from the peripheral retina where nonspectrally opponent ganglion cells with
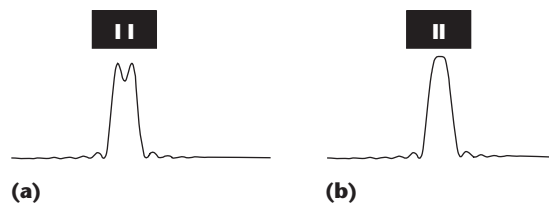
**4** The limits of visual resolution. The retinal intensity profiles of (a) resolvable and (b) unresolvable line targets.

large receptive fields and transient temporal characteristics are dominant. The remaining *parvocellular* layers take primary input from the foveal region where spectrally opponent cells with small receptive fields and sustained temporal characteristics are dominant. The striking differences in the functional properties of the magno- and parvo-cellular layers suggest that the eyes may in fact be serving two visual processing systems. One is a fast-responding, achromatic system, sensitive to motion, but with low spatial resolution. The other is a slow-responding, trichromatic system, relatively insensitive to motion but with high resolution.[8]

From the LGN, fibers project to the *visual cortex*. The primary visual cortex is known alternately as V1, area 17, and striate cortex. Cells in the visual cortex have distinct sensitivities. Some cells are sensitive to a target's color or contrast but not to its shape or motion. Others are selective for a target's orientation but are insensitive to its color and motion. In addition, other cells are selective for orientation and direction of motion but not color. The functional specificity observed in V1 and other areas of the visual cortex has led to speculation that the visual system is divided into "what" (identification) and "where" (localization) systems.[9] Case studies that show that brain damage can produce losses in one type of function without affecting the other support this conjecture.

## The eye as an optical system

The cornea, iris, and lens comprise an optical system that forms an image on the retinal surface. As with any optical system, aberrations in the components and diffraction produced by the entry aperture limit the image's *resolution*. Here resolution means the fidelity with which object features are represented in an image. Features smaller than the resolution limit aren't discernable. Measurements show that the resolving power of the eye's optical system is limited to about 30 seconds of visual angle.[10]

The image formed by the eye's optics falls on the retinal photoreceptors. The photoreceptors are arrayed in a rough hexagonal grid with highest density in the fovea.[11] The photoreceptors sample the retinal image to produce a neural image representation. In terms of sampling theory, the spacing of photoreceptors in the fovea is matched well to the eye's optics. The lowpass filtering provided by the optics lets the photoreceptors create a faithful representation of the continuous retinal image at the sampling intervals given by the spacing of cells in the retinal mosaic.[12]

Optical filtering, receptor sampling, and the receptive field organization of early visual processing determine the resolution with which the visual system represents the patterns of light arriving at the eye. The psychophysical measure of this resolution is known as *visual acuity*.

### Visual acuity

From a bright, thin line in the visual field, the eye's optics will produce a retinal image that has a slightly blurred intensity profile. If two bright lines lie side by side, their retinal intensity profiles will overlap, producing a composite distribution with a central minimum like the one shown in Figure 4a. As the two lines are brought closer together the central minimum's intensity will increase (see Figure 4b). The smallest distance at which the two lines can be visually discriminated is a measure of the visual system's resolving power and the observer's acuity. Figure 4 shows that visual acuity is a function of *contrast sensitivity*. The acuity limit is determined by the visual system's ability to detect the small contrast in the center of the composite distribution. Contrast sensitivity limits this kind of visual acuity to approximately 30 seconds of visual angle.[10]
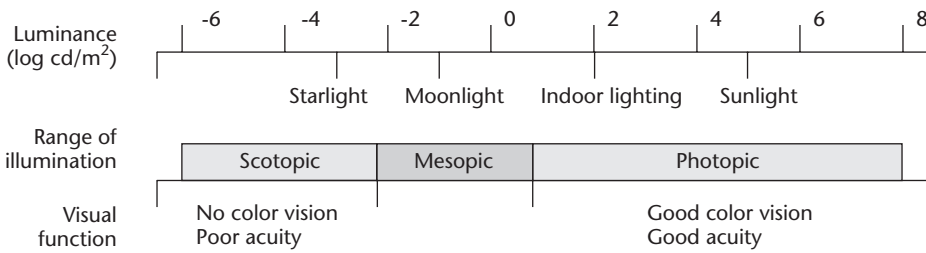
There's another important measure of visual acuity that isn't a function of resolution, but instead specifies the visual system's ability to localize the positions of objects in the visual field. This is known as *vernier acuity* or *hyperacuity*. If two bright lines are laid end to end, observers can detect misalignments of the lines as small as 4 to 6 seconds of visual angle.[13] This precision is remarkable because it corresponds to approximately one fifth of the distance between the foveal photoreceptors. There's still much speculation on how the visual system produces such a fine-grained representation of position.[14] Hyperacuity plays an important role in the visibility of aliasing artifacts in digital images.

## Light sensitivity and visual adaptation

The range of light energy we experience in the course of a day is vast. The light of the noonday sun can be as much as 100 million times more intense than starlight. Figure 5 shows the range of luminances we encounter in the natural environment and summarizes some visual parameters associated with this luminance range. Our visual system copes with this huge luminance range by adapting to the prevailing conditions of illumination. Through *adaptation* the visual system functions over a range of nearly 10 log units.

Adaptation is achieved through the coordinated action of mechanical, photochemical, and neural processes in the visual system. The pupil, the rod and cone systems, bleaching and regeneration of receptor photopigments, and changes in neural processing all play a role in visual adaptation.

Although adaptation provides visual function over a wide range of illumination levels, this doesn't mean that we see equally well at all levels. For example, under dim illumination our eyes are very sensitive, and we're able to detect small differences in luminance. However, our acuity for pattern details and our ability to distinguish colors are both poor. This is why it's difficult to read a newspaper at twilight or to correctly choose a pair of colored socks while dressing at dawn. Conversely, in

Luminance
(log cd/m²)

| -6 | -4 | -2 | 0 | 2 | 4 | 6 | 8 |

Starlight    Moonlight    Indoor lighting    Sunlight

Range of
illumination

| Scotopic | Mesopic | Photopic |

Visual
function

No color vision                 Good color vision
Poor acuity                     Good acuity

**5** **The range of luminances in the environment and associated visual parameters. (Adapted from Spillman and Werner.[15])**

daylight we have sharp color vision, but absolute sensitivity is low and luminance differences must be large to be detectable. This is why it's impossible to see the stars against the sunlit sky.

Further, adaptation doesn't happen instantaneously. Nearly everyone has experienced the temporary blindness that occurs when entering a dark theater for a matinee. It can sometimes take a few minutes before you can see well enough to find an empty seat. Similarly, once you've adapted to the dark theater, going back out into the daylight is at first dazzling, but within about a minute, you can see normally again.

### Changes in sensitivity

Visual adaptation is often measured psychophysically in a *detection threshold* experiment. Typically, subjects are seated in front of a blank screen that fills a large portion of their field of view. To determine the absolute threshold, the screen is made dark. To determine the contrast threshold, a large region of the screen is illuminated to a particular background level. Before testing begins, the subjects fixate on the screen until they have completely adapted to the background level. On each trial a disk of light is flashed near the center of fixation for a few hundred milliseconds. The subjects report whether they see the disk. If they don't see the disk, its intensity is increased on the next trial. If they do see the disk, its intensity is decreased. In this way, the detection thresholds for seeing the target disk against different backgrounds can be measured.

As the background luminance in a detection threshold experiment is increased from zero, the luminance difference between target and background required for detection increases in proportion to the background luminance. Plotting the detection threshold against the corresponding background luminance gives a *threshold-versus-intensity* (TVI) function.

Figure 6 shows TVI functions for the rod and cone systems. At luminance levels below about −4 log candelas per square meter (cd/m²), the rod curve flattens to a horizontal asymptote. This indicates that the background luminance has little effect on the threshold, which approaches the visual system's absolute sensitivity limit. At levels above 2 log cd/m², the curve approaches a vertical asymptote. This indicates that the rod system is being overloaded by the background with the result that no amount of luminance difference between them is detectable.

The function is linear over a wide middle range covering 3.5 log units of background luminance. We can describe this relationship, known as Weber's law,[16] by the function $\Delta L = kL$, where $L$ is luminance and $k$ is an
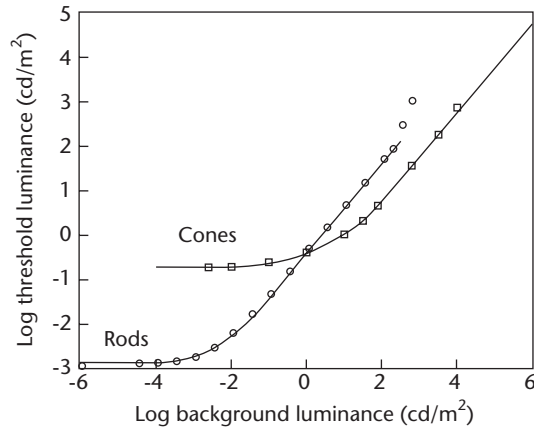


**6** **Threshold versus intensity (TVI) functions for the rod and cone systems.**
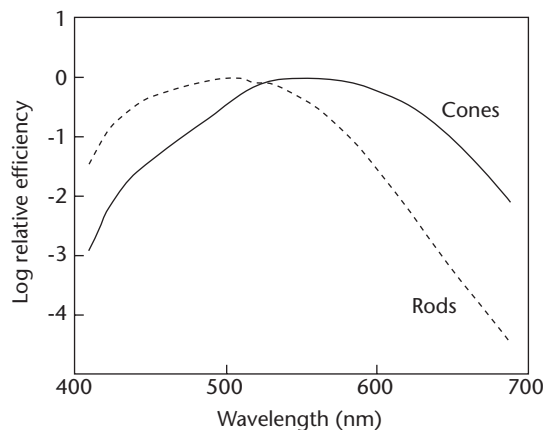


**7** **Photopic (cones) and scotopic (rods) luminous efficiency functions.**

experimentally defined constant. Weber's law behavior is indicative of a system that has constant contrast sensitivity, since the increase in threshold with background luminance corresponds to a luminance pattern with constant contrast.
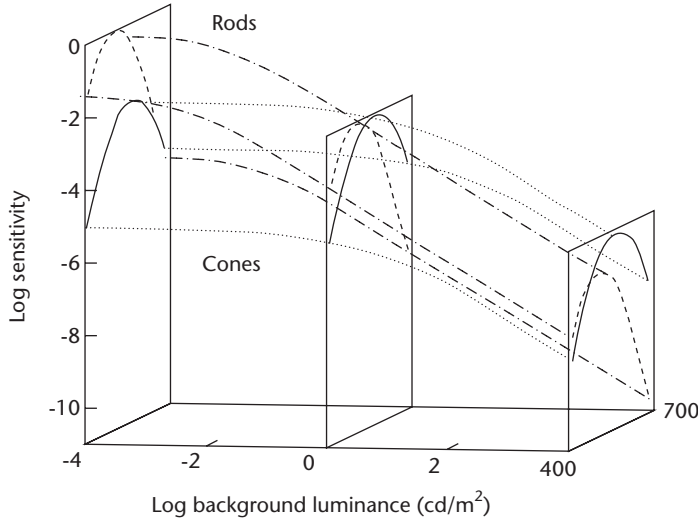
The other curve in Figure 6 shows the TVI function for the cone system. In many ways, the rod and cones show similar patterns of response. At levels below −2.6 log cd/m², the cone TVI function is essentially flat, indicating that the cones are operating at their absolute levels of sensitivity. At background levels above 2 log cd/m² the function is linear, indicating Weber's law behavior and constant contrast sensitivity.
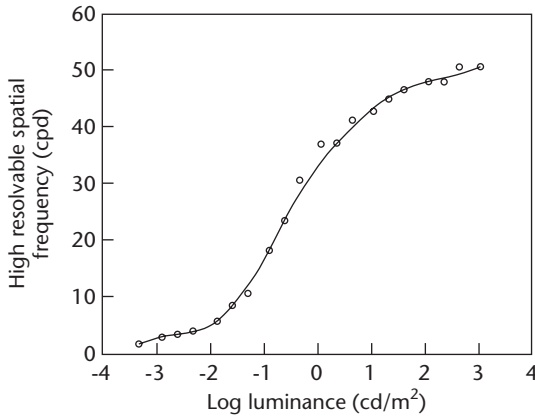
### Changes in color appearance

The scotopic and photopic *luminous efficiency functions* shown in Figure 7 describe, respectively, the spec-
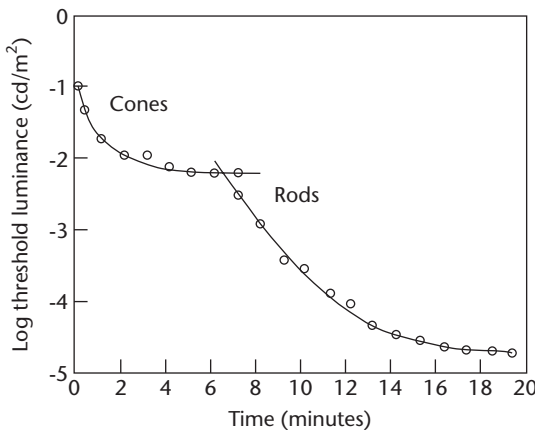
**8** Changes in spectral sensitivity at different luminance levels.



ing luminance levels. The vertical panels show cross sections through this spectral sensitivity versus luminance surface.
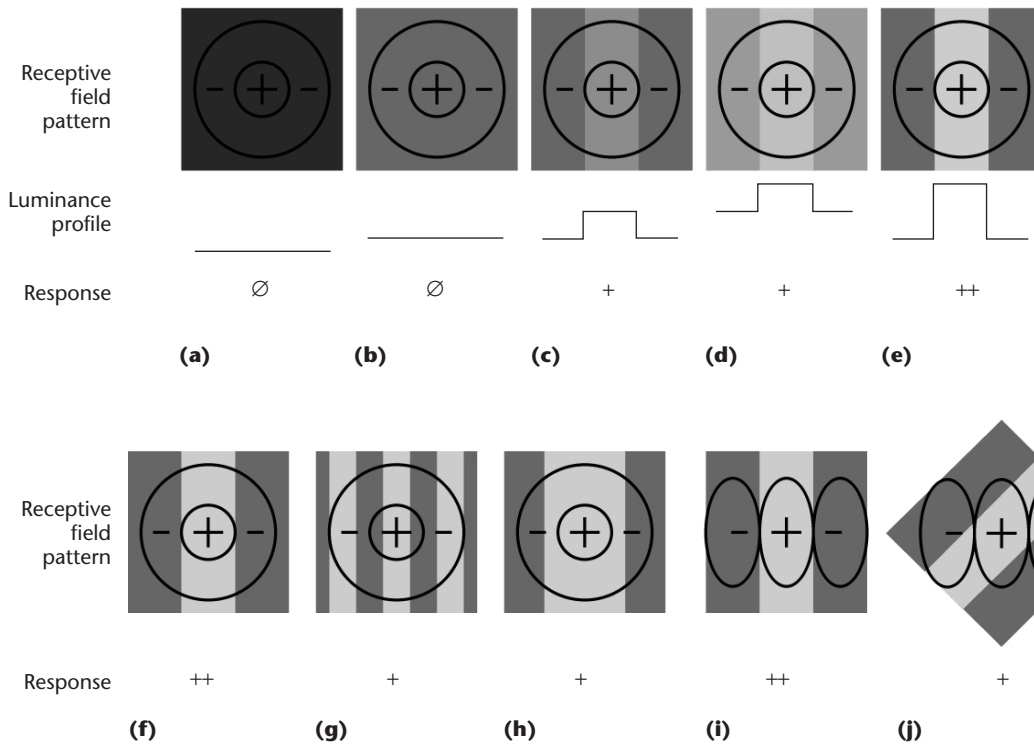
This model of the changes in spectral sensitivity with changing luminance levels can account for a number of different color appearance phenomena observed over the scotopic to photopic range. First, at low luminance levels vision is achromatic because detection at all wavelengths is served by the rod system. As the luminance level rises, the cone system becomes active and colors become visible, beginning with the long wavelength reds and progressing toward the middle wavelength greens. Only at relatively high luminances do short wavelength blue targets begin to appear colored.

**9** Changes in grating acuity (highest resolvable spatial frequency) as a function of background luminance.



### Changes in acuity

Adaptation also affects visual acuity, which is lower at scotopic levels of illumination than at photopic levels. The curve in Figure 9 shows how visual acuity changes with background luminance. The data range from bright daylight down to starlight. The experiment measured acuity by testing the detectability of square-wave gratings of different spatial frequencies. The graph shows that the highest frequency grating that can be resolved drops from about 50 cycles per degree (cpd) at 3 log cd/m$^2$ to about 2 cpd at –3.3 log cd/m$^2$. This is equivalent to a change in acuity from almost 20/10 high at daylight levels to nearly 20/300 under starlight conditions.

### The time course of adaptation

Adaptation doesn't happen instantaneously. If you're seated in a room and the lights are suddenly switched off it can take many minutes before your visual system adjusts to the new illumination level. This process is known as *dark adaptation*. Figure 10 shows the time course of dark adaptation that Hecht[17] measured. In this experiment, the observer was first adapted to a high background luminance level and then plunged into darkness. Detection thresholds were measured continuously over 20 minutes. The graph shows the detection threshold as a function of time in the dark. The kinked threshold curve is actually the envelope of the curves for the separately tested rod and cone systems. In the first 5 minutes after the adapting field is switched off, the threshold drops rapidly. Between 5 and 7 minutes, the threshold levels off at a relatively high level because although the cone system has reached its greatest sensitivity, the rod system still hasn't recovered significantly. After about 7 minutes the rod system sensitivity surpasses that of the cone system and the threshold begins to drop again. This point is known as the *Purkinje break*[16] and indicates the transition from detection by the cone system to detection by the rods. Changes in the threshold can be measured out to about 35 minutes, at which point the visual system has reached its absolute

**10** The time course of dark adaptation.



tral sensitivities of the rod and cone systems. In graphs, the functions are typically normalized, which masks the fact that the rod and cone systems differ greatly in sensitivity and operate over different luminance ranges.

Figure 8 shows the luminous efficiency functions positioned with respect to the rod and cone TVIs at different luminance levels. This 3D graph shows how the visual system's spectral sensitivity varies with chang-

**11** Properties of visual system receptive fields: (a–e) contrast processing, (f–h) spatial frequency tuning, and (i–j) orientation tuning.

levels of sensitivity, and the threshold has dropped nearly 4 log units.

The inverse of dark adaptation is *light adaptation* where the visual system adjusts to a rapid transition from lower to higher illumination levels. The time course of light adaptation is generally more rapid than dark adaptation although complete light adaptation may also take several minutes.[15]

## Spatial vision

We see by the patterns of light projected into our eyes by objects and surfaces in the visual field. Variations in the color and intensity of these patterns are essential for visual perception. If we eliminate spatial structure by uniformly illuminating the visual field we may have a sensation of light, but we don't "see" anything and our visual experience is amorphous.

The goal of *spatial vision* research is to understand the visual mechanisms that transform the light patterns in the retinal image into the colors, sizes, shapes, locations, and motions of the 3D objects we perceive in the world around us. The field has a long tradition that draws on both physiological studies of the responses of cells in the visual pathways of primates and lower animals as well as on psychophysical studies of the responses of human observers to simple visual stimuli.

### Physiology of spatial vision

One of the fundamental findings in the study of spatial vision is that the rod and cone photoreceptors aren't independent of one another but interact to form the receptive fields of retinal ganglion cells. To understand the properties of these neural networks, Kuffler[3] made electrophysiological measurements of the responses of retinal ganglion cells in a cat. He found that each ganglion cell took its input from a spatially localized receptive field with an antagonistic center/surround organization.

### Contrast processing in receptive fields

Center/surround antagonism in receptive fields results in ganglion cells that respond primarily to *contrast* rather than to simple light intensity. Figures 11a through 11e show the response of an idealized ganglion cell to various types of stimuli. In the dark (Figure 11a), the ganglion cell fires spontaneously at its base rate. If the intensity of light falling on the ganglion cell's receptive field is raised uniformly (Figure 11b), the excitatory and inhibitory regions of the field cancel and the cell continues to fire at its base rate. However, if a bar pattern with contrast between the bar and the background (Figure 11c) is introduced, then the central excitation will exceed the surround inhibition and the cell will increase its firing rate. Figures 11d and 11e show that the cell's response depends on the pattern's contrast rather than its absolute intensity. In Figure 11d the luminance of the bar and background have both increased but the cell continues to give the same response. However, when the contrast between the bar and background increases (Figure 11e), the response goes up as well.

### Spatial frequency tuning

Researchers have also found that different ganglion cells have receptive fields of different sizes. These receptive fields overlap in the retina so that at any retinal location, receptive fields of many sizes can be found.[4] Different-sized receptive fields result in ganglion cells that are selectively responsive to patterns of different

**(a)**

**(b)**
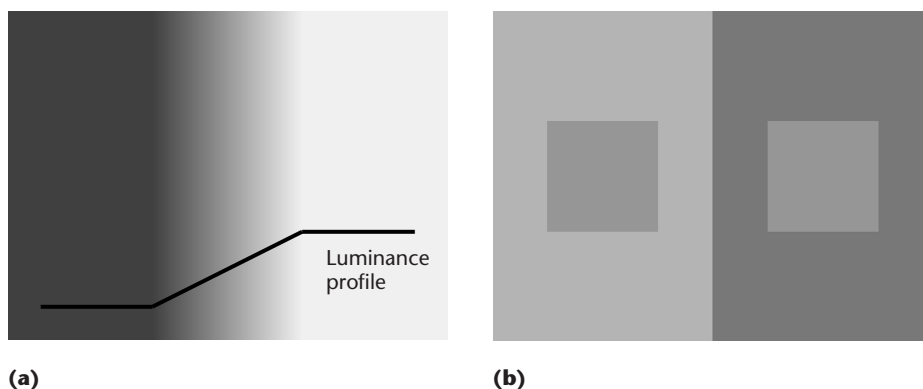
**12** **Visual phenomena related to contrast processing in receptive fields. (a) Mach bands: A luminance ramp joins two regions of differing but uniform luminance. At the transition from the dark region to the ramp, you can see a darker vertical bar. At the transition from the ramp to the light region, you can see a lighter vertical bar. These dark and light bars aren't in the image but are a product of lateral interactions in visual processing. (b) Simultaneous contrast: A middle gray target seen against a light gray surround (left) will appear darker than the same gray target seen against a dark gray surround (right). The differences in appearance are due to antagonistic interactions between the neural signals produced by different regions in the retina.**

**13** **Contrast sensitivity functions. The solid line shows a normal contrast sensitivity function. The symbols show the contrast sensitivity function after adaptation to a sine-wave grating of 7.1 cpd. The arrow marks the depression in sensitivity near the adapting frequency. (Adapted from Blakemore and Campbell.[22])**

scales. Figures 11f through 11h illustrate this *spatial frequency tuning*.

The ganglion cell's receptive field has an excitatory center and inhibitory surround. If we illuminate the receptive field with the grating pattern shown in Figure 11f, where the spatial frequency of the grating matches the width of the center and surround, there will be significant excitation from the center and not much inhibition from the surround. As a result, the cell will respond near its maximal rate. If we raise or lower the grating's spatial frequency as shown in Figures 11g and 11h, there will be both less central excitation and more

surround inhibition so the cell will respond at a lower rate. A cell's spatial frequency tuning depends on the size of its receptive field. Cells with smaller receptive fields will respond to higher ranges of spatial frequencies. Cells with larger fields will respond to lower ranges.

### *Orientation tuning*

Hubel and Wiesel[18,19] conducted electrophysiological studies of cells in the visual cortex of the cat and monkey, mapping the properties of cortical receptive fields. At this level of the visual system, cells show greater selectivity for specific features of visual patterns. For example, Hubel and Wiesel found cells that respond to edges rather than bars, showing selectivity for pattern symmetry. They also found cells that respond to motion in one direction but not in the other, bringing this selectivity to the temporal domain. One characteristic that many cells showed was selectivity for orientation. Figures 11i and 11j illustrate orientation selectivity in cortical cells.

Figure 11i shows an idealized receptive field for a cortical cell. The receptive field still shows an antagonistic center/surround organization but the field is elongated in a particular direction. This field's elongation accounts for the cell's orientation selectivity. If a grating pattern of the right spatial frequency and orientation stimulates the cell's receptive field, then there will be significant excitation and little inhibition. As a result, the cell will respond maximally. However, if we change the orientation of the grating as in Figure 11j, then there will be a mix of excitation and inhibition and the response will be reduced. Thus the cell exhibits *orientation tuning*.

## Psychophysics of spatial vision

Given the physiological evidence that visual mechanisms in animals are selective for contrast, spatial frequency, and orientation, psychophysicists began to test for the existence of similar mechanisms in human vision.

### *Contrast processing in receptive fields*

The physiological evidence for contrast processing mechanisms in human vision has a long history. Mach[20] suggested that *lateral inhibition* could account for the bright and dark *Mach bands* seen at discontinuities in luminance profiles (Figure 12a). Hering proposed that antagonism between visual mechanisms was a fundamental principle of perception that could explain important visual phenomena such as *simultaneous contrast* and *color constancy* (see Hurvich[6] for a review).
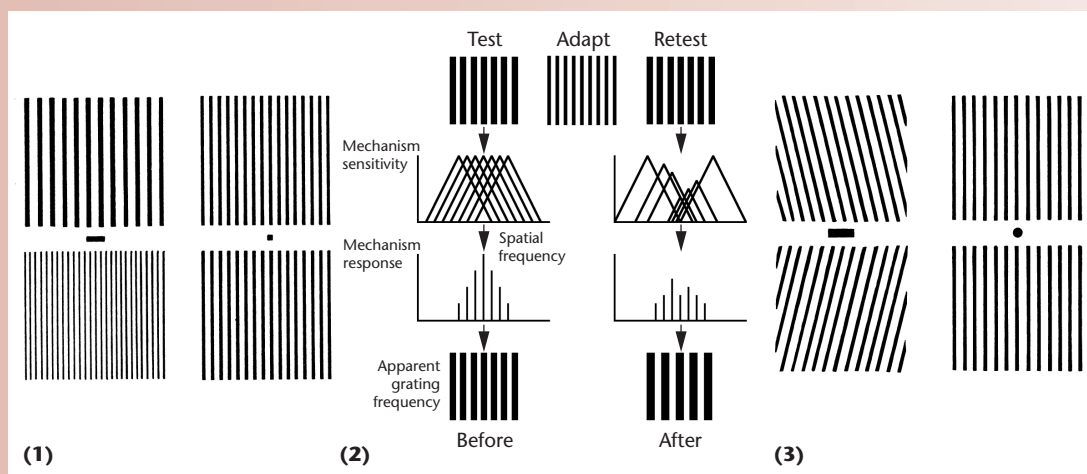
Campell and Robson tested contrast thresholds for sine-wave gratings over a range of spatial frequencies and plotted the contrast sensitivity function shown as the solid line in Figure 13. In the fovea, at the $100 \, \text{cd/m}^2$

luminance level tested, contrast sensitivity peaks at about 4 to 5 cpd where we can detect a contrast of 0.5 percent. The graph shows that threshold contrast sensitivity declines for both higher and lower spatial frequencies. At high spatial frequencies, the decline in sensitivity closely follows losses in physical image contrast due to limitations in the eye's optics. At low spatial frequencies, the decline can be at least partly explained by the limits on the sizes of the largest receptive fields.

*Spatial frequency tuning*

As shown in the previous section, the receptive field organization of visual processing in cats and primates leads to visual mechanisms tuned to different ranges of spatial frequencies. Blakemore and Campbell[22] conducted a series of psychophysical experiments to see if frequency-tuned mechanisms exist in human vision.

Their experiments used an adaptation paradigm. Prior to the experiment, they measured the subject's contrast sensitivity function. They then had the subject inspect a grating of a particular spatial frequency for one minute, instructing the subject to move his or her eyes constantly to avoid afterimages. They then remeasured the subject's contrast sensitivity function. The filled-in symbols in Figure 13 show their results.

Contrast sensitivity is reduced for spatial frequencies close to the adapting frequency. The loss of sensitivity is greatest at the adapting frequency but is also reduced within a 2-octave band around the adapting frequency. Sensitivity outside this range is unaffected. Blakemore and Campbell repeated the adaptation experiment at a number of different spatial frequencies and found a similar pattern of results in each case. Figure A1 in the sidebar "Spatial Frequency and Orientation-Specific Adaptation Afteraffects" shows a visual demonstration of spatial frequency tuning.
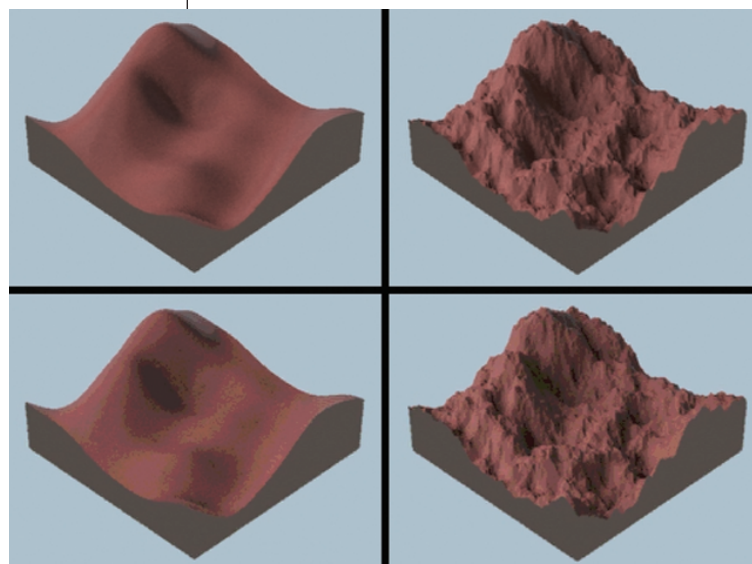
Wilson and Gelb[23] performed a set of related experiments on spatial frequency discrimination to estimate the spatial frequency tuning of visual mechanisms in the fovea. They proposed a multiple mechanism model to account for their data. The model illustrated in Figure 14 (next page) has six spatial frequency-tuned mechanisms with different peak frequencies and spatial bandwidths.

While there's ongoing debate about the number, peak frequencies, and bandwidths of spatially tuned mechanisms in human vision, the general form of the results presented by Blakemore and Campbell and Wilson and Gelb has been corroborated in numerous subsequent experiments (see Wilson[24] for a review). These results provide strong psychophysical evidence for spatial-fre-
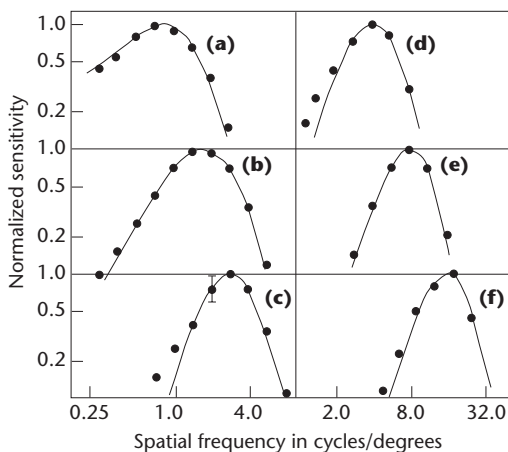
**14** Model of spatial frequency-tuned mechanisms in the human visual system. The curves show difference-of-Gaussian (DOG) fits to the data for each mechanism. Mechanisms in Figures 14a to 14f are in order of increasing peak spatial frequency. Each curve is plotted on a normalized sensitivity scale. Note that the scales in the right and left halves of the figure are different.
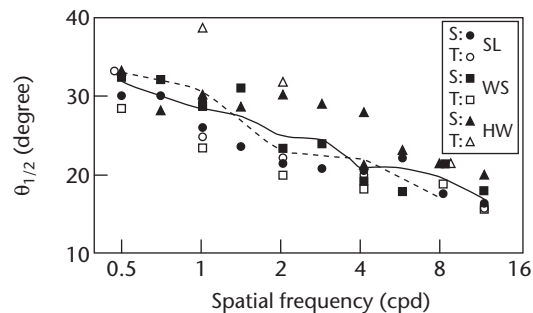


**15** Bandwidth estimates of orientation-tuned mechanisms in the human visual system. The data show the 50 percent amplitude, half-bandwidths of orientation-tuned visual mechanisms at different spatial frequencies. Different symbols are used for each of the three subjects. The filled symbols represent sustained presentations. The open symbols represent transient presentations. The solid line runs through the average half-bandwidth value at each spatial frequency. The dashed line compares these results to physiological data from primates.[27] Note that the orientation bandwidths of the mechanisms become progressively narrower with increasing spatial frequency. (Adapted from Phillips and Wilson.[26])



**16** Masking in computer graphics. The upper images are quantized to 8 bits. The lower images are quantized to 4 bits. Banding is visible in the smooth surface on the lower left but not in the rough surface on the lower right because of masking effects. (From Bolin and Meyer.[28])

quency-tuned mechanisms in human vision. Although researchers can't pinpoint the particular physiological locus of these tuned mechanisms in the human visual system, the experiments show that their influence can be measured by concrete changes in visual performance.

### Orientation tuning

We can see a similar pattern of results in psychophysical experiments that test the orientation tuning of mechanisms in human vision. Campbell and

Kulikowski[25] measured contrast sensitivity for a vertical test grating superimposed on a background grating that varied in orientation, and found evidence for orientation tuning. Phillips and Wilson[26] performed a related set of experiments to determine the orientation tuning of human visual mechanisms at different spatial frequencies. The test pattern was a spatially localized grating patch superimposed on a background grating that varied in orientation. Figure 15 shows the orientation tuning half-bandwidth of the visual system at different spatial frequencies. The results show that the visual system is more tightly tuned to orientation at high spatial frequencies than at low spatial frequencies. At a spatial frequency of 0.5 cpd the orientation bandwidth of the visual system is approximately 60 degrees (half-bandwidth times 2). At 11 cpd it has narrowed to approximately 30 degrees. This pattern of results is consistent with estimates from Campbell and Kulikowski's experiments as well as from physiological studies of the primate visual cortex.[25] Figure A3 presents a visual demonstration of orientation tuning in human vision.

### Masking

For years graphics researchers have observed that visual texture can hide artifacts in images caused by noise, aliasing, geometric tesselation, or quantization. Figure 16 shows a recent example from Bolin and Meyer[28] where banding due to quantization is much more apparent in the smooth surface on the lower left than in the rough surface on the lower right. Here the visual texture produced by the rough surface masks the banding artifact.

*Masking* is a robust perceptual phenomenon that physiologists and psychologists have studied for more than 30 years. Masking was first observed in auditory
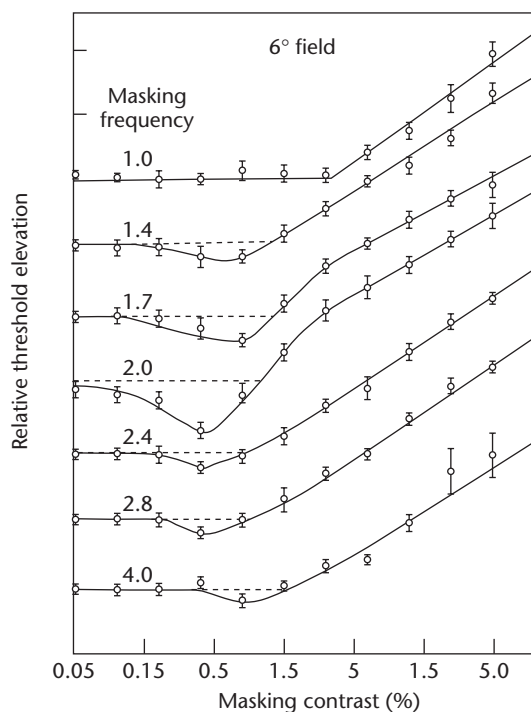
**(a)**        **(b)**

**(c)**        **(d)**
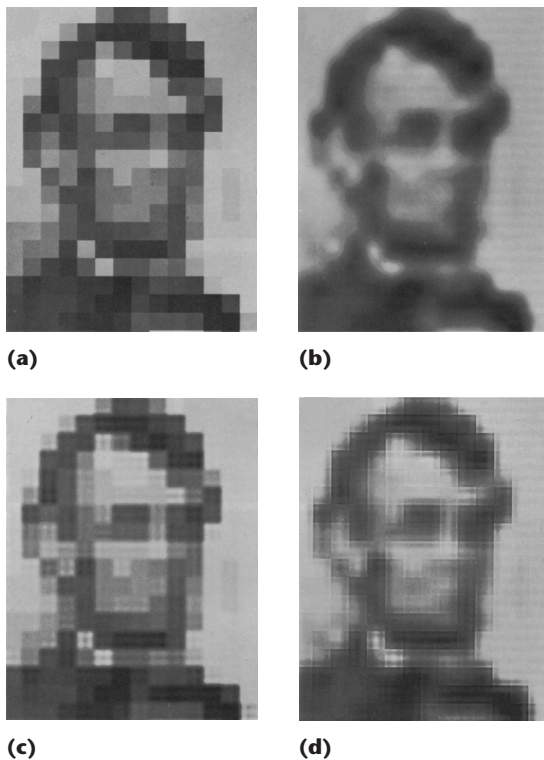
**17** Demonstration of critical band masking. (Adapted from Harmon and Julesz.[30])



**18** Threshold changes produced by masking. From Legge and Foley.[31]

perception but analogues in the visual domain were soon discovered.[29] We can define visual masking as the situation in which the presence of one visual pattern changes the visibility of another.

Figure 17, from a classic study by Harmon and Julesz,[30] illustrates the characteristics of visual masking. They lowpass filtered a continuous tone photograph of Abraham Lincoln to 10 cycles per picture height and then coarsely sampled and quantized it to produce the image in Figure 17a. Notice how this processing disturbs our ability to recognize the subject. If this blocky image is once again low-pass filtered as in Figure17b, recognition is improved. Thus it first appears that the image discontinuities introduced by high spatial frequencies in the block edges interfere with recognition. However, Harmon and Julesz showed that it's not simply high frequencies that disturb recognition but frequencies adjacent to the picture spectrum.

They termed this *critical band masking*. Thus in Figure 17c where spatial frequencies above 40 cycles have been removed, the block edges are softened but recognition is still difficult. However, in Figure 17d where frequencies between 12 and 40 cycles have been removed, the block edges are still apparent, but the subject is identifiable. This shows that masking is due to interactions within the limited spatial frequency bands because removing the critical band of frequencies directly adjacent to the picture's 10-cycle limit eliminates the masking effect but eliminating higher frequencies doesn't.

Legge and Foley[31] performed a series of experiments to determine the parameters of visual masking. In their experiments they tested how the presence of a masking grating affects the threshold for detecting a test grating. The sine wave test grating had a spatial frequency of 2.0 cpd. The masks ranged in frequency from 1.0 to 4.0 cpd. For a range of mask contrasts, they measured the contrast required to detect the test grating. Figure 18 shows their results.

The individual curves show the results for each mask frequency. Each curve is plotted on its own vertical scale showing in arbitrary units, the relative threshold changes produced by the masking grating at different contrasts. The general form of the results is that very low mask contrasts have no significant effect on the visibility of the test grating. However, as mask contrasts increase, at first the threshold drops slightly, but then rises showing a loss in sensitivity (threshold elevation) for seeing the test grating in the presence of the mask.

The curves in Figure 18 also show the spatial frequency tuning of masking. Loss of sensitivity is greatest when the mask and test gratings have the same spatial frequency. As the spatial frequencies of the mask and test gratings diverge, greater and greater mask contrasts are necessary to produce the same threshold elevation.

Legge and Foley's masking results provide evidence for a contrast nonlinearity in the visual system that has important implications for how the features of the world are coded by the visual system. See Graham[32] for a comprehensive review of masking and other contemporary issues in spatial vision.

## Conclusion

To a large extent, the properties of early visual mechanisms determine both the limits and capabilities of visual perception. This tutorial has surveyed some of the fundamental findings in the study of early vision. An

## Further Reading

This tutorial has barely scratched the surface of issues in vision research that have relevance for the field of computer graphics. Fortunately, a number of good resources are available. Both Palmer[1] and Wandell[2] have written excellent recent texts that survey not only early vision but also higher order issues in visual processing. Journals such as *Vision Research, Perception and Psychophysics, Journal of Experimental Psychology,* and *Journal of the Optical Society of America* regularly publish cutting-edge research in the field. Finally, conferences such as ACM Siggraph and SIGCHI and the International Society for Optical Engineering/Society for Imaging Science and Technology (SPIE/IS&T) conference on human vision and electronic imaging are attracting growing numbers of researchers and practitioners interested in the potential synergies between human vision research and advanced computer graphics techniques.

### References

1. S. Palmer, *Vision Science: Photons to Phenomenology,* MIT Press, Cambridge, Mass., 1999.
2. B. Wandell, *Foundations of Vision,* Sinauer Associates, Sunderland, Mass., 1995.

understanding of early visual processing is currently driving the development of perceptually based algorithms that are improving both the efficiency and the effectiveness of graphics methods. Further study of both early and higher levels of visual processing should provide new insights that will allow us to solve many important problems in computer graphics. ∎

## Acknowledgments

## References

1. E.N. Pugh, "Vision: Physics and Retinal Physiology," *Steven's Handbook of Experimental Psychology,* 2nd ed., R.C. Atkinson, ed., John Wiley & Sons, New York, 1988, pp. 75-163.
2. R.C. Atkinson, ed., *Steven's Handbook of Experimental Psychology,* 2nd ed., John Wiley & Sons, New York, 1988.
3. S.W. Kuffler, "Discharge Patterns and Functional Organization of the Mammalian Retina," *J. Neurophysiology,* vol. 16, 1953, pp. 37-68.
4. C. Enroth-Cugell and J.G. Robson, "The Contrast Sensitivity of Retinal Ganglion Cells of the Cat," *J. Physiology,* vol. 187, 1966, pp. 517-552.
5. R.L. DeValois and K.L. DeValois, "Neural Coding of Color," *The Handbook of Perception,* vol. 5, E.C. Carterette and M.P. Friedman, eds., Academic Press, New York, 1975, pp. 117-162.
6. L. Hurvich, *Color Vision,* Sinauer Assoc., Sunderland, Mass., 1981.
7. R. Sekuler and R. Blake, *Perception,* McGraw-Hill, New York, 1994.
8. P. Lennie, "Recent Developments in the Neurophysiology of Color," *Trends in Neuroscience,* vol. 7, 1984, pp. 243-248.
9. M. Mishkin, L.G. Ungerleider, and K.A. Macko, "Object Vision and Spatial Vision: Two Critical Pathways," *Trends in Neuroscience,* vol. 6, 1983, pp. 414-417.
10. J.P. Thomas, "Spatial Resolution and Spatial Interaction," *The Handbook of Perception,* vol. 5, E.C. Carterette and M.P. Friedman, eds., Academic Press, New York, 1975, pp. 233-263.
11. G. Osterberg, "Topography of the Layer of Rods and Cones in the Human Retina," *ACTA Ophthamologica Supplementum,* vol. 6, 1935, pp. 11-97.
12. A.W. Snyder and W.H. Williams, "Photoreceptor Diameter and Spacing for Highest Resolving Power," *J. Optical Soc. of America,* vol. 67, no. 5, 1977, pp. 696-698.
13. G. Westheimer, "Spatial Frequency and Light Spread Descriptions of Visual Acuity and Hyperacuity," *J. Optical Soc. of America,* vol. 67, no. 2, 1977, pp. 207-212.
14. R.J. Watt and M.J. Morgan, "Mechanisms Responsible for the Assessment of Visual Location: Theory and Evidence," *Vision Research,* vol. 23, 1983, pp. 97-109.
15. L. Spillman and J.S. Werner, eds., *Visual Perception: The Neurophysiological Foundations,* Academic Press, San Diego, 1990.
16. L.A. Riggs, "Vision," *Woodworth and Schlosberg's Experimental Psychology,* 3rd ed., J.W. Kling and L.A. Riggs, eds., Holt, Rinehart, and Winston, New York, 1971, pp. 273-314.
17. S. Hecht, "Vision II: The Nature of the Photoreceptor Process," *A Handbook of General Experimental Psychology,* C. Murchison, ed., Clark University Press, Worchester, Mass., 1934, pp. 78-93.
18. D.H. Hubel and T.N. Wiesel, "Receptive Fields, Binocular Interaction, and Functional Architecture in the Cat's Visual Cortex," *J. Physiology,* vol. 160, 1962, pp. 106-154.
19. D.H. Hubel and T.N. Wiesel, "Receptive Fields and Functional Architecture of Monkey Striate Cortex," *J. Physiology,* vol. 195, 1968, pp. 215-243.
20. F. Ratliff, *Mach Bands: Quantitative Studies on Neural Networks in the Retina,* Holden-Day, San Francisco, 1965.
21. F.W. Campbell and J.G. Robson, "Application of Fourier Analysis to the Visibility of Gratings," *J. Physiology,* vol. 197, 1968, pp. 551- 566.
22. C. Blakemore and F.W. Campbell, "On the Existence of

Neurones in the Human Visual System Selectively Sensitive to the Orientation and Size of Retinal Images," *J. Physiology*, vol. 203, 1969, pp. 237-260.

23. H.R. Wilson and D.J. Gelb, "Modified Line-Element Theory for Spatial-Frequency and Width Discrimination," *J. Optical Soc. of America*, vol. 1, 1984, pp. 124-131.

24. H.R. Wilson, "Psychophysical Models of Spatial Vision and Hyperacuity," *Spatial Vision*, D. Regan, ed., vol. 10, CRC Press, Boca Raton, Fla., 1991, pp. 64-86.

25. F.W. Campbell and J.J. Kulikowski, "Orientation Selectivity of the Human Visual System," *J. Physiology*, vol. 187, 1966, pp. 437-445.

26. G.C. Phillips and H.R. Wilson, "Orientation Bandwidths of Spatial Mechanisms Measured by Masking," *J. Optical Soc. of America*, vol. 1, 1984, pp. 226-232.

27. R.L. DeValois, E.W. Yund, and N. Hepler, "The Orientation and Direction Selectivity of Cells in Macaque Visual Cortex," *Vision Research*, vol. 22, 1982, pp. 531-544.

28. M.R. Bolin and G.M. Meyer, "A Frequency Based Ray Tracer," *Proc. Siggraph 95*, ACM Press, New York, 1995, pp. 409-418.

29. A. Pantle, and R.W. Sekuler, "Contrast Response of Human Visual Mechanisms Sensitive to Orientation and Direction of Motion," *Vision Research*, vol. 9, 1969, pp. 397-406.

30. L.D. Harmon and B. Julesz, "Masking in Visual Recognition: Effects of Two-Dimensional Filtered Noise," *Science*, vol. 180, 1973, pp. 1194-1197.

31. G.E. Legge and J.M. Foley, "Contrast Masking in Human Vision," *J. Optical Soc. of America*, vol. 70, 1980, pp. 1458-1470.

32. N.V. Graham, *Visual Pattern Analyzers*, Oxford University Press, New York, 1989.

*James A. Ferwerda is a research associate in the Program of Computer Graphics at Cornell University where he leads an interdisciplinary group studying perceptual issues in computer graphics. His current work focuses on developing computational models of human vision from psychophysical experiments and implementing graphics algorithms based on these visual models. He received a BS in 1980, an MS in 1987, and a PhD in 1998, all from Cornell University. He is a member of Siggraph and the Society for Imaging Science and Technology (IS&T).*

*Readers may contact Ferwerda at the Program of Computer Graphics, Cornell University, 580 Rhodes Hall, Ithaca, NY 14853, email jaf@graphics.cornell.edu.*

For further information on this or any other computing topic, please visit our Digital Library at http://computer.org/publications/dlib.