Psychological Review
1989, Vol. 96, No. 2, 267-314

0033-295X/89/$00.75

Sequential Ideal-Observer Analysis of Visual Discriminations

Wilson S. Geisler
University of Texas

Visual stimuli contain a limited amount of information that could potentially be used to perform a given visual task. At successive stages of visual processing, some of this information is lost and some is transmitted to higher stages. This article describes a new analysis, based on the concept of the ideal observer in signal detection theory, that allows one to trace the flow of discrimination information through the initial physiological stages of visual processing, for arbitrary spatio-chromatic stimuli. This ideal-observer analysis provides a rigorous means of measuring the information content of visual stimuli and of assessing the contribution of specific physiological mechanisms to discrimination performance. Here, the analysis is developed for the physiological mechanisms up to the level of the photoreceptor. It is shown that many psychophysical phenomena previously attributed to neural mechanisms may be explained by variations in the information content of the stimuli and by preneural mechanisms.

The purpose of vision is to extract and represent information about the physical environment from the light that is emitted, transmitted, or reflected by objects and surfaces. In order to extract useful information, a visual system must be able to encode and classify changes that occur in the visual image. This article presents a new theoretical approach to analyzing the role of specific physical and physiological mechanisms in the simplest form of classification task—the discrimination or binary classification task, of which the detection task is an important special case. Although the emphasis here is on detection and discrimination in the human visual system, the general approach should be widely applicable to other classification tasks[1] and to other visual systems.

The physical and physiological machinery underlying human vision is extremely complex. This complexity is hinted at in Table 1, which lists some of the important physiological mechanisms, or factors, governing the initial stages of visual processing.

The processing begins with an optical system that forms, in perspective projection, a two-dimensional image on an array of cone and rod photoreceptors, whose size, collection area, and spatial arrangement varies with eccentricity (as do the optics). Each cone contains one of three photopigments; the rods all contain the same photopigment. These photopigments set limits on the receptors' absolute sensitivities and determine their relative sensitivities as a function of wavelength. Complex physiological processes within the receptors determine their response profiles (over time) as a function of the input intensity profile and previous adaptation conditions. Then, within the retina, the outputs of the photoreceptors are processed at a series of neural stages. Each neuron in each retinal stage combines the output of the previous (and/or current) stage over some region of space and time. Several more stages of parallel neural processing then occur along the primary visual pathway, first in the lateral geniculate nucleus (LGN) and then in area 17 of the visual cortex. At each stage, there are generally a number of different classes of neuron (neural channels), each presumably combining the output of the previous stages in a different fashion.

The information-processing operations carried out by the neurons at each stage are often described by receptive-field properties. However, it is important to recognize that the spatial, temporal, and chromatic receptive-field properties of a neuron at any stage are determined by the cumulative effect of all the earlier stages beginning with the optical system. In other words, multiple mechanisms at various levels of the visual system contribute to the physiological response of later stages and to overall visual performance.

It is not surprising that the complexity of the visual system

This research was supported by National Institutes of Health Grant EY02668 and by an equipment grant from the College of Liberal Arts at the University of Texas.

I am indebted to Martin Banks, Karen Davila, Dennis McFadden, David Williams, and the two anonymous reviewers for extensive comments on this article. Randy Diehl also provided useful comments and discussion. Gordon Hanka wrote the graphics subroutines used to create the figures showing the patterns of photon absorptions in the receptors. Larry Stern organized the references and helped with the plotting of data and predictions. Gail Geisler drafted the figures.

Correspondence concerning this article should be addressed to Wilson S. Geisler, Department of Psychology, University of Texas at Austin, Mezes Hall 330, Austin, Texas 78712.

[1] The prominence of classification tasks (particularly discrimination tasks) in the study of vision stems from several factors. First, they provide objective, repeatable measures of visual performance. Second, behavioral classification tasks can be used to rigorously test physiological hypotheses. As Brindley (1970) pointed out, if a physiological hypothesis predicts that the neural responses to two (or more) stimuli become identical at some stage along the visual pathway, then the stimuli must be indistinguishable to the organism. Therefore, if a classification experiment shows that they are distinguishable, then there is no recourse but to reject the physiological hypothesis. Third, it is clear, on a priori grounds, that a primary (if not *the* primary) goal of visual processing is to quickly and accurately classify (i.e., recognize) regions or parts of visual images.

has hindered development of rigorous, general theories of visual processing and, at the same time, encouraged specialization in subareas of vision science such as color vision, visual adaptation, temporal and flicker sensitivity, and spatial vision. The major goal of this article is to present a theoretical framework and analysis that may provide a rigorous method for analyzing complex visual systems in a straightforward and unified fashion (Geisler, 1984; Geisler, 1987a; Geisler & Davila, 1985). This method, which is based on the theory of ideal observers (de Vries, 1943; Green & Swets, 1974; Peterson, Birdsall, & Fox, 1954; Rose, 1942), allows one to determine the precise roles of specific physical and physiological mechanisms in limiting the discrimination performance of later physiological stages and of the system as a whole. Because this ideal-observer analysis can be applied to almost arbitrary discrimination tasks, it provides a unifying perspective that cuts across the many specialized areas of visual science.

The plan of the article is as follows. First, the logic and rationale of the ideal-observer analysis is presented. Second, a review of what is currently known about the mechanisms involved in image formation by the optics of the eye and in image sampling by the photoreceptors is given. These include approximately the first half of the mechanisms listed under optics and receptors in Table 1. The emphasis is on these mechanisms because they are the best understood at the quantitative level. Third, an ideal-observer analysis of these initial mechanisms is developed and applied to a wide range of spatial and chromatic discrimination tasks. Fourth and finally, the general implications and further uses of ideal-observer analysis are discussed.

Sequential Ideal–Observer Analysis

In the proposed analysis, the early visual system is regarded as a sequence of information-transmission or processing stages.

Table 1
Some Mechanisms of Initial Visual Processing

Optics and receptors
Quantal fluctuations
Optics of the eye
Eye movements
Receptor size, shape, and waveguide properties
Receptor sampling lattice
Receptor spectral sensitivities
Photopigments and photochemistry
Internal transduction process
Temporal response properties
Receptor nonlinearities
Receptor adaptation/gain
Retina and lateral geniculate nucleus
Spatial, temporal, and chromatic receptive-field shapes
Receptive-field sampling lattice
Response dynamic-range nonlinearities
Spatial, temporal, and chromatic nonlinearities
Neural noise
Neural adaptation
Visual cortex
Everything at the retinal and geniculate level
Binocular interactions
Multiple representations
Cortical interactions

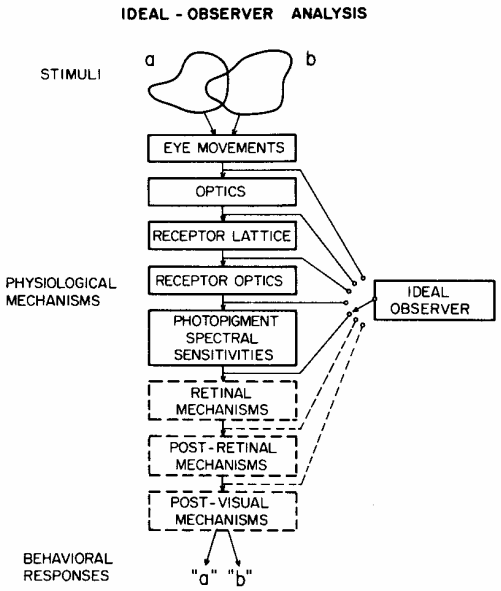


Figure 1. Sequential ideal-observer models consist of three parts: a complete representation of the stimuli, a model of the physiological mechanisms, and an ideal observer designed to process optimally the output of a given physiological mechanism.

Before entering the eye, a visual stimulus contains some total amount of information that potentially may be used to perform a classification task. This information is then processed by a long series of visual mechanisms (each of which operate spatially in parallel). The goal of the proposed analysis is to measure quantitatively, with ideal observers, the total amount of information relevant to the classification task that is lost or delayed in transmission at each step along the visual pathway. Thus, one can, in principle, determine exactly how much and in what way each visual mechanism contributes to visual performance.

The logic of this sequential ideal-observer analysis is not entirely new. Indeed, it is widely recognized that many of the important discoveries about human visual processing have involved determination of the information lost at some physiological stage or stages of processing. A classic example was the discovery that the psychophysical laws of color mixture (according to which physically different stimuli can match exactly) are the result of having only three types of photopigments within three classes of photoreceptor that each obey the *univariance* principle.[2] What is new about the present approach is the use of the theory of ideal observers, in conjunction with physiological data and models, to compute rigorously the information transmitted through various levels in the visual system for arbitrary spatial and chromatic stimuli.

The structure of a sequential ideal-observer analysis is illustrated in Figure 1. There are three major steps in carrying out

[2] The univariance principle states that once a photon is effectively absorbed, it has the same physiological effect on the receptor regardless of the photon's wavelength. The principle appears to hold.

an analysis. The first is to obtain precise representations of the stimuli that are to be classified. These representations must include a description of their spatial, temporal, and chromatic properties, as well as any sources of noise or random variation. Even under the most controlled conditions, visual stimuli randomly change from moment to moment because of the inherent Poisson randomness of light. Thus, as shown in Figure 1, the two nominal stimuli in a discrimination experiment must be represented as sets.

The second step is to specify a quantitative model of the physiological mechanisms under consideration. The solid boxes in Figure 1 show the initial sequence of mechanisms in the human visual system up to the absorption of light in the receptor photopigments. These are the only mechanisms considered in this article.

The third step is to compute both the total information relevant to the classification task and how much of that information is transmitted by the mechanism or sequence of mechanisms under consideration. This can be done by designing a machine that uses all of the transmitted information to achieve optimal classification performance. By definition, the performance of this machine would be a measure of the total transmitted information. Fortunately, the design of such optimal machines is a common problem in statistical decision theory and is exactly the sort of problem that was solved in developing many ideal-observer models in both vision (Barlow, 1957; de Vries, 1943; Rose, 1942; Tanner & Clark-Jones, 1960) and audition (Green & Swets, 1974; Peterson, Birdsall, & Fox, 1954; Tanner & Clark-Jones, 1960). Thus, the well-understood methods of statistical decision theory can, in principle, provide us with a precise metric of the classification information available at any stage in the physiological model.

There are four important ways in which carrying out such an ideal-observer analysis can help one understand human visual discrimination.

1. By switching the location of the ideal observer (see Figure 1), one can determine how much information is lost or transmitted by a specific stage or sequence of stages. This is done by comparing the performance of the ideal observer when it is applied before and after the stage(s). Note that in general, the design of the ideal observer will have to be different at each stage.

2. By comparing the performance of the ideal observer with human performance for identical stimuli, one can determine how much information was lost in the physiological stages following the one tapped by the ideal observer.

If the ideal discriminator is placed behind a perfect optical system (i.e., if all physiological mechanisms are excluded), comparison of ideal and real performance reveals the absolute efficiency of the real visual system for the discrimination. This special case is the classical form of ideal-observer theory used in vision and audition.

3. The proposed analysis permits rigorous comparison of human performance across different tasks. Thus, it becomes possible to ask and to answer meaningfully questions of the following sort: Is intensity discrimination better or worse than color discrimination? Is vernier acuity better than Snellen (eye-chart) acuity? What spatial patterns does the eye see best?

4. Perhaps most important, ideal-observer analysis provides us with a better understanding of the stimuli used in discrimination tasks. An appropriate ideal discriminator placed at an early level of the visual system provides a precise metric of all of the information available to the later stages (assuming that the model of the included physiological mechanisms is sound). If altering the stimuli in some discrimination task results in little or no change in the information available, then we are not surprised if human performance changes little. Similarly, if the information content changes a great deal, then we are not surprised if human performance is greatly affected. On the other hand, if human and ideal performance are not affected similarly by changes in the stimuli, we learn that the human visual system is failing to use some aspect(s) of the information contained in the stimuli. In other words, ideal-observer analysis tells us when we should attribute changes in performance to higher level mechanisms and when we should attribute them to differences in the information content of the stimuli at the lower level being analyzed. Ideal-observer analysis tends to be more informative when real and ideal performance are close in absolute level. When they are relatively close, strong contraints are placed necessarily on the nature of the higher level mechanisms.

It is important to keep in mind that the present ideal-observer analysis only measures the information in the stimuli that is relevant for the particular discrimination task under consideration. The measurement of stimulus information in most real-world tasks is enormously more complicated (and may be impossible) because it depends not only on the set of possible stimuli but on the prior probabilities of their occurances. In simple laboratory discrimination tasks, the prior probabilities are precisely controlled and hence known.

De Vries (1943) and Rose (1942, 1948) were the first to apply the concept of the ideal observer to visual psychophysics. They found that the optimal intensity threshold for detecting a spot of light against a background of light is proportional to the square root of the background intensity (the so-called square-root or de Vries-Rose law). De Vries (1943) also suggested developing ideal observers for acuity tasks but did not pursue this idea (it is developed more fully here). Barlow (1958b) and Tanner and Clark-Jones (1960) derived the intensity threshold for the ideal observer as a function of target area and duration. They found that threshold energy was proportional to the square root of target area (Piper's law) and to the square root of target duration (Pieron's law). During this same period, there was an even more sophisticated development of ideal-observer theory in auditory psychophysics (see Green & Swets, 1974, for a review; Peterson, Birdsall, & Fox, 1954; Van Meter & Middleton, 1954).

Despite these important developments, ideal-observer theory subsequently received less attention in visual psychophysics than one might have expected given the theory's rigorous physical basis. There are two probable reasons for this. The first is that ideal-observer theory generally does not predict human performance very well. For example, Barlow (1958b) and others showed that the aforementioned predicted relations between threshold energy and background intensity, target area, and target duration hold only over limited ranges of these variables. It is, of course, quite natural to ignore theories that do not fit the data. However, as the developers of ideal-observer theory recognized, this is not a valid reason for dismissing an ideal-observer

theory, because it is not a model of human performance but is a model of the physical factors that limit human performance. Thus, the theory can only be rejected if the physics or the calculations are proven wrong. If the physics and calculations are correct, then the theory necessarily stands as one correct piece of a complete theory. Furthermore, a correctly computed ideal-observer analysis can be just as useful when it provides a bad fit to data as when it provides a good one. In the former case, one learns that subsequent mechanisms contribute significantly to performance by creating more information loss. In the latter case, one learns that the mechanisms included in the analysis are the major source of information loss and, hence, that the subsequent mechanisms are of high uniform efficiency.

The second major reason for the relatively small impact of ideal-observer theory in visual psychophysics is that most early applications of the theory dealt with just one (albeit important) task—the detection of test spots against uniform backgrounds. In recent years, however, the range of applications and interest in ideal-observer theory has expanded so that it now represents an important tool for vision research. Vos and Walraven (1972) developed an ideal-observer theory for chromatic discrimination that will be described in more detail later. Cohn and his colleagues (Cohn, 1976; Cohn & Lasley, 1974, 1986; Cohn & Wardlaw, 1985) have compared human and ideal observers on increment and decrement detection tasks and on increment detection tasks with spatial uncertainty. Watson, Barlow, and Robson (1983), Pelli (1985), Kersten (1984), and Banks, Geisler, and Bennett (1987) compared human and ideal performance for detecting sine-wave grating targets. Geisler (1984), Geisler and Davila (1985), and Geisler (1987a) developed and applied ideal-observer theory to various acuity tasks. A related early effort at developing an ideal observer for position acuity was carried out by Andrews, Butcher, and Buckley (1973). However, because their model does not consider photon noise and includes other hypothesized sources of internal noise, it does not qualify as an ideal-observer theory in the present sense. Nonetheless, as we shall see later, some of the physiological factors considered in their model (optical blur and receptor collection area) can be sensibly incorporated into a sequential ideal-observer analysis.

Barlow (1978), van Meeteren and Barlow (1981), Burgess, Wagner, Jennings, and Barlow (1981), Kersten (1984), and others have also applied ideal-observer theory to tasks involving the detection of targets in high levels of computer-generated pixel noise. One goal of using pixel noise in psychophysical tasks is to overwhelm the effects of photon noise and other peripheral sources of information loss. Under such conditions, differences between real and ideal performance may measure information loss in cortical processing. This interesting work will not be considered further here because my primary concern is with the limitations on performance imposed by photon noise and peripheral physiological mechanisms.

The present work extends the previous work with ideal observers in two ways. First, it expands the application of the ideal observer to a much wider range of tasks. The basic mathematical results allowing this expansion are summarized here. They provided the basis for a computer package that allows the performance of the ideal observer to be computed in almost arbitrary discrimination tasks. Second, the present work extends the rigorous use of ideal observers to include physiological factors. These extensions may increase the usefulness of ideal observers in analyzing the visual system. A similar development might be of value in analyzing other sensory systems.

Image Formation and Image Sampling Mechanisms

In the sections that follow I will attempt to summarize what is currently known about the image-formation and receptor-sampling properties of the human eye that affect visual discrimination. The emphasis is on quantitative measurements and analyses that support development of the sequential ideal-observer analyses presented later. Although this literature review will focus on studies of the human eye, some physiological or anatomical measurements are difficult or impossible to obtain on humans. In these cases, our best information comes from studies of the visual system of macaque monkey. Anatomical and psychophysical studies suggest that the macaque visual system is remarkably similar to that of the human visual system in both structure and function (DeValois, Morgan, & Snodderly, 1974; Harwerth, Boltz, & Smith, 1980).

Before proceeding, it is important to mention that in recent years, the mathematical and engineering techniques associated with linear systems have been used extensively in the empirical investigation of optical and neural mechanisms and in modeling their effects on visual performance. A fair portion of the studies reviewed later were based on linear-systems techniques. We have also used linear-systems techniques in applying our ideal-observer theories. Because some readers may be unfamiliar with these techniques, a brief description of the major concepts is provided in Appendix A. For a more detailed treatment, the reader should consult one of the many excellent texts (e.g., Bracewell, 1978; Gaskill, 1978).

Photon Noise

Visible light propagates in the form of electromagnetic waves, but it is emitted and absorbed in finite quanta of energy called photons. The emission and absorption of photons is a random process that can usually be adequately modeled as a simple Poisson process (D. L. Snyder, 1975), although there are a few cases in which this is not so (Teich, Prucnal, Vannucci, Breton, & McGill, 1982).[3]

Because the emission and absorption of light is described by the Poisson process, the probability that a given number of photons will be absorbed or emitted in a fixed time interval over a fixed area is described by the Poisson density:

$$p(z) = \alpha^z \exp(-\alpha)/z! \quad z = 0, 1, 2, \ldots, \tag{1}$$

[3] Cathode-ray tubes (CRTs) are examples of non-Poisson light sources. The electrons released from the cathode are Poisson. However, each electron hitting the CRT phosphor has a high probability of releasing more than one photon. Because the photons are emitted in "bunches," they do not satisfy the independence property of the Poisson process. Fortunately, at reasonable viewing distances the probability is extremely small of receiving, through the pupil, more than one quantum from a bunch. Under these conditions, the photons absorbed in the retinal photopigments are still described by a Poisson process.

where z is the number of photons absorbed, and α is the mean number of photons absorbed (or emitted) in a fixed time interval over a fixed area. The variance of the Poisson density is also equal to $-\alpha$. Thus, as light level goes up, the variance of the absorbed and emitted photons increases. Photon noise places a physical limit on human visual discrimination performance. The Poisson density is useful for computing this limit for the stimulus conditions that will be of interest here.

Image Formation

Electromagnetic waves emitted from a direct source (e.g., the filament of a lamp) or an indirect source (e.g., the surface of an object) usually propagate in all directions. The purpose of the optical system in the eye is to unscramble the light waves so that light emitted at a given distance along each line of sight is imaged as well as possible at a unique point on the receptor array. For example, consider a very small patch of surface (small enough to be considered a point source) producing diverging spherical wavefronts. A perfect eye would convert these wavefronts into converging spherical wavefronts that collapse to a single point at the inner segment layer of the photoreceptors. Of course, this image-formation process is not perfect in real eyes and thus adds further limits to visual sensitivity. Imperfections in the optics of the eye generally work to blur the image—light that should be focused at a single point is spread out to some extent. (Some imperfections produce distortions instead of blur.)

The major structures affecting image formation in the mammalian eye are the cornea, the iris, and the crystalline lens. The front surface of the cornea is the main refracting surface of the eye. The iris forms a circular aperture or pupil of adjustable size. The lens is a secondary refracting structure with adjustable power that is used to reduce focusing errors for objects at different distances.

Image formation in the human eye is complex and is understood best for stimuli falling in or near the fovea, when accommodation is accurate and fixation is steady. I will consider mainly these conditions because they are understood best and because they are the most common in psychophysical experiments. Recall that the goal is to obtain a reasonably precise model of image formation in the human eye so that it can be included in a quantitative ideal-observer analysis. In the subsections that follow, various sources of information loss in the image-formation process are reviewed. The combined effect of all of these image-forming factors can be represented in a single *point-spread function* that describes how a point of light is spread out across the retina.

Monochromatic and chromatic aberrations. One class of errors in the image-formation process is the monochromatic aberrations (aberrations that would occur in monochromatic light). Monochromatic aberrations are usually expressed as deviations from an ideal spherical wavefront converging to a point in the desired image plane (e.g., the receptor layer). Almost any form of deviation is possible, but for most optical systems they can be described adequately in terms of a fourth-order two-dimensional Taylor polynomial (Goodman, 1968; Howland & Howland, 1977). These include prismatic aberrations (first order), focusing and astigmatic aberrations (second order), comatic (not chromatic) aberrations (third order), and spherical aberrations (fourth order). In monochromatic light, prismatic aberrations have no effect on image clarity because they only produce a lateral shift of the image. Focusing aberrations occur when the eye is not optimally accommodated. Astigmatism occurs when there is a cylindrical component to the shape of the cornea or lens, making the power of the optical system greater along one axis of the image plane. The result is that accommodation can only eliminate focusing errors in one axis at a time. When eyeglasses are prescribed, the intent is to reduce astigmatism and/or to reduce focusing errors (due to myopia or hyperopia) by bringing the power of the eye within the accommodation range of the lens. In the normal or corrected eye, residual focusing errors and astigmatism and the higher order aberrations still limit retinal image quality and, hence, produce some information loss. Unfortunately, even in the normal eye there are substantial individual differences in residual astigmatism (Rempt, Hoogerheide, & Hoogenboom, 1971), coma (Walsh, Charman, & Howland, 1984), and spherical aberration (van Meeteren, 1974; Walsh, Charman, & Howland, 1984). This makes it difficult to arrive at a single representative model of the monochromatic aberrations. Fortunately, other factors dominate image quality under normal viewing conditions, thereby reducing the relative importance of individual differences in the monochromatic aberrations (after correction for focusing and astigmatic errors).

Naturally occurring stimuli always contain a wide range of wavelengths, and different wavelengths of light are refracted by different amounts in passing from one optical medium into another. This introduces another source of imaging error, chromatic aberration. Longitudinal chromatic aberration occurs because only one wavelength of light can be in precise focus at a given time. The level of defocus at other wavelengths increases, the further the wavelengths are from the one that is in focus. Lateral chromatic aberration occurs because different wavelengths from objects that are off the optical axis of the eye are imaged at different lateral positions. Lateral chromatic aberrations are generally most severe for eccentric stimuli, but are larger than one might expect in the fovea because the fovea is slightly eccentric to the optical axis (van Meeteren, 1974). Chromatic aberrations are well understood, easily measured, and quite consistent in magnitude from one eye to the next (Bedford & Wyszecki, 1957; Charman & Jennings, 1976; Wald & Griffin, 1947). In spectrally broad-band (white) light, chromatic aberration is probably the dominant aberration (van Meeteren, 1974).

Pupil diffraction. Even in an ideal, aberration-free optical system, a pupil or aperture will diffract light waves, thereby reducing image quality. The effect of diffraction is illustrated in Figure 2 for several pupil sizes. These are for artificial pupils placed just in front of the cornea. The dashed lines show cross-sections of the retinal intensity distributions (diffraction patterns) that would be produced by thin lines of incoherent white light lying near the optical axis in an aberration-free eye. As can be seen, the amount of diffraction increases as the size of the pupil decreases. An incoherent point source produces a similar diffraction pattern that is radially symmetric. The solid curves show the actual retinal intensity distributions measured by Campbell and Gubisch (1966) with white light, for four pupil

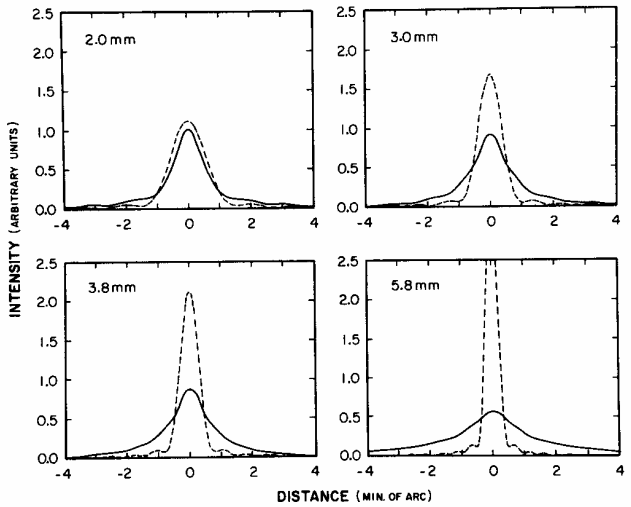


Figure 2. Optical line-spread functions in the fovea for artificial pupil diameters ranging from 2.0 mm to 5.8 mm. (Solid curves are objective data obtained by measuring the light distributions formed outside the eye from an intense line stimulus imaged on the retina; data from Campbell & Gubisch, 1966. Dashed curves are theoretical diffraction-limited line-spread functions for artificial pupils placed just in front of the eye. All of the line-spread functions have been normalized to the same arbitrary area.)

sizes. These data (which are described in more detail later) represent the combined effect of all the optical sources of blur. All of the retinal intensity distributions in Figure 2 (including the theoretical diffraction-limited functions) have been normalized to the same area. It is apparent from Figure 2 that for foveal stimuli and pupil sizes of 2 mm or less, the blurring effects of diffraction exceed the combined effects of all the aberrations.

Scattered light. Another factor affecting image formation is the scattering of light by various structures in the eye. It is useful to distinguish between two types of scattered light: (a) light diffusely scattered by preretinal structures (e.g., by the cornea and lens) and by reflection off the back of the retina and (b) light scattered more locally in its forward passage through the outer layers of the retina (there may also be a localized component of scattered light reflected from the back of the retina).

The diffuse component of scatter fills the entire eye with a veil of light that is easily seen by viewing an intense light source against a totally dark background. Psychophysical measurements in young eyes indicate that the effective luminance of the diffusely scattered light from a point source is inversely proportional to the square of the distance (in degrees) to the source (Stiles, 1929; also see Vos, 1984; and Wyszecki & Stiles, 1982). This relation is meaningful only for angular separations greater than around 0.5° because the psychophysical estimates of scatter become more questionable as the likelihood of lateral neural effects from the point source increases. The psychophysical measurements indicate that about 10% of the light arriving at the retina is diffusely scattered light. At first this may seem to be a rather large percentage, but it is really rather small given the large area over which the light is scattered.

There are few useful physiological measurements of diffuse scatter in the human eye, largely because it is difficult to make and interpret measurements made on excised eyes (DeMott & Boynton, 1958). (However, for animal eyes, the fiber-optic technique of Robson & Enroth-Cugell, 1978, appears to work well.) Using a retinal densitometer on human subjects, Rushton and Gubisch (1966) measured the amount of photopigment bleached inside an intense annulus of light and compared this with the amount bleached by uniform fields. Using Equation 4 (7.8.1) in Wyszecki and Stiles (1982), one finds that Rushton and Gubisch's estimate of the scattered light under their conditions agrees quite well with calculations based on the psychophysical estimates.

The diffuse component of scattered light is small enough that its effects are only significant for detection or discrimination of relatively dim targets in the presence of intense glare sources. These situations do arise in the environment but not in the psychophysical experiments of direct interest here.

The localized component of scatter is more difficult to measure. Ohzu and Enoch (1972) attempted to measure the optical transfer function of the human foveal region of the retina immediately following removal of the eye during surgery. In principle, this transfer function should reflect the combined effects of forward scatter and spatial integration over the collection area of the photoreceptors. Ohzu and Enoch found that grating contrast is reduced by around 70% over the spatial frequency range of 30 to 60 c/deg (cycles per degree). D. R. Williams (1985b) argued that Ohzu and Enoch's measurement technique is likely to have overestimated the amount of scatter. Furthermore, using an elegant psychophysical technique, MacLeod, Williams, and Makous (1985) obtained evidence that forward retinal scatter is in fact minimal in the fovea. They did this by measuring threshold for beat-frequency patterns created by tilting interference-fringe components that, individually, were above the resolution limit. They found that threshold for a fixed beat frequency as a function of the component frequency was predicted accurately by the receptor aperture alone. Substantial forward retinal scatter would have produced a steeper fall off in threshold at high frequencies than was observed.

Point-spread function. Through linear-systems analysis, the combined effects of the various aberrations, pupil diffraction, and scatter on image formation can be neatly and usefully described with a point-spread function (or equivalently, with an optical transfer function, which is the Fourier transform of the point-spread function). For linear shift-invariant systems (i.e., ones that produce the same output shape regardless of input position), the point-spread function can be used to compute the output for an arbitrary input. (Point-spread functions, transfer functions, and their use in linear systems analysis are briefly described in the Appendix A.) However, aberrations, diffraction, and scatter each vary with eccentricity, pupil size, wavelength composition, and the state of accommodation. Thus, image formation in the human eye can only be accurately described by a family of point-spread functions. There are two general approaches for estimating point-spread functions. One is direct measurement; the other is calculation from measurements of the aberrations and other parameters of the human optical system.

When an image is formed on the retina, some of the diffusely reflected light passes back through the optics and forms an im-

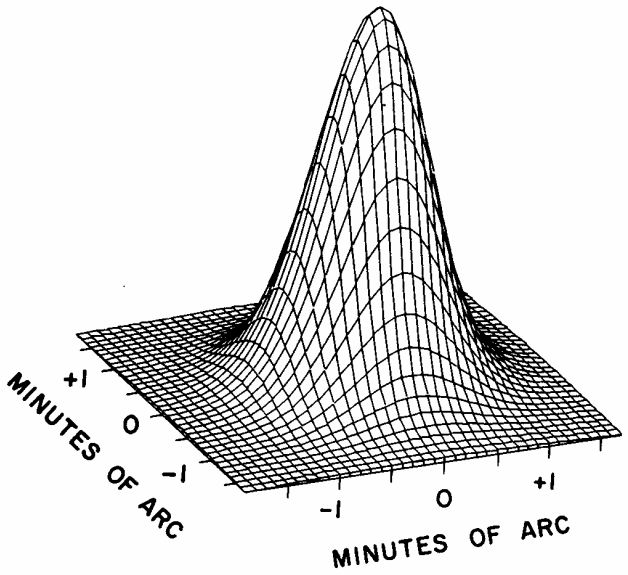


Figure 3. Optical point-spread function for a 2-mm pupil, calculated from the appropriate line-spread function in Figure 2, assuming radially symmetric optics.

age outside the eye. Were the intensity distribution of this image measurable, it would be possible to obtain an estimate of the actual retinal intensity distribution by correcting for the double passage through the optics. This is the principle behind the fundus reflection technique for directly measuring the point-spread function (Campbell & Gubisch, 1966; Flamant, 1955; Jennings & Charman, 1981; Westheimer & Campbell, 1962). The obvious target to use would be a point source; however, so little light is reflected out of the eye that more extended targets are always used. The most common is a very thin, bright line. With this target, one obtains an estimate of the line-spread function. For a linear shift-invariant system, the point-spread function can be computed from line-spread functions measured at many different orientations (e.g., see Gaskill, 1978). Furthermore, if the optical system is radially symmetric about the line of sight (which is approximately true for foveal stimuli), then the point-spread function can be computed from a single line-spread function. For example, Figure 3 shows the point-spread function derived from the line-spread data in Figure 2, for the 2-mm pupil.

A psychophysical method of measuring the point-spread function is based on laser interferometry (Arnulf & Dupuy, 1960; Campbell & Green, 1965; Westheimer, 1960). When two spatially separated, coherent point images are formed in the plane of the pupil, the addition and cancelation of the resulting wavefronts forms a sinusoidal grating pattern (like Figure A1 in Appendix A) on the retina. The spatial frequency (the number of stripes per degree of visual angle) of the grating is controlled by varying the separation of the point images.

The retinal contrast of interference gratings is unaffected by the monochromatic aberrations and pupil diffraction (e.g., see Saleh, 1982). Therefore, the ratio of the contrast thresholds for interference gratings and normally imaged monochromatic gratings, measured as a function of spatial frequency, gives an estimate of the modulation-transfer function (MTF; see Appendix A) of the ocular media, excluding the components that are due to chromatic aberration and scatter. Chromatic aberrations are excluded because the laser light is monochromatic. Scatter is excluded because it is much the same for normal and interference gratings (Vos, 1963) and thus will cancel out in the ratio. Assuming radial symmetry about the line of sight, the point-spread function (minus the effects of scatter) can be obtained from the estimated MTF.

The point-spread functions obtained by interferometry are somewhat narrower than those obtained by fundus reflection. This is explained in part because, as was mentioned earlier, interferometry excludes the effects of chromatic aberration and scatter. Campbell and Gubisch (1966) demonstrated that chromatic aberration could account for many of the differences between their fundus-reflection data and the interferometry measurements of Campbell and Green (1965). A possible factor contributing to the differences between the two methods is laser speckle in the interferometer, which serves as a potent masker of sine-wave gratings (D. R. Williams, 1985b). Laser speckle reduces the differences between the coherent and incoherent thresholds, thereby leading to an overestimate of the quality of the optics. On the basis of Williams's measurements, laser speckle probably only affected the results of Campbell and Green (1965) and Arnulf and Dupuy (1960) in the range of 40 to 60 c/deg because both studies varied the contrast of the target gratings by adding a background of incoherent light, and only at the highest spatial frequencies was the target component small relative to the background.

A number of investigators have attempted to compute foveal point- or line-spread functions from physical measurements of the aberrations and other optical parameters of the human eye. The calculations of van Meeteren (1974) included estimates of chromatic aberration, pupil diffraction, spherical aberration, irregular aberrations, and Stiles-Crawford apodization (see the section ahead on image sampling by receptors). Estimates of the comatic (i.e., third order) aberrations were unavailable at that time, and so were excluded from the calculations. Similarly, Navarro, Santamaría, and Bescós (1985) calculated line-spread functions for a schematic eye constrained by essentially the same factors that were considered by van Meeteren. The line-spread functions derived by van Meeteren and Navarro et al. are somewhat narrower than those measured by Campbell and Gubisch (1966). van Meeteren suggested that the difference might be accounted for by retinal scatter. Indeed, he showed that Ohzu and Enoch's (1972) estimate of forward retinal scatter could account for the differences. However, as mentioned earlier, Ohzu and Enoch most likely overestimated the effects of forward retinal scatter.

Another factor that might explain the differences are the comatic (third order) aberrations. An example of comatic aberration is the comet-shaped blur pattern that can be seen when one tries to create an image (e.g., of the sun) with a lens that is not perpendicular to the line connecting the object and the image. Walsh, Charman, and Howland, (1984) attempted to measure simultaneously all of the monochromatic aberrations with an interesting new technique, and found that both spherical- and comalike aberrations were substantial in normal young eyes. In

calculating line-spread functions for the conditions of Campbell and Gubisch (1966), Walsh et al. included their own measurements, plus the other factors considered by van Meeteren, except, rather curiously, chromatic aberration. If the effects of chromatic aberration are added into their calculations, one obtains somewhat broader line-spread functions than those measured by Campbell and Gubisch. However, this discrepancy might be explained because Walsh et al. also included Ohzu and Enoch's (1972) estimate of retinal scatter.

In sum, there seems to be a consensus that for white light the Campbell and Gubisch (1966) foveal line-spread functions are close to correct, although there is some disagreement about which factors contribute most significantly to defocus. van Meeteren (1974) argued that for white light, pupil diffraction and chromatic aberration are the main factors contributing to the line-spread functions, and that there are negligible individual differences in these two factors. Other factors of some importance are the spherical and comalike aberrations, diffuse scatter, and perhaps the Stiles-Crawford apodization (although probably not as much in the fovea, see later). Forward retinal scatter is probably not a major factor.

It appears that for a well-accommodated eye, one can use with reasonable confidence the line-spread functions of Campbell and Gubisch (1966) to compute foveal intensity distributions for broad-band, achromatic stimuli. Because of individual differences in the monochromatic aberrations, calculations are less certain (in relative terms) for narrow-band or monochromatic stimuli. Nonetheless, representative line-spread functions for narrow-bandwidth light can be obtained by correcting the Campbell and Gubisch line-spread functions for chromatic aberration. For stimulus conditions with glare sources, one must add a term representing the diffuse scatter.

The only substantial study measuring image quality in the peripheral retina was carried out by Jennings and Charman (1981). Using the fundus-reflection technique, they found that image quality was uniformly high for the region within 12° of the optical axis and deteriorated at greater eccentricities. The usefulness of their data for computing retinal intensity distributions was greatly reduced because apparatus limitations required that all measurements be obtained with a fully dilated (7.5-mm) pupil. It is possible that the increased aberrations that occurred with the large pupil size may have masked substantial changes in the line-spread function that occur over the central 12° with more natural pupil sizes. Nonetheless, Jennings and Charman's data do provide a lower bound on image quality in the periphery for the well-accommodated eye.

Ocular transmittance. Not all light striking the cornea reaches the retina. Some of it is reflected off the refracting surfaces of the eye, and some is absorbed in the ocular media. Near the visual axis, the most important factors are absorption in the lens and macular pigment (a yellow pigment that is most concentrated over the fovea). In the periphery, macular pigmentation decreases, leaving lens absorption as the primary factor. There are individual differences in macular-pigment density, and the density of the lens increases with age; nonetheless, the shapes of the absorption curves are fairly similar across individuals and ages. The combined effect of all optical transmittance factors in the average young eye is shown in Figure 4. This figure gives the fraction of photons transmitted to the foveal receptors as a function of wavelength. Note that the short wavelengths are the ones most severely attenuated. A summary and evaluation of the data concerning preretinal absorption is given in Wyszecki and Stiles (1982).

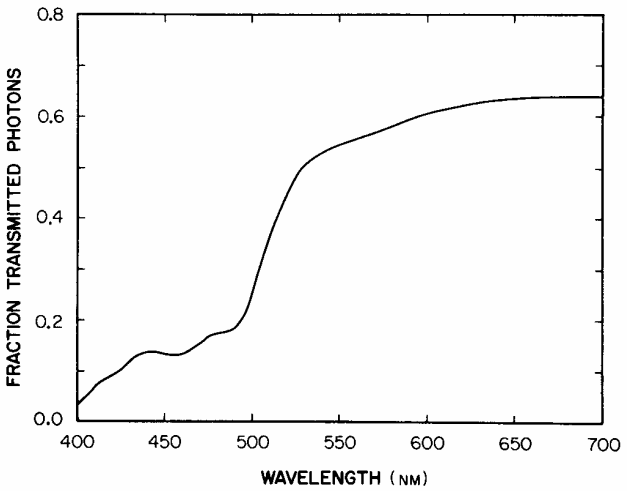


Figure 4. Average optical transmittance function in the fovea for young eyes. (Data from Wyszecki & Stiles, 1982.)

Pupil response. The size of the pupil is strongly affected by the luminance of the visual field and to a lesser extent by a variety of other factors, including the size and position of the visual stimuli, the state of accommodation, and the convergence of the eyes. Changes in pupil size affect both the intensity and the optical quality of the retinal image. The retinal intensity produced by a point source is proportional to the apparent area of the pupil when the pupil is viewed from the direction of the source. The pupil also affects depth of focus, the range of distances around the point of accommodation over which an object remains in good focus. Also, recall from Figure 2 that the size of the pupil affects the point-spread function even when the eye is in optimal focus.

In addition to all of the factors affecting pupil size, there are large individual differences in the size of the pupil response. However, it is not difficult to measure or control pupil size, and some normative data are available for estimating pupil size for those studies in which pupil size was neither controlled nor measured. Figure 5, which was derived from Wyszecki and Stiles (1982), gives average pupil size as a function of luminance for large stimulus fields. In sum, pupil size is an important factor limiting visual discrimination; its effects are well understood and easily included in a sequential-mechanisms analysis.

Accommodation. Another important factor affecting image formation is the accommodation response. Objects that are not in the plane of accommodation produce blurred images. Even under well-controlled psychophysical conditions, subjects generally will not achieve optimal focusing. The subject's accuracy varies with the form of the target stimuli (Charman & Tucker, 1977; Heath, 1956), the size of the pupil, and the distance of the target from the end points of the subject's accommodation range (Charman & Tucker, 1977; M. W. Morgan, 1944). In addition, the accommodation response undergoes continuous mi-

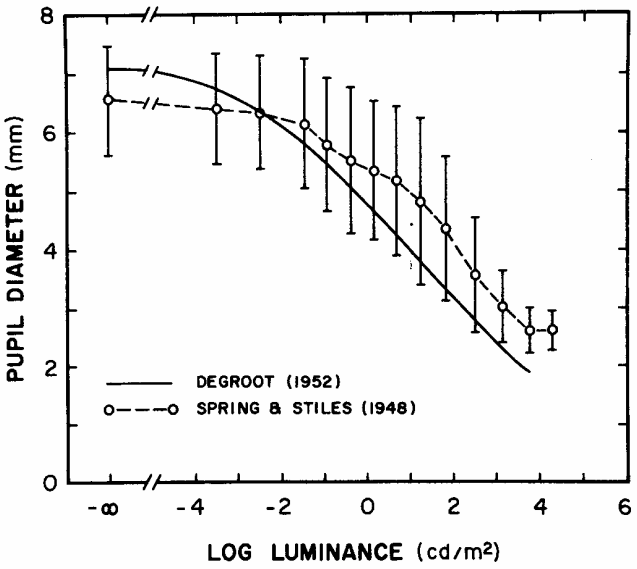


Figure 5. Average pupil diameter as a function of background luminance for large field stimuli. (Solid curve is the average of several studies computed by deGroot & Gebhard, 1952. Circles are the average of 12 observers from Spring & Stiles, 1948. Error bars indicate ±1 *SD* computed across subjects.)

crofluctuations at a rate of 2 to 4 Hz. These fluctuations are quite small (approximately 0.1 diopters) and hence should result in a small amount of image degradation. It has been suggested that the purpose of the fluctuations may be to provide information to the accommodation system about when and in what direction to respond (Campbell, Robson, & Westheimer, 1959).

The effects of the accommodation errors on image quality (i.e., on the point-spread function) are complicated—they depend on pupil size and on the magnitudes of the optical aberrations (Campbell & Green, 1965). Focusing errors due to accommodation can be essentially eliminated by paralyzing the accommodation mechanism with cycloplegic eye drops and placing appropriate lenses in front of the eye. However, only a small fraction of psychophysical experiments use this procedure. When considering other experiments, one must consider the probable effects of focusing errors on image formation on a study-by-study basis.

Eye movements. When the eyes move, the retina slides around underneath the retinal image. As mentioned earlier, the stimulus conditions of primary interest here are those with well-controlled fixation. However, even under steady fixation, the eyes are constantly making slow drifts and small correcting saccades (Ditchburn, 1973; Rattle, 1969; Steinman, 1965). Nachmias (1959), Krauskopf, Cornsweet, and Riggs (1960), Steinman and others showed that the standard deviation of eye position under steady fixation is around 3′ to 5′ (3–5 min of arc) in the typical subject for fixation trials of 30 s. This is quite accurate considering that a foveal cone is about 0.5′ in diameter. This level of accuracy holds for many different fixation targets. Steinman had subjects fixate the centers of circular targets and found that fixation was equally accurate for diameters ranging from 2′ to 87′. Rattle (1969) found that subjects were nearly as accurate in fixating the midpoint of two dots as in fixating the center of a circle and, also, that there was little variation in fixation accuracy for circle diameters or interdot distances ranging from 19′ to 230′.

The design of the visual system beyond the level of the receptors is such that residual eye movements are essential to the transmission and processing of visual information—without eye movements, a human soon becomes effectively blind (Ditchburn & Ginsborg, 1952; Riggs, Ratliff, Cornsweet, & Cornsweet, 1953). However, at the level of the receptors, eye movements result in two potential forms of information loss. First, the temporal integration of the photoreceptors results in an effective spatial smearing of the visual image. This factor is not relevant here because (a) the present ideal-observer analysis only includes mechanisms up to the level of photon absorption, and (b) most of the experiments to be analyzed used brief stimulus presentations. Second, if the visual system does not keep precise information about the positions of the eyes, then some information loss will result from position uncertainty. The experimental conditions of most interest here are those providing good fixation targets. It is not known whether the visual system monitors variations in eye position within the 3′ to 5′ range of eye position occurring during steady fixation. Thus, the only sensible strategy is to carry out ideal-observer analyses with and without the uncertainty factor included. As it turns out, ideal-observer calculations show that a position uncertainty of 3′ to 5′ produces only modest losses of information in detection tasks and essentially no information loss in most discrimination tasks (Geisler & Davila, 1985; also see ahead).

Image Sampling by Receptors

The retinal image is sampled spatially and spectrally by a lattice of discrete photoreceptors. There are four major classes of photoreceptor (the rods and three classes of cone) that differ in spectral sensitivity and distribution within the lattice. The image sampling properties of the receptor lattice result in further losses of information useful for discrimination.

Receptor aperture. Light passes through all of the layers of the retina and most parts of the receptor—the synaptic termination, cell nucleus, and inner segment—before reaching the light-sensitive photopigment in the outer segment. Because the receptors collect light over a finite spatial region (the receptor aperture), they perform some spatial integration. This phenomenon, known as the *aperture effect,* produces an effective blurring of the retinal image.

The rods are essentially cylindrical in shape with a diameter of around 1 micron, which translates to an aperture of about 12″ (12 s of arc). On the other hand, the cones are tapered from the inner segment through the outer segment, reaching their smallest diameter at the end of the outer segment. Anatomical evidence (Sidman, 1957) indicates that the refractive index of the inner segment (in rods and cones) differs sufficiently from that of the interstitial medium that the receptors should behave like optical wave guides. If so, much of the light entering the inner segment would be funneled into the outer segment, and the diameter of the cone aperture would equal the diameter of the inner segment at its maximum.

This conclusion must be qualified, especially in the fovea where the cone inner segments are tightly packed together. The change in the refractive index at the inner segment is probably not sufficient for wave-guide behavior until the inner segment has begun to taper (Miller & Bernard, 1983; Yamada, 1969). On the basis of Miller's (1979) phase-contrast microscopy, Miller and Bernard found that the change in refractive index of foveal cones becomes substantial enough to identify individual inner segments when they have tapered to about 2.3 microns (80% of maximum) or 28″. They also cited Byram's (1944) psychophysical measurements of resolution for interference fringes as support for a foveal cone aperture of 2.3 microns. Miller and Bernard noted that because interference fringes are unaffected by the optics of the eye, the only optical limit for resolution is that set by the receptor aperture. The MTF for such an aperture is given by a first-order Bessel function of the first kind. For an aperture of 2.3 microns, this MTF first reaches a value of zero at 150 c/deg, which is near the upper limit of resolution reported by Byram. An aperture larger than 2.3 microns would produce a zero in the MTF at a lower frequency. D. R. Williams (1985b) reached a similar conclusion on the basis of his measurements of interference-fringe resolution. However, these calculations must be viewed with some reservation because they assume a single fixed aperture diameter for all the receptors under the grating. A zero in the MTF would not, in general, occur if there are variations in receptor diameter.

The inner segment diameters of the cones increase substantially with eccentricity, whereas the rod diameters remain essentially the same. Thus, the cone aperture effect increases with eccentricity, whereas the rod aperture effect remains approximately constant. One can reasonably estimate the size of the cone aperture effect at any eccentricity by assuming an aperture diameter of 60% to 80% of the inner segment diameter. (The aperture may be a somewhat larger percentage of inner segment diameter in the periphery because the inner segments are not packed as tightly as in the fovea.)

The previous discussion gives the impression that having a finite receptor aperture is primarily detrimental. However, the aperture also determines how much light a receptor can absorb. All other things being equal, the more light absorbed, the more discrimination information transmitted. This is because photon noise is described by the Poisson process; thus, as we shall see later, the signal-to-noise ratio for a discrimination increases as quantum catch goes up. It should also be added that receptor-aperture blur is small compared with that produced by the optics.

The Stiles-Crawford effect. Photopic (cone) sensitivity is greatest for light that enters the eye near the center of the pupil and decreases monotonically as the entry point moves toward the edge of the pupil (Stiles & Crawford, 1933). For example, parafoveal cone thresholds for light entering the eye 3 mm from the center of the pupil are 4 times higher than in the center. This phenomenon, known as the Stiles-Crawford effect of the first kind (SCE1), probably results from the wave-guide properties of the cones—their shape and index of refraction gives them directional selectivity (Stiles & Crawford, 1933; Snyder & Pask, 1973; Westheimer, 1967a). Wave-guide theory also correctly predicts that cones in the central fovea and rods should display less SCE1 than do peripheral cones because they are more cylindrical in shape (Snyder & Pask, 1973; Westheimer, 1967a).

The SCE1 has two important effects on the transmission of discrimination information through the eye. First, there is a reduction in the number of photons reaching the cone photopigments. The following equation (LeGrand, 1957; Wysecki & Stiles, 1982) gives the fraction of photons reaching the photopigments (due to the SCE1) as a function of pupil radius:

$$\sigma = (1 - 10^{-0.05p^2})/0.115p^2. \quad (2)$$

This equation is appropriate for foveal stimulation where the cones are more cylindrical. For small pupils (1–3 mm in diameter), σ is greater than 0.9, so the Stiles-Crawford effect can be ignored.

The second consequence of the SCE1 is an effective improvement in image quality for large pupil sizes. Because light near the edge of the pupil is attenuated, the SCE1 is essentially equivalent to reducing the pupil diameter. This is called Stiles-Crawford *apodization.* Campbell and Gubisch (1966; see Figure 3) showed that image quality is optimal for pupil diameters around 2 to 3 mm. Thus, for larger pupils, a reduction in effective pupil size will improve image quality. Most of the stimulus conditions considered here were for pupil diameters in the range of 2 to 3 mm, so the SCE1 plays a relatively minor role in determining image quality. Furthermore, the diffuse, polarized component of light reflected off the back of the retina displays the Stiles-Crawford effect (Röhler, Miller, & Aberl, 1969). Thus, the Campbell and Gubisch line-spread functions may already include the effects of Stiles-Crawford apodization. No adjustments of the Campbell and Gubisch functions for Stiles-Crawford apodization were included in the present analysis.

In addition to the directional effect on sensitivity, there is also a directional effect on color appearance, the Stiles-Crawford effect of the second kind (e.g., see Enoch & Stiles, 1961). However, it is a minor effect that was not included in the present ideal-observer analysis.

Photopigment absorptance spectra. Not all light entering a photoreceptor is effectively absorbed by the photopigment. The fraction of photons absorbed is dependent on the concentration (c) of the photopigment, the length (l) of the outer segment, and the photopigment's extinction coefficient $\epsilon(\lambda)$. In particular, the proportion of incident photons absorbed as a function of wavelength (the absorptance spectrum) is given by the Beer-Lambert law,

$$a(\lambda) = 1 - e^{-\epsilon(\lambda)cl} \quad (3)$$

(e.g., see Hsia, 1965).

The most direct measurements of the absorptance spectra of the cone and rod photopigments have been made with microspectrophotometry (Bowmaker, Dartnall, Lythgoe, & Mollon, 1978; Brown & Wald, 1964; Dartnall, Bowmaker, & Mollon, 1983; Marks, Dobelle, & MacNichol, 1964). This technique involves passing light of different wavelengths through the outer segments of single excised photoreceptors. As a consequence, the measurements are relatively noisy.

If the univariance principle (see Footnote 2) were to hold, then more precise determinations of the absorptance spectra could be made. In particular, the absorptance spectrum of a

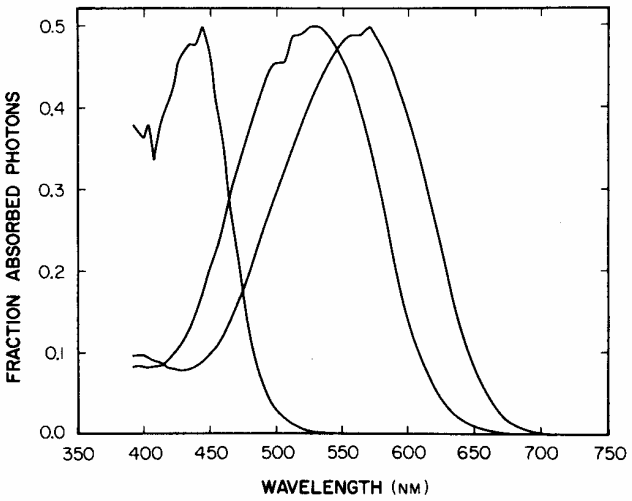


Figure 6. Absorptance spectra of the cone pigments after correction for ocular transmittance. (Data from Estevez, 1979, as reported in Wyszecki & Stiles, 1982.)

photopigment can be deduced by measuring an action spectrum—the intensity at each wavelength necessary to produce a criterion response from the photoreceptor. The relative absorptance spectrum is obtained by normalizing the reciprocal of the action spectrum. Using an elegant suction-electrode technique, Baylor, Nunn, and Schnapf (1984) and Nunn, Schnapf, and Baylor (1984) have verified the univariance principle and measured the action spectra of macaque rods and cones. The measured spectra agree reasonably well with the results of microspectrophotometry but are more precise and cover a wider range of wavelengths.

Long before the development of microspectrophotometry and suction electrodes, there were several attempts to estimate the cone absorption spectra from psychophysical color-matching data of normal and color-blind observers. In the color-matching experiment, the observer adjusts the intensities of three fixed-wavelength lights (primaries) so that their mixture appears identical to the test light. Observers can obtain such identity matches for arbitrary test lights and can do so with great precision and reliability (see Boynton, 1979, for a readable discussion of color-matching experiments and data). The relative absorptance spectra deduced by Estevez (1979; but see Wyszecki & Stiles, 1982) from color-matching data agree quite well with the spectra measured with the suction electrode technique by Nunn et al. (1984). Figure 6 shows the absorptance spectra of Estevez, adjusted to a reasonable peak absorptance of 0.5. It is highly likely that these are close to the true absorptance spectra of cones in the normal human retina.

The fact that there are only three cone photopigments, with broad overlapping spectra, implies that under some circumstances considerable information relevant for chromatic discriminations will be lost at the photoreceptors. Thus, the photopigment absorptance spectra represent an important mechanism limiting visual discrimination. The Estevez (1979) spectra in Figure 6 were used in the present ideal-observer analysis because they provide a good fit to the measurements of Nunn et al. (1984), although there is one exception, described later, in which the Smith and Pokorny (1975) spectra were used. For most of the discrimination tasks considered here, the Estevez and the Smith and Pokorny spectra give very similar results.

Dark noise. It has long been suggested that intrinsic noise in the visual system may be an important factor limiting visual sensitivity (Barlow, 1956, 1977). Barlow (1977) listed several potential sources of intrinsic noise: (a) thermal isomerization (spontaneous bleaching of photopigment molecules that produces receptor responses that are just like those produced by absorbed photons), (b) other receptor noise, (c) neural and synaptic noise, (d) impulse-quantitization error due to the spike-generation process of ganglion cells and central neurons, and (e) criterion fluctuations and other sources of central inefficiency. The only intrinsic noise source that might be appropriate to include in the present development is thermal isomerization (or an equivalent photonlike noise).

Using the suction-electrode technique mentioned earlier, Baylor et al. (1984) have measured the intrinsic-noise properties of macaque rods. They distinguish two types of receptor dark noise: a continuous low-amplitude noise and a discrete impulse noise that appears similar to the responses produced by single absorbed photons (photonlike noise). Baylor et al. argued that the continuous noise might be filtered out at the receptor synapse. If so, only the photonlike noise would limit visual performance. Their measurements indicate that the rate of these photonlike events is around 6×10^{-3} events/s/receptor. This noise level equals approximately the psychophysical estimates of the intrinsic noise of the whole rod system (Barlow, 1977), suggesting that it may be the main source of intrinsic noise. These photonlike events are most likely due to thermal isomerizations of the rod photopigment molecules; however, from an information-processing viewpoint their source is irrelevant, as long as the noise events are indistinguishable from true photon absorptions.

The Baylor et al. (1984) study provides a firm estimate of the photonlike dark-noise level in macaque rods, but there are no comparable estimates for macaque cones. Thus, at least for the cone system (which is the focus of the present article), it is premature to try to include dark noise in a rigorous ideal-observer analysis. However, some level of photonlike noise is certain to exist in cones if for no other reason than an occasional thermal isomerization. Thus, a small level of dark noise was included in generating some of the ideal-observer predictions in order to illustrate its qualitative effect. These instances will be noted in the text or in the figure captions.

Lattice geometry. Figure 7 shows cross-sections of the receptor layer in macaque monkey at eccentricities of 0°, 5°, and 10° (from Perry & Cowey, 1985). In the fovea, the cones form a tightly packed lattice that, over small regions, is nearly a perfect triangular array (Borwein, Borwein, Medeiros, & McGowan, 1980; Hirsch & Hylton, 1984; Miller, 1979; Perry & Cowey, 1985; Williams, 1988). As eccentricity is increased, the cones become larger and fewer in number, the regularity of the lattice declines, and rods begin to fill in spaces between the cones.

Figure 8 shows the number of cones per degree of visual angle as a function of eccentricity in humans (Osterberg, 1935) and in the macaque monkey (Perry & Cowey, 1985). The apparent differences between macaque and human receptor densities are

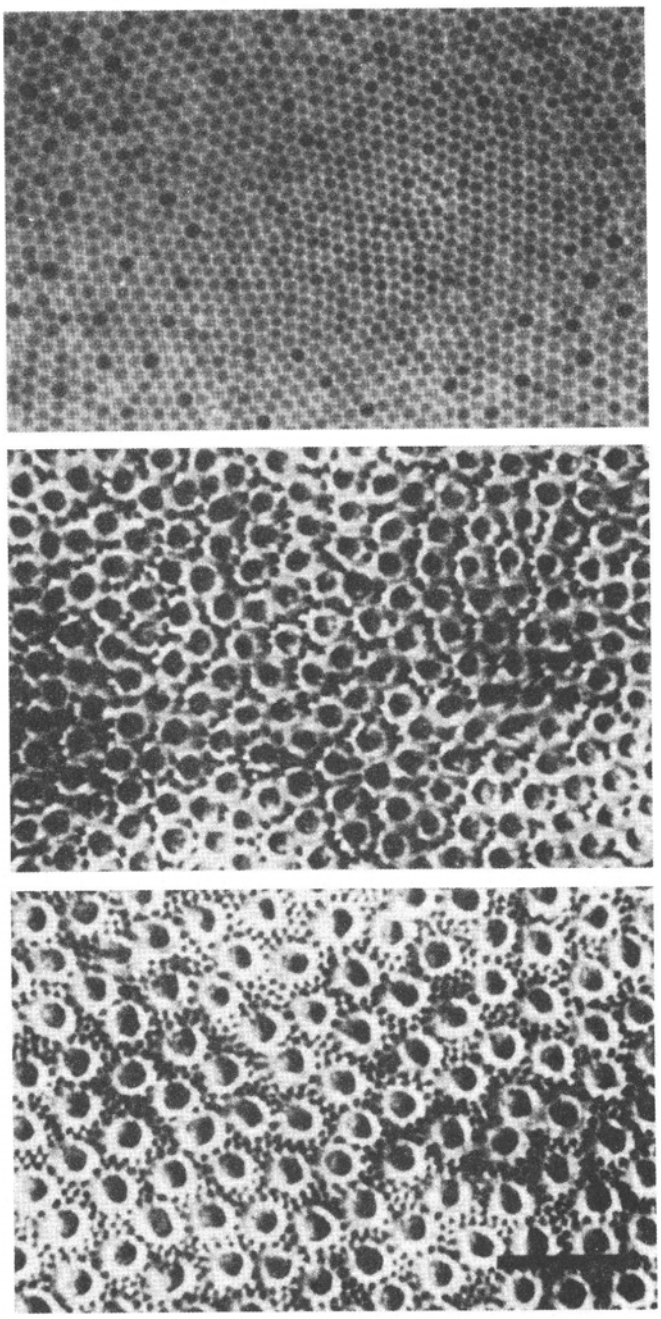


Figure 7. Microphotographs of the cone lattice in the macaque retina in the fovea (top), at 5° eccentricity (middle), and at 10° eccentricity (bottom). (Horizontal bar represents approximately 5 min of arc; from Perry & Cowey, 1985.)

due in part to the differences in eye size—they are more similar when expressed in receptors/mm^2. There is, however, some inherent uncertainty in the anatomical results because of unavoidable tissue shrinkage and other potential forms of histological artifact. Fortunately, Williams (1988) has recently developed an elegant (psychophysical) laser-interferometry technique for measuring the topography of the retinal cone mosaic in the living human eye, and has obtained results that are in good agreement with the anatomical studies.

Receptor sampling rate is a significant factor limiting the information available for visual discriminations. Classical sampling theory provides some insight into these limitations. The basic result is the Whittaker-Shannon sampling theorem (e.g., see Bracewell, 1978). The theorem implies that a retinal intensity distribution, whose Fourier transform does not contain spatial frequencies above a cutoff frequency of W c/deg, can be completely represented and (if desired) reconstructed by taking 2W or more regularly spaced samples per degree of visual angle (that is, $4W^2$ samples per square degree). This critical sampling rate is known as the *Nyquist rate.* Sampling below the Nyquist rate results in information loss because more than one intensity distribution could then give rise to the same sample values. This source of confusion is called *aliasing.* Sampling above the Nyquist rate is unnecessary, in the strict sense of the sampling theorem, because no additional information is gained.

It has been widely noted that the human foveal lattice is well matched to the optics of the eye (Miller & Bernard, 1983; Snyder & Miller, 1977; Yellott, 1984). Recall that the point-spread function in the fovea, for the optimal pupil size of 2 to 3 mm, has a cutoff frequency (W) of around 60 c/deg. In other words, retinal intensity distributions can never contain spatial frequencies above 60 c/deg. Therefore, a sampling rate of 120 samples/deg would be optimal for foveal images. Figure 8 shows that this is almost exactly the rate reached in the central fovea.

The close match between the optics and receptor density has

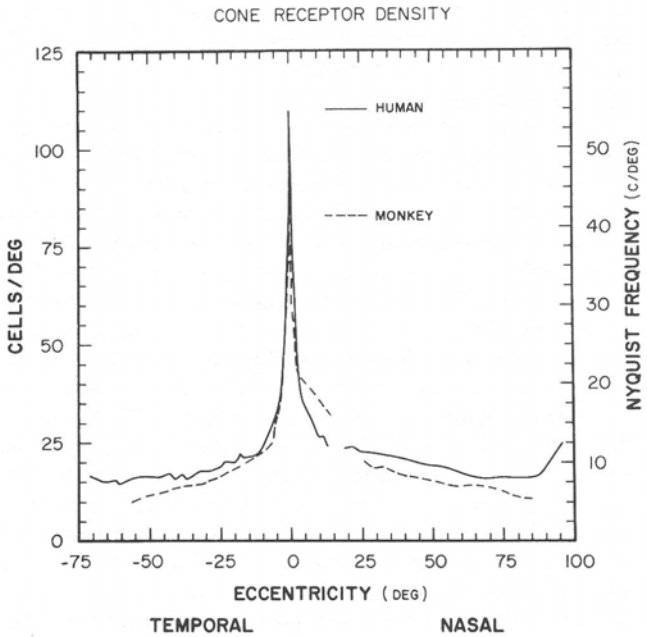


Figure 8. Density of cones in the human (solid curve) and macaque (dashed curve) retina as a function of eccentricity. (Left axis is the square root of the number of cells per square degree. Right axis shows the highest spatial frequency that could be reliably transmitted by the lattice according to the Whittaker-Shannon sampling theorem. Human data are from Osterberg, 1935; macaque data are from Perry & Cowey, 1985.)

been confirmed in the living human retina by Byram (1944) and D. R. Williams (1985a). They created interference fringes on the retina at frequencies above 60 c/deg. Their observers reported the appropriate perceptual side effects of aliasing and moire patterns at frequencies just above 60 c/deg.

In the periphery, the cone sampling rate appears to be below the Nyquist rate implied by the optics. For example, the point-spread functions measured by Jennings and Charman (1981) suggest that the optical quality of the eye is fairly uniform out to eccentricities of 12°, but as shown in Figure 8, the sampling rate drops precipitously over this range.

Deductions about information transmission through the retinal lattice that are based solely on the sampling theorem are incomplete because they do not consider the effects of image noise (Bossomaier, Snyder, & Hughes, 1985; Geisler & Hamilton, 1986). In addition, sampling theorems have so far been proved only for regular sampling lattices. Indeed, there is currently some debate about whether the irregularity of the receptor lattice is an advantage or a disadvantage from an information-transmission viewpoint (Hirsch & Hylton, 1984; Maloney, 1988; Miller & Bernard, 1983; Snyder et al., 1986; Yellott, 1984). However, as will be shown, the ideal-observer analysis developed here provides a complete and precise method for evaluating the effects of lattice structure on specific visual discriminations.

There is reasonably good agreement among the anatomical studies that have measured receptor size and density and lattice regularity. Thus, these aspects of the lattice structure can be incorporated with some confidence into an ideal-observer analysis. Specifically, lattices such as those in Figure 7 can be directly encoded and loaded into the computer program carrying out the analysis. For the fovea, I have found that idealized triangular lattices (Figure 9) with the appropriate receptor diameter behave equivalently to the real lattices in all discrimination tasks I have examined so far.

Unfortunately, there is less solid anatomical and physiological evidence concerning the relative numbers and distributions of the three classes of cone within the receptor lattice. By selectively staining the blue cones, De Monasterio, McCrane, Newlander, and Schein (1985) showed, in agreement with earlier psychophysics (D. R. Williams, MacLeod, & Hayhoe, 1981) and anatomy (Marc & Sperling, 1977), that the blue-sensitive (B) cones are absent in the central 10′ to 20′ of the fovea and make up only a small percentage (2%–10%) of the cones elsewhere in the retina. Psychophysics and cell counts from microspectrophotometry indicate that there are more red-sensitive (R) cones than green-sensitive (G) cones (Ingling & Martinez-Uriegas, 1983). However, the ratio could conceivably be anywhere within the range of 2:1 to 1:1. Figure 9 shows an idealized, but not unreasonable, lattice that was proposed by Walraven (1974). In this lattice, the ratio of R to G to B cones is 32:16:1. An appropriate lattice for the central fovea would be similar but would only contain R and G cones.

For many stimulus conditions of interest, the stimuli are chromatically broad band (e.g., white); hence, they stimulate the R and G cones about equally well (and the sparse B cones don't contribute significant information for the discrimination). Under these circumstances, the receptor lattice can be regarded as a homogeneous array and the relative numbers of the cone types becomes irrelevant. For other stimulus conditions, it is necessary to vary the ratios (and spatial arrangement) of the cone types over the plausible ranges in order to determine confidence intervals on the amount of information transmitted.

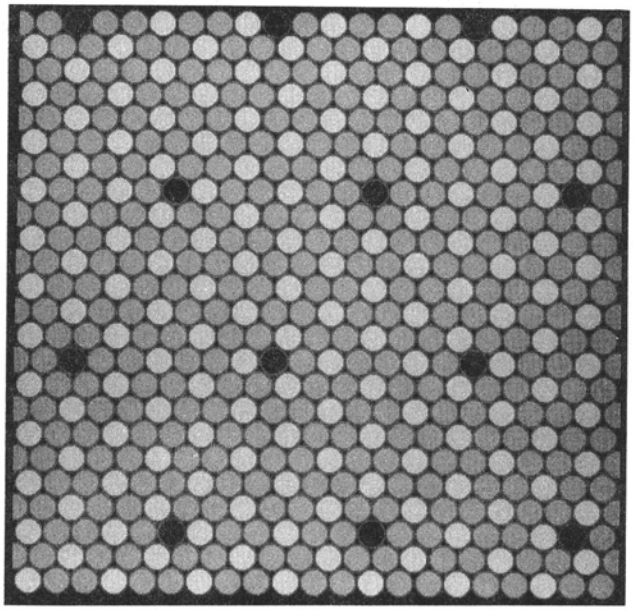


Figure 9. Idealized receptor lattice used in the ideal-observer analysis. (The brightest circles are the middle-wavelength [green] receptors, the medium-bright circles are the long-wavelength [red] receptors, and the dark circles are the short-wavelength [blue] receptors. The red, green and blue receptors are in a 32:16:1 ratio; Walraven, 1974.)

Ideal–Observer Model

The preceding sections have summarized and evaluated our current knowledge of the factors up to the level of photon capture in the receptors that are likely to affect visual discriminations. These factors are now well-enough understood at the quantitative level to warrant a rigorous analysis of their effects on visual discrimination. In this section, the ideal-observer analysis illustrated in Figure 1 is developed for the ideal observer placed at the level of photon absorption in the receptors (at the output of the last solid box in Figure 1).

Model of the Stimuli

In the two-alternative discrimination experiment, there are two nominal stimuli, a and b, that the subject must discriminate. We presume that the stimuli are presented for a fixed duration, d, in either a single-alternative or a two-alternative forced-choice procedure. We also presume, with little loss of generality, that the two stimuli are generated by a display device that creates the stimuli by mixing three primaries (e.g., the red, green, and blue "guns" of a color TV monitor). Then, the arbitrary stimulus can be described by a spectral distribution (amount of light emitted as a function of wavelength) and a spatial intensity distribution for each primary. Let the spectral dis-

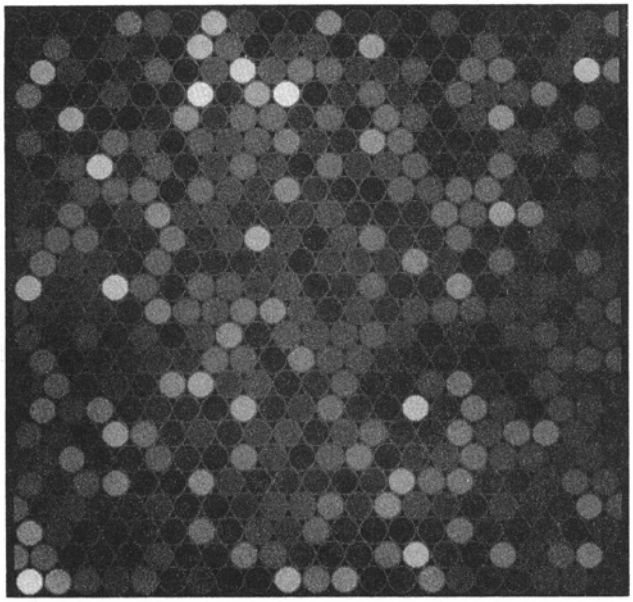


Figure 10. Illustration of the Poisson randomness of light. (Typical pattern of photon absorptions produced in the photoreceptors by a uniform flash of 1×10^6 quanta/deg^2 [530 nm] at the cornea. The brightness of a photoreceptor represents the relative number of photons absorbed.)

tributions of the three primaries, normalized to an area of one quantum, be $n_r(\lambda)$, $n_g(\lambda)$, and $n_b(\lambda)$. Let the spatial intensity distributions for Stimulus a be $r(x, y)$, $g(x, y)$ and $b(x, y)$, and for Stimulus b be $r'(x, y)$, $g'(x, y)$ and $b'(x, y)$. (For modeling purposes, the most convenient units for these functions are quanta/s/deg^2 at the cornea per mm^2 of pupil.) For the special case when the visual stimuli are monochrome and broadband (e.g., generated on a single-phosphor TV monitor), it is convenient to define a single spectral distribution, $n_w(\lambda)$, and spatial intensity distributions, $w(x, y)$ and $w'(x, y)$.

Recall that even under well-controlled psychophysical conditions, the inherent Poisson randomness of light implies that the same nominal stimulus will give rise to a different physical distribution of photons on each presentation. For much of the analysis ahead, we assume that the sole source of stimulus variability is that produced by the Poisson randomness of light. Figure 10 illustrates Poisson noise—it shows the pattern of photon absorptions produced by a single presentation of a uniform green field at an intensity 10^6 quanta/deg^2 at the receptor layer. The brightness of each receptor indicates the number of absorbed photons.

The assumption that Poisson noise is the only noise source in the stimulus is the simplest, and it holds under most well-controlled psychophysical and electrophysiological studies. However, it is worthwhile to first develop the ideal-observer analysis for more general and natural stimulus conditions. The theory for the conditions with only Poisson noise will fall out as a special case.

When performing binary classification tasks in the real world, the stimuli often randomly vary along many different dimensions. For example, an inspector on an assembly line may have to classify parts as defective or not defective. The parts within the two categories may vary in size, shape, position, orientation, or color. Even if the parts are all the same, and the defect is only in one particular feature, the parts may still vary in position and orientation. To describe these more complicated stimulus conditions, we define a vector, $\mathbf{v} = \langle v_1, v_2 \cdots, v_m \rangle$, of uncertainty parameters, and a pair of uncertainty functions, $f(\mathbf{v}|\mathrm{a})$ and $f(\mathbf{v}|\mathrm{b})$. The uncertainty functions give, for each nominal stimulus, the joint probability of obtaining any particular value of $v_1, \ldots, v_m$ on a trial. In this general case, we represent the corneal intensity distributions as $r(x, y, \mathbf{v})$, $g(x, y, \mathbf{v})$, $b(x, y, \mathbf{v})$, and $r'(x, y, \mathbf{v})$, $g'(x, y, \mathbf{v})$, $b'(x, y, \mathbf{v})$. To precisely specify the stimuli in this more general case, it is necessary to know the uncertainty functions and the corneal intensity distributions for every possible uncertainty vector.

Model of Image Formation and Sampling

If an ideal observer is to be placed at the level of photon capture in the photoreceptors, we need to compute the probability of absorbing any given number of photons in any given receptor on a trial. To do this, we must compute the mean number of photons absorbed in each receptor for each of the two stimuli (Geisler, 1984). Let $a_i(\mathbf{v})$ and $b_i(\mathbf{v})$ be the mean number of photons absorbed in the ith receptor of the lattice to Stimuli a and b, respectively, given a particular value of the uncertainty vector $\mathbf{v}$. Computation of $a_i(\mathbf{v})$ and $b_i(\mathbf{v})$ requires inclusion of all the mechanisms and factors described in the first part of this article.

The first thing to note is that the number of absorbed photons will depend on whether the receptor is red-, green-, or blue-sensitive. If we let (x_i, y_i) represent the coordinates (in degrees of visual angle) of the ith receptor, then

$$a_i(\mathbf{v}) = \begin{cases} R(x_i, y_i, \mathbf{v}), & \text{if } i\text{th receptor is red sensitive} \\ G(x_i, y_i, \mathbf{v}), & \text{if } i\text{th receptor is green sensitive} \\ B(x_i, y_i, \mathbf{v}), & \text{if } i\text{th receptor is blue sensitive} \end{cases}$$

$$b_i(\mathbf{v}) = \begin{cases} R'(x_i, y_i, \mathbf{v}), & \text{if } i\text{th receptor is red sensitive} \\ G'(x_i, y_i, \mathbf{v}), & \text{if } i\text{th receptor is green sensitive} \\ B'(x_i, y_i, \mathbf{v}), & \text{if } i\text{th receptor is blue sensitive} \end{cases} \quad (4)$$

where, R, G, B, R', G', and B' are the mean numbers of photons absorbed in each of the cone classes to each of the two stimuli given a particular uncertainty vector $\mathbf{v}$.

Next, note that the number of photons absorbed in a particular receptor is the sum of the numbers absorbed from each primary:

$$R(x_i, y_i, \mathbf{v}) = R_r(x_i, y_i, \mathbf{v}) + R_g(x_i, y_i, \mathbf{v}) + R_b(x_i, y_i, \mathbf{v}) + de_R$$

$$G(x_i, y_i, \mathbf{v}) = G_r(x_i, y_i, \mathbf{v}) + G_g(x_i, y_i, \mathbf{v}) + G_b(x_i, y_i, \mathbf{v}) + de_G$$

$$B(x_i, y_i, \mathbf{v}) = B_r(x_i, y_i, \mathbf{v}) + B_g(x_i, y_i, \mathbf{v}) + B_b(x_i, y_i, \mathbf{v}) + de_B$$

$$R'(x_i, y_i, \mathbf{v}) = R'_r(x_i, y_i, \mathbf{v}) + R'_g(x_i, y_i, \mathbf{v}) + R'_b(x_i, y_i, \mathbf{v}) + de_R$$

$$G'(x_i, y_i, \mathbf{v}) = G'_r(x_i, y_i, \mathbf{v}) + G'_g(x_i, y_i, \mathbf{v}) + G'_b(x_i, y_i, \mathbf{v}) + de_G$$

$$B'(x_i, y_i, \mathbf{v}) = B'_r(x_i, y_i, \mathbf{v}) + B'_g(x_i, y_i, \mathbf{v}) + B'_b(x_i, y_i, \mathbf{v}) + de_B \quad (5)$$

where the subscripts r, g, and b indicate the red, green, and blue primaries, d is the duration of the stimulus, and e_R, e_G, and e_B are the intensities of the photonlike dark noise in events/s/receptor. For those occasions in which dark noise was included in generating predictions, it was assumed that the dark noise was the same in all three cone classes: $e_R = e_G = e_B$.

The problem is now reduced to computing the number of photons absorbed in a particular receptor from a particular primary. To do this, note first that the intensity distribution for a primary as a function of position and wavelength is obtained by multiplying the spatial intensity distribution by the spectral distribution (e.g., $r(x, y, \mathbf{v})n_r(\lambda)$). Proceeding in the logical sequence, one begins by multiplying the input function by the duration of the stimulus, d, and then by the area of the pupil, s, and the ocular transmittance function, $t(\lambda)$. The resulting function is then convolved with the point-spread function for the primary (e.g., $h_r(x, y)$). (To ensure that no quanta are added or lost, the point-spread function is always normalized to a volume of 1.0.) The result of this convolution is the *retinal-intensity distribution.* Next, the retinal intensity distribution is convolved with the receptor-aperture function, $k(x, y)$. (The aperture function is a cylinder with a height of 1.0 and a diameter equal to the receptor aperture; thus, its volume equals the area of the aperture.) The function resulting from this convolution is then multiplied by the Stiles-Crawford correction factor, σ, and the absorptance spectrum of the receptor photopigment (e.g., $a_R(\lambda)$). Finally, the entire result is integrated with respect to wavelength.

For example, photon absorption in the R cones for the r primary is given by the following:

$$R_r(x, y, \mathbf{v}) = sd\sigma r(x, y, \mathbf{v}) ** h_r(x, y) ** k(x, y) \int n_r(\lambda)t(\lambda)a_R(\lambda)d\lambda, \quad (6)$$

where ** represents the operation of two-dimensional convolution (see Appendix A). Note that the terms in the equation have been rearranged to put together the quantities that depend on wavelength. If the ith receptor is an R cone, Equation 6 can then be evaluated at the position of the ith receptor to obtain $R_r(x_i, y_i, \mathbf{v})$. Similar equations are used to compute photon absorption for all combinations of cone types and primaries. If the stimulus is monochrome and broad band, then the point-spread function $h_w(x, y)$ is used for all cone types.

When evaluating Equations 4 to 6 to generate predictions, I have tried to use the best estimates of the physiological parameters and functions available in the literature reviewed earlier. If pupil size was not reported for an experiment under consideration, it was estimated using the average data of Wyszecki and Stiles (1982; Figure 5 here). The point-spread functions (e.g., Figure 3) were derived from the line-spread functions of Campbell and Gubisch (1966; Figure 2 here). Unless reported otherwise, accommodation was assumed to be near optimal. If it was not near optimal, the point-spread functions were appropriately modified. Ocular transmittance in the fovea was taken from Wyszecki and Stiles (1982; Figure 4 here). For peripheral stimulation, in which the macular pigment is absent, only the lens transmittance function was used. At any given eccentricity, the receptor aperture was taken to be 80% of the average inner segment diameter. The Stiles-Crawford correction factor was calculated from Stiles's formula (Equation 3). The receptor absorptance spectra were those of Estevez (1979), assuming an absorptance of 0.5 at the peak wavelength (see Figure 6). For most of the predictions, the lattice at a given eccentricity was taken to be a perfectly regular triangular array (Figure 9), with the center-to-center distance between receptors equal to the average at that eccentricity. To check on the influence of lattice irregularity, I also directly used the macaque lattices of Perry and Cowey (1985; Figure 7 here). In the central fovea it was assumed that the B cones were absent; elsewhere it was assumed that they constituted 2% of the cones (Figure 9). For most predictions it was assumed that the R cones were twice as numerous as the G cones—however, the ratio was varied for some experiments in order to determine its effect.

Ideal Discriminator

The last step in developing the ideal-observer analysis is to derive the ideal discriminator. On each trial of the experiment, the ideal discriminator must decide which of the two alternatives (a or b) was presented. It is well-known (see Green & Swets, 1974) that the optimal discriminator bases its decision on the likelihood ratio or some monotonic transformation of it. The intuition behind the use of the likelihood ratio is the following. On each trial of the experiment, there is some pattern of photon absorptions across the photoreceptors that can be represented by the random vector $\langle Z_1, Z_2, \ldots, Z_n \rangle$, where Z_i is the number of photons absorbed in the ith receptor. The ideal discriminator first computes the probability of getting this pattern of absorptions assuming that Stimulus a was presented, and then assuming that Stimulus b was presented. It then simply picks the more likely alternative. Equivalently, the discriminator computes the ratio of the probabilities (the likelihood ratio); if the result exceeds 1.0, one of the alternatives is picked, otherwise the other alternative is picked. (This assumes that the alternatives are equally likely and the discriminator is trying to maximize the proportion of correct responses. If the alternatives are not equally likely, a criterion other than 1.0 is used; Green & Swets, 1974.)

It can be shown (Geisler & Davila, 1985) that under the most general conditions, an ideal discriminator (when placed at the level of photon capture in the photoreceptors) uses the following decision variable:

$$L = \frac{\int_{\mathbf{v}} \exp\{\sum Z_i \ln[b_i(\mathbf{v})] - \sum b_i(\mathbf{v})\} \cdot f(\mathbf{v}|\mathrm{b})d\mathbf{v}}{\int_{\mathbf{v}} \exp\{\sum Z_i \ln[a_i(\mathbf{v})] - \sum a_i(\mathbf{v})\} \cdot f(\mathbf{v}|\mathrm{a})d\mathbf{v}}, \quad (7)$$

where ln is the natural logarithm, and the integral is over all the possible values of the uncertainty vector. If L exceeds some constant c, then Stimulus b is picked, otherwise Stimulus a is picked. Note that the quantities in the exponents are approxi-

mately the cross correlations (Gaskill, 1978) of the stimulus (the Z_is) with the natural logarithm of the expected stimulus shape, minus the volume under the expected shape, $\sum b_i(\mathbf{v})$. In other words, the likelihood of a stimulus alternative is computed by exponentiating the cross correlation of the signal with the logarithm of the stimulus shape (minus its volume), multiplying by the uncertainty function, and summing over the uncertainty space. The derivation of the aforementioned decision rule is straightforward and is similar to the derivation for the case of Gaussian noise given in standard texts on statistical decision theory (e.g., Van Trees, 1968). The performance of the ideal discriminator for this general case can only be calculated using Monte Carlo simulations. Geisler and Davila (1985) have carried out some of these calculations for the case of position uncertainty (i.e., the stimulus randomly appears in different positions on each trial). As might be expected, such calculations are difficult and time consuming.

Our primary concern here, however, is with stimulus conditions in which the only source of noise or uncertainty is that due to quantal fluctuations (photon noise). In this case, the maximum-likelihood decision rule is equivalent to the following rule (Geisler, 1984; Helstrom, 1964). Compute the quantity Z by the formula

$$Z = \sum Z_i \ln(b_i/a_i), \tag{8}$$

where b_i is the mean number of photons absorbed in the ith receptor to Stimulus b, and a_i is the mean number of photons absorbed in the ith receptor to Stimulus a. If Z exceeds a criterion c, then respond that the stimulus was b, otherwise that it was a.

Notice that Z is a weighted sum of the photon catch in each photoreceptor. Thus, the ideal decision rule can be implemented by creating a single weighting function (or receptive field) for the particular pair of stimuli to be discriminated. We will call this the *ideal receptive field*. Consider, for example, two-point resolution—a classic discrimination task that was of great interest to early astronomers and psychophysicists. Figure 11A shows the average pattern of photon absorptions produced in the foveal photoreceptors by two perfectly superimposed green point sources of light. Figure 11B shows the pattern of absorptions when the two point sources are separated by 60 s of arc (approximately the two-point resolution threshold for humans). The brightness of each receptor indicates the mean number of photons absorbed on a trial. These photon absorptions were computed using Equations 4, 5, and 6 for a 2-mm pupil size. Because the figures are in gray tone, the different classes of cone are not distinguishable. However, because the stimuli are green in the present case, the G receptors can be picked out because they appear somewhat brighter than the R receptors. (Because the stimuli are in the central fovea, there are no B cones.) On each trial of the experiment, the human or ideal observer must decide whether the stimulus was a sample from the pattern in Figure 11A or a sample from the pattern in Figure 11B. (Recall that on each trial the same stimulus will be randomly different due to photon noise; see Figure 10.) The ideal receptive field for making this discrimination is shown in Figure 11C. The receptors that are brighter than the background are weighted positively and can be thought of as excitatory; the receptors that are darker than the background are weighted negatively and can be thought of as inhibitory; the receptors that are at the background gray level make no contribution to the decision. Thus, a single postreceptor unit that sums the receptor outputs with a receptive field like that in Figure 11C will behave as an ideal discriminator for this two-point resolution task. Notice that it is not possible to pick out the R and G receptors in the ideal receptive field. This is because the weight assigned to each receptor depends only on the ratio of the photon catches from the two stimuli (see Equation 8). The weights vary with the cone types only when the task contains some component of chromatic discrimination (see Figure 25).

In general, for every different discrimination there is a different ideal receptive field (although they do not change when the stimulus intensity distributions are scaled by a multiplicative factor). It follows that some discriminations require ideal receptive fields that are similar to those that have been found in the visual pathway and others require ones that are very dissimilar. An intriguing hypothesis is that human discrimination tends to be most efficient in those discrimination tasks for which the ideal receptive fields are similar to the actual receptive fields found early in the visual pathway. This and other aspects of ideal receptive fields will be discussed later in the article.

Ideal-observer models that apply to conditions where Equation 8 is appropriate will be referred to here as Stimulus-Defined-Exactly (SDE) models. For the more general cases, where Equation 7 holds, they are referred to as Stimulus-Defined-Statistically (SDS) models (Geisler & Davila, 1985).

The predictions of an SDE model can be computed without resorting to Monte Carlo simulations. Equations 4, 5, and 6 are evaluated to compute the mean number of photons absorbed in each receptor (the a_is and b_is) for both stimuli. Performance accuracy can then be obtained using the following closed-form expression for d' (Geisler, 1984):

$$d' = \frac{\sum (b_i - a_i) \ln(b_i/a_i)}{[\sum (b_i + a_i) \ln^2(b_i/a_i)]^{1/2}}. \tag{9}$$

Equation 9 gives d' for the two-alternative forced-choice task. If the stimulus presentation probabilities are equal for the two intervals, and the ideal observer is trying to maximize percentage correct, then each value of d' corresponds to a unique percentage correct (e.g., see Green & Swets, 1974). In this article, the main interest is in predicting discrimination thresholds. In most experiments, threshold is defined to be that difference between the stimuli that results in 75% correct discrimination performance. This occurs when $d' = 1.36$.

The derivation of Equation 9 assumes that the sum of the scaled Poisson random variables in Equation 8 is approximately normal. Monte Carlo simulations with Equation 8 show that Equation 9 is very accurate even at low intensity levels.

Predictions of the more general SDS models can be computed by Monte Carlo simulation. In particular, a percentage-correct psychometric function is determined as a function of the difference in the stimuli along the relevant dimension. Threshold is estimated from the 75% correct point on the psychometric function (which corresponds to a d' of 1.36).

Calculations for the ideal-observer analysis were carried out on a VAX computer, using VAX FORTRAN and the IMSL

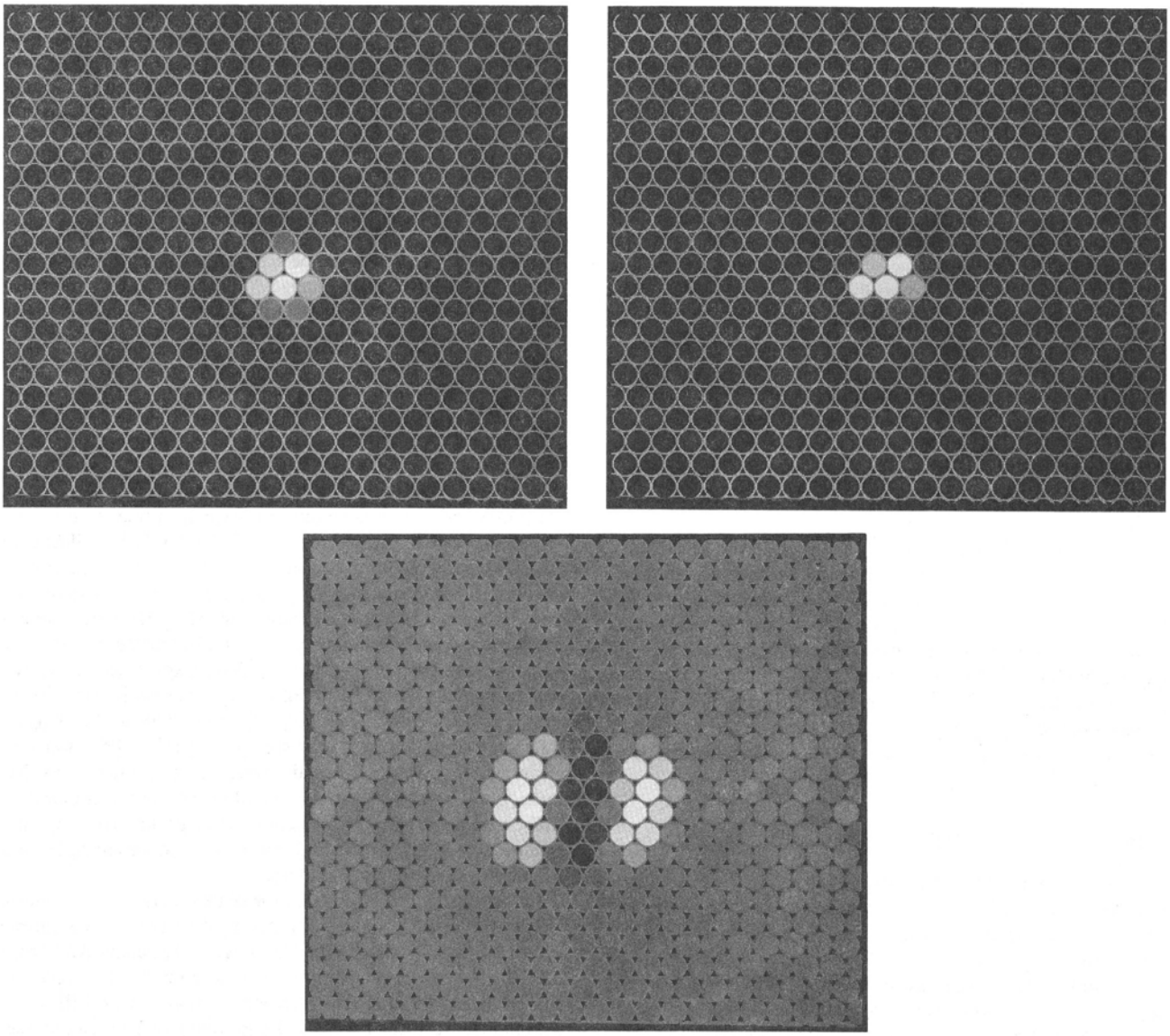


Figure 11. Two-point resolution. (A. Average pattern of photon absorptions in the receptors produced by two superimposed green point sources. B. Average pattern of absorptions produced by two point sources separated by 50 s of arc. C. Ideal receptive field for discriminating the patterns in A and B.)

(Houston, TX 77036) mathematical and statistical library. A brief discussion of how the actual computations were carried out is given in Appendix B.

Visual-Discrimination Data and Predictions

This section describes applications of the SDE ideal-observer model developed earlier to a wide variety of visual discrimination tasks. It is important to keep in mind that the purpose of the ideal-observer model is not to obtain quantitative fits to psychophysical data by estimating parameters. Indeed, there are no free parameters except those used to fit smooth curves through some of the physiological data for purposes of interpolation. Rather, the purposes are (a) to determine the contributions of the preneural mechanisms to various visual discriminations, (b) to determine the discrimination information available at the level of the photoreceptors, and hence, by comparison with human observers, the relative efficiency of the postreceptor mechanisms across the various tasks, and (c) to evaluate current models of visual discrimination in light of the ideal-observer analysis.

Because of the wide range of discrimination tasks to be considered, it is impossible to do justice to the recent models, hypotheses, and experimental research described in the vision literature. When possible, general references to this literature are given.

Ultimately, all visual discrimination tasks can be rightly regarded as discriminations of the distribution of photons across the retina over time. However, for historical and practical reasons it is useful to categorize the discriminations as follows: intensity and contrast discriminations, position and shape discriminations, chromatic discriminations, and temporal discriminations (or some combination of these). We use these particular categories to partition the discrimination literature into smaller chunks. However, such categories can be misleading by drawing attention away from common mechanisms. For example, it is typical to distinguish between shape and position discrimination. This is rather arbitrary because one can regard changes in the positions of objects as a change in the shape of the image as a whole, and conversely, one can often regard a change in the shape of a single object as changes in the positions of features within the object. From the standpoint of ideal-observer analysis, this problem is largely irrelevant because it treats all discriminations equivalently and because the purpose of the analysis is, in fact, to determine how physiological mechanisms contribute to performance across this wide range of tasks.

The following discussion will be confined to the variables of intensity, contrast, position, shape, and wavelength composition. Temporal variables (such as stimulus duration, adaptation time, and flicker frequency) and temporal measurements (such as response latency) will not be considered because the preneural mechanisms contribute little to the temporal-processing properties of the visual system. Unless stated otherwise, the predictions of the SDE model presented below are for foveal presentation in a well-accommodated eye, with a 2- or 3-mm pupil diameter.

Comparing Real and Ideal Performance

In what follows, the performance of the ideal observer will be compared with that of real observers. It is a foregone conclusion that the ideal observer is more sensitive than the human observer and, hence, that the ideal-observer model will not provide a good fit to the psychophysical data. In order to visualize what aspects of human performance might be accounted for by the factors incorporated into the ideal observer, it is useful to compare the shapes of measured threshold functions with the predicted functions. The method adopted here is to reduce the sensitivity of the ideal observer uniformly until the predicted and observed threshold functions are sufficiently overlapping for easy shape comparison. Sensitivity is reduced, in effect, by placing a neutral density filter (e.g., a pair of neutral-colored sunglasses) at the input to the ideal-observer model. This method of comparing shapes will be referred to as the *relative-efficiency method* because it is closely related to the concept of *quantum efficiency* (Barlow, 1962, 1977).[4] The quantum efficiency of a real observer for a particular stimulus is defined as the percentage transmittance of a neutral density filter (placed in front of the ideal observer's eye) required to equate the real and ideal thresholds. Thus, the difference between quantum and relative efficiency is that one uses an ideal observer placed at the cornea and the other uses one placed at some level in the visual system.

In the following examples, the value of the neutral density filter used to adjust ideal performance was picked rather arbitrarily to bring the real and ideal curves close enough together for easy comparison of shapes. When possible, the value of the neutral density filter (or the ratio of real-to-ideal performance) is reported in the figure caption. However, in some cases the duration and/or retinal illumination of the stimuli was not specified well enough in the original report for an accurate calculation of relative efficiency.

Intensity and Contrast Discrimination

In order to see, one must be able to encode and reliably discriminate local intensity changes in space (and time). Hence, intensity and contrast discrimination have been studied vigorously since the beginnings of vision science. In the intensity or contrast-discrimination experiment, the shapes, positions, and wavelength compositions of the two test patterns remain fixed, whereas the intensities or contrasts are varied.[5]

Sine-wave grating detection. Over the past 20 years, much effort has been directed toward measuring the detectability and discriminability of sine-wave grating stimuli. Sine-wave grating stimuli are of considerable interest because of their central role in linear-systems analysis (see Appendix A). It is well-known from Fourier's theorem that any (monochrome) image can be decomposed into a sum of sine-wave gratings varying in amplitude, frequency, phase, and orientation. Thus, if a visual system is approximately linear, its behavior can be characterized by its responses to the set of sine-wave gratings. Experience has shown that even when a visual system is not behaving linearly its response to the sine-wave gratings often provides useful insights into its general behavior. The sine-wave grating detection task is given a somewhat more detailed treatment here than are other tasks, because it has proved to be of fundamental importance in the measurement and prediction visual performance and because it provides a good initial example of the reasoning behind sequential ideal-observer analysis.

In the grating detection task, the observer must discriminate between a uniform field and a patch of grating. (An example sine-wave grating is shown in Figure A1 in Appendix A.) Figure 12A shows the pattern of photon absorptions produced in the central fovea by a uniform green background. Figure 12B shows the pattern produced by a 20 c/deg grating patch with a contrast of 1.0. Figure 12C shows the ideal receptive field for the discrimination.

The contrast-sensitivity functions (CSFs) of the ideal-discriminator are shown in Figure 13. The vertical axis is contrast sensitivity (the reciprocal of contrast threshold) on a log scale.

---

[4] An alternate definition of efficiency is Tanner and Birdsall's (1958) η, which is the square of the ratio of d'-ideal to d'-real (also see Barlow, 1978). Examination of Equation 9 shows that η is exactly the transmittance of the neutral density filter required to equate real and ideal performance. Thus, the present definition of efficiency is identical to Tanner and Birdsall's (1958).

[5] The usual definition of the contrast (C) within some region is

$$C = (I_{max} - I_{min})/(I_{max} + I_{min}),$$

where I_{max} and I_{min} are the maximum and minimum intensities within the region. Thus, if I_{max} and I_{min} are the same, the contrast is zero; if I_{min} is zero, the contrast is one.

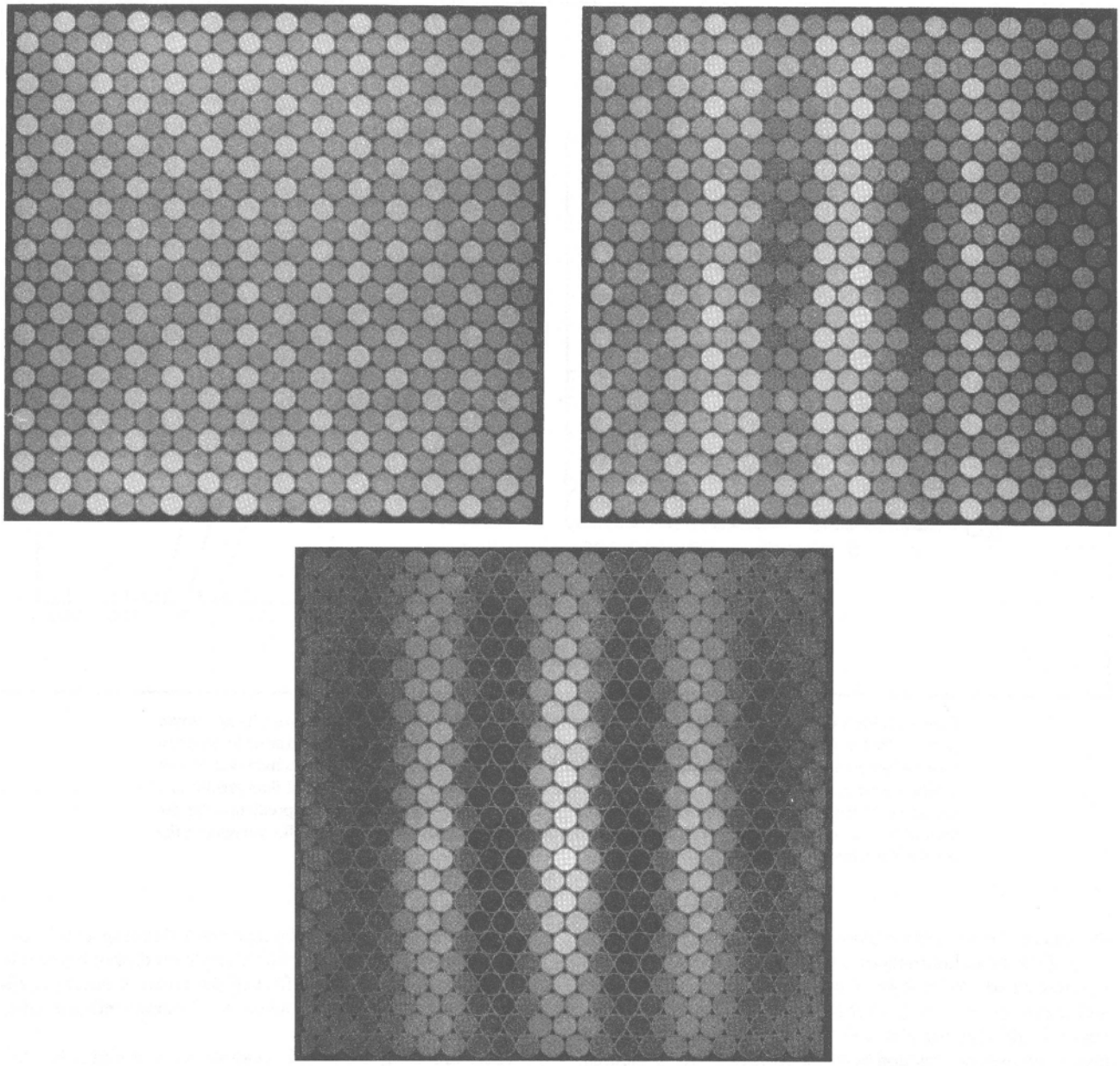


Figure 12. Luminance grating detection. (A. Average pattern of photon absorptions in the receptors produced by a uniform green background. B. Average pattern of absorptions produced by a 100% contrast, Gaussian-damped, sine-wave grating patch with a frequency of 20 c/deg [cycles per degree]. C. Ideal receptive field for discriminating the patterns in A and B.)

Thus, a contrast sensitivity of 10 represents a contrast threshold of 0.1. The horizontal axis is spatial frequency on a log scale. These predictions are for circular grating patches with a fixed number of cycles; thus, the patch size decreases with increasing spatial frequency. It is sensible to consider gratings with a fixed number of cycles because they match the cycle-summation properties of the visual system (see ahead). (The predictions are for patches with 7 cycles, but the predicted shapes are not very dependent on the number of cycles.)

The solid curves in Figure 13A show the sensitivity functions of the ideal discriminator at three mean luminances, for a perfect optical system and arbitrarily small and densely packed photoreceptors. These curves thus represent the absolute physical limit of contrast sensitivity for grating patches with a fixed number of cycles. The solid curves in Figure 13B show the CSFs with all the preneural factors included.

The curves in Figure 13A illustrate two important properties of ideal observers limited only by quantal noise. First, note that

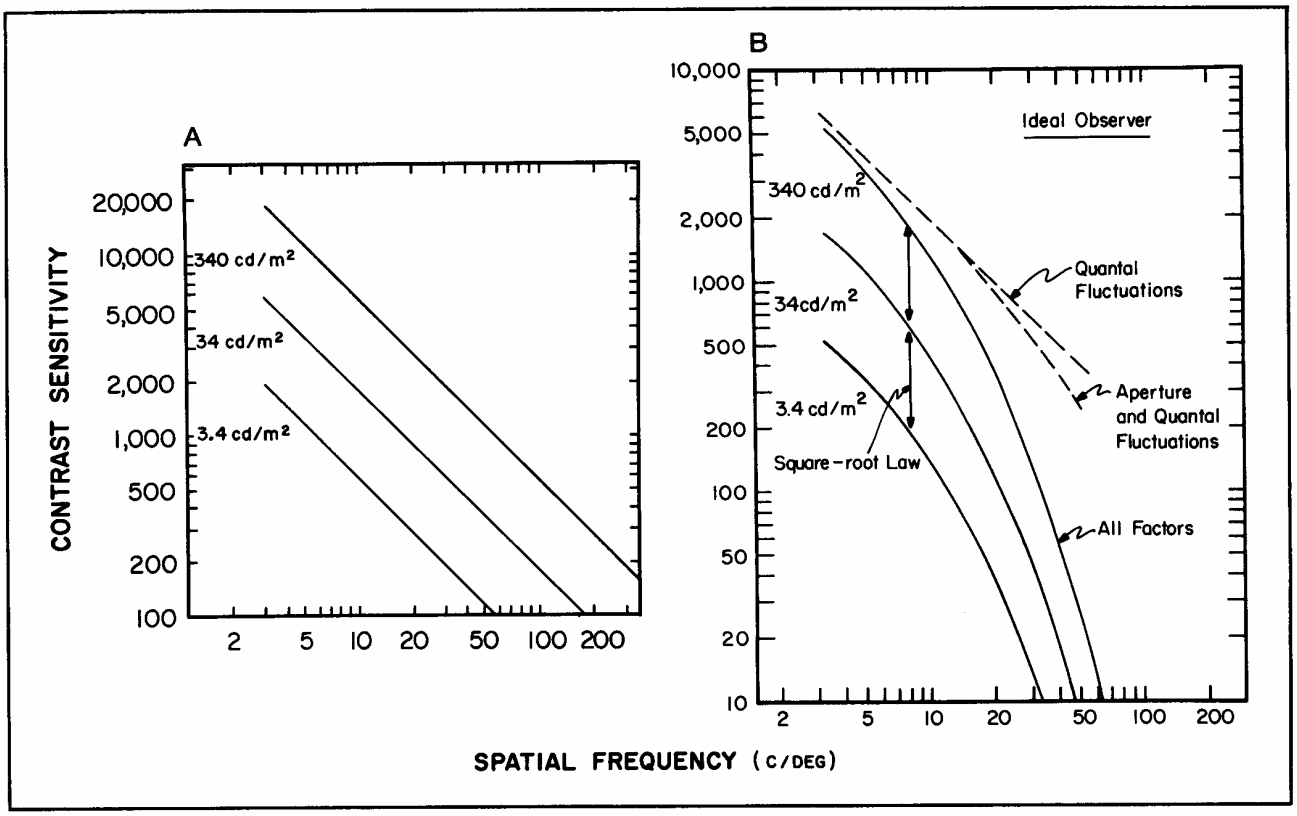


Figure 13. Ideal-observer predictions of contrast sensitivity for white sine-wave gratings with a fixed number of cycles (in this case, 7 cycles). (A. Predicted contrast-sensitivity functions at three background levels when the ideal observer is placed at the cornea. B. Predictions with various physiological factors included. Upper dashed curve is the prediction for the highest background luminance with only quantal fluctuations and the transmittance of the ocular media; cf. top curve in A. Lower dashed curve is the prediction for the highest background luminance with the additional effect of the receptor aperture. Solid curves are the predictions when all factors are included.)

the spacing between the curves equals the square root of the ratio of the mean luminances. This is an example of the classic square-root (de Vries-Rose) relation between threshold, ΔN, and mean luminance, N. Second, note that the slopes of the curves in this log–log plot are -1. This is an example of the inverse square-root relation between threshold and target area, A (Barlow, 1958b). These two properties are contained in the following relation that is due to de Vries (1943):

$$\Delta N = K_s(N/A)^{1/2}.$$

This relation holds independently of stimulus shape, which only affects the proportionality constant, K_s. Note that the effects shown in Figure 13A represent differences in the information content of the stimuli at the cornea because no physiological mechanisms are included.

The upper dashed line in Figure 13B shows the predictions for the highest mean luminance of Figure 13A, when the only factors included in the analysis are quantal fluctuations, the transmittance of the ocular media, and the photopigment absorptance spectra. As can be seen by comparing this curve with the upper solid curve in Figure 13A, the effect of the ocular transmittance function and the cone absorptance spectra is simply to translate the CSF downward. The lower dashed line shows the result of including the effect of the receptor aperture. Finally, the upper solid curve shows the combined effect of all of the factors.

There are many published measurements of CSFs, but few were conducted with sufficient stimulus control for rigorous comparison with ideal-observer models. Thus, Banks et al. (1987) carefully measured foveal CSFs for gratings with a fixed number of cycles, in a two-alternative temporal forced-choice procedure. The gratings were damped vertically and horizontally to prevent spatial transients and were presented for 100 ms to ensure that eye movements could not be initiated during presentation. Focusing of the eye was controlled by viewing the display through a 2-mm artificial pupil after dilating the pupil and paralyzing accommodation with 1% cyclopentolate. Optimal focusing was then obtained by placing appropriate lenses in front of the eye.

The open symbols in the upper panels of Figure 14 show the measured contrast sensitivities of two subjects, from 5 to 40 c/deg, at three mean luminance levels. The solid curves represent

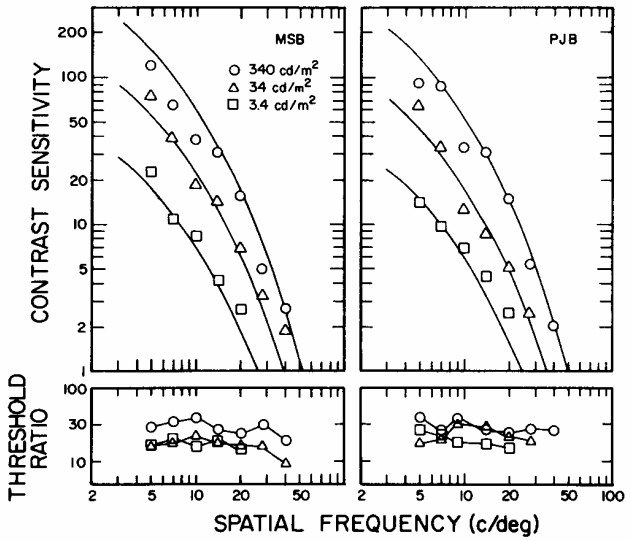


Figure 14. Comparison of contrast-sensitivity functions at high spatial frequencies for human and ideal observers. (Symbols are the contrast-sensitivity functions of two human observers measured at three background luminances, for damped sine-wave gratings with a fixed number of cycles. Solid curves are predictions of the ideal-observer model shifted by the relative-efficiency method for the purposes of comparing shapes. Specifically, the predicted curves were translated vertically 1.4 log units; neutral density value = 2.8 log units. Lower panels show the ratio of real to ideal threshold.)

the performance of the ideal discriminator after translating the curves vertically by a fixed amount, without changing their relative positions. (This method of comparing shapes corresponds to the relative-efficiency method described earlier.) As can be seen, the shapes of the predicted and observed CSFs are almost identical. The similarity of the shapes is illustrated further in the bottom panels, which show the ratios of real to ideal thresholds.

There are two important conclusions one can draw immediately from the Banks et al. (1987) results. First, in agreement with earlier studies (Kelly, 1972; Van Nes & Bouman, 1967), the square-root law holds over the whole midphotopic range for spatial frequencies above 5 c/deg. Second, and more surprising, the shapes of the CSFs are accurately predicted by the preneural mechanisms, as specified in the ideal-observer model, plus the differences in the information content of the stimuli at the cornea (Figure 13A). This result suggests that the efficiency of visual processing beyond the photoreceptors (for detection of fixed-cycle grating patches) is roughly constant over the whole range of mean luminances and spatial frequencies tested. In other words, under these conditions, there is no relative attenuation of high spatial frequencies by the neural mechanisms in the visual system—the neural-transfer function (NTF) is flat. (Note that NTF is defined here as the ratio of real-to-ideal sensitivity as a function of spatial frequency.)

The NTF found by Banks et al. (1987) seems, at first thought, to differ from earlier estimates of the NTF obtained by comparing the optical transfer function (Fourier transform of the point-spread function) with large-field CSFs, or by measuring contrast sensitivity for large-field gratings created directly on the retina with laser interferometry (Campbell & Green, 1965; Kelly, 1977; Westheimer, 1960; D. R. Williams, 1985b).[6] From these earlier estimates of the NTF, it has been concluded that neural mechanisms strongly attenuate high spatial frequencies (Snyder & Srinivasan, 1979).

The results of Banks et al. (1987) are, in fact, consistent with this conclusion. The difference in outcomes is almost surely explained by the fact that the earlier studies obtained CSFs with gratings of large, *fixed* spatial extent. Specifically, Banks et al. suggested that the difference in the results can be explained by the effects of neural spatial summation across cycles of the sine-wave grating targets. Several studies have shown that the detectability of sine-wave gratings increases with the number of cycles up to some critical number (Howell & Hess, 1978; Koenderink, Bouman, Bueno de Mesquita, & Slappendel, 1978; Robson & Graham, 1981). For medium and high spatial frequencies, this critical number of cycles is constant, regardless of spatial frequency, and is of the same extent both parallel and perpendicular to the orientation of the grating (Howell & Hess, 1978). This property of the visual system is correlated with the fact that cortical neurons tuned to higher spatial frequencies tend to have smaller spatial extents even over the same part of the visual field (De Valois, Albrecht, & Thorell, 1982; Movshon, Thompson, & Tolhurst, 1978).

In order to minimize the effect of neural summation across cycles on the computed NTF, Banks et al. (1987) deliberately measured contrast sensitivities with grating patches of a fixed number of cycles. However, they would have reached a different conclusion had they measured CSFs with gratings of large fixed spatial extent. The reasoning is as follows. Because the ideal observer summates perfectly across all cycles of the grating, the ideal CSFs for gratings of large, fixed extent are much flatter than those in Figure 13. On the other hand, because human observers summate only over a fixed number of cycles, their CSFs would not be flatter for gratings of large, fixed extent. Therefore, the ratio of real to ideal threshold would have decreased with spatial frequency (cf. Figure 14). The conclusion is that the cycle-summation properties of the visual system predict a flat NTF for CSFs measured with a fixed numbers of cycles but a decreasing NTF for CSFs measured with fixed spatial extents. Thus, the Banks et al. results are perfectly consistent with the fact that the NTF is not flat for CSFs measured with fixed spatial extents.

What the Banks et al. (1987) results suggest is that the form of the neural summation is summation over cycles, not over adjacent receptors. A flat NTF for gratings matched to the cycle-summation properties of the visual system, suggests that center mechanisms (and perhaps other subregions) of many postretinal receptive fields centered on the fovea consist of a single cone or column of cones. This conclusion is consistent

[6] For experiments with a fixed patch size, the definition of the neural-transfer function (NTF) used by Banks et al. (1987; i.e., the ratio of real to ideal sensitivity) is equivalent to the classic definition of the NTF. That is, it will yield the same shape as the contrast-sensitivity function measured when bypassing the eye's optics (except for the relatively minor effect of the receptor aperture). The definitions are not equivalent for the experiments with a fixed number of cycles.

with the anatomical evidence (Boycott & Dowling, 1969; Polyak, 1941) that the center mechanisms of bipolar and ganglion cells in the fovea are driven by a single cone.

A possible counterargument might be that the widths of center mechanisms measured by electrophysiology are always larger than a single cone (e.g., Derrington & Lennie, 1984; De Valois et al., 1982). However, this is expected because the measured receptive fields include the contributions of the preneural factors. Measurements of Ricco's area in the fovea (Davila & Geisler, 1987) have confirmed the conclusion that foveal center mechanisms are as small as a single cone (see ahead). Another possible counterargument might be that constant relative efficiency doesn't really imply anything about the sizes of center mechanisms because the later visual system could simply undo the effects of summation over adjacent receptors (without information loss) by some operation such as deconvolution. However, this argument would only be valid if the visual system were noise free. For example, if the ganglion cells summed over many receptors and if spikes were lost in transmission from the retina to the LGN at random, much like photons are lost in passing through a neutral density filter, then there would be no way to recover the information lost because of summation.

The preceding conclusions from the ideal-observer analysis must be tempered by two other considerations. First, the Banks et al. (1987) study focused deliberately on the spatial frequencies above 5 c/deg. This was because preneural mechanisms have little interesting effect on contrast sensitivities at lower spatial frequencies. Furthermore, neural mechanisms do have a large effect at low spatial frequencies. Figure 15 shows contrast-sensitivity functions measured for the low to medium frequencies by Kelly (1972). Below 3 c/deg, the CSFs flatten and then turn downward—a clear violation of the trends predicted by the ideal observer. Furthermore, the spacing between the CSFs below 3 c/deg is consistent with Weber's law ($\Delta N \propto N$) not the square-root (de Vries-Rose) law. As has been recognized for a long time, such effects must be produced by neural mechanisms. For example, the common interpretation of the falloff in contrast sensitivity at low frequencies is the spatially antagonistic receptive fields of retinal and cortical neurons (e.g., see Kelly, 1977). Weber's law is most likely due to gain adjustment (multiplicative adaptation), response compression, and noise mechanisms of the receptors and other neurons along the visual pathway (e.g., see Geisler, 1983; Shapley & Enroth-Cugell, 1984). It is also worth noting that the smooth transitions from square root to Weber behavior as a function of mean luminance could easily be a natural consequence of these mechanisms (e.g., see Geisler, 1984).

The second consideration is that even at high spatial frequencies, the preneural mechanisms are only responsible for some aspects of the contrast-sensitivity function. One obvious unexplained aspect of the data in Figure 14 is the factor of 20–30 reduction in sensitivity that must be produced by neural mechanisms. There are several factors that singly or in combination may contribute to this disparity. A highly likely factor is that the visual system is integrating over fewer than the 7 cycles that were used in the Banks et al. (1987) study. The visual system may integrate with efficiency only over a few cycles. Another simple possibility is multiplicative attenuation, which is equivalent to placing a pair of neutral-colored sunglasses over the eyes. This would occur, for example, if only some fraction of the photons absorbed in the photopigment were effective in eliciting a receptor response. Another possible explanation is observer uncertainty. If the human observer is uncertain about the position or spatial frequency of the target grating, there will be a fairly uniform loss of sensitivity. (Monte Carlo simulations show that position uncertainty produces slightly more effect at the highest spatial frequencies.) This could explain some of the differences between real and ideal performance because the SDE model does not include position or frequency noise. However, as will be seen in the discussion of contrast discrimination later, the effect of uncertainty is too small to explain much of the disparity between real and ideal performance. Another possibility is that there are sources of spatially and temporally uncorrelated noise within the visual system that approximately mimic the effects of photon noise (e.g., grow in proportion to the square root of background or signal level). A slightly more complex possibility is that there is an appropriate combination of internal noise and nonlinear transduction to produce a uniform sensitivity loss across spatial frequency.

The difference between ideal and real performance above 5 c/deg could be explained by any of these mechanisms or by some combination of them. In fact, Crowell, Banks, Anderson,

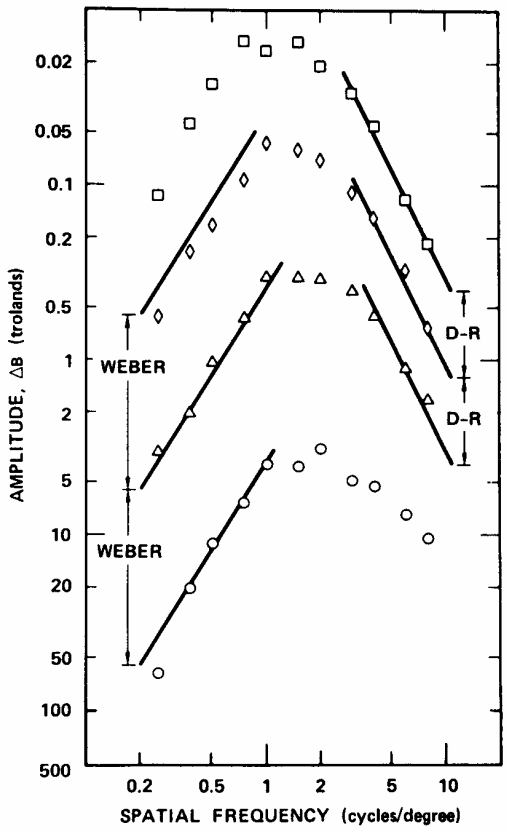


Figure 15. Contrast threshold for low- and midfrequency sine-wave gratings measured at several background levels. (Background luminance: ○ = 720 td, △ = 72 td, ▽ = 7.2 td, □ = 0.72 td; from Kelly, 1972.)

and Geisler (1988) found recently that the ratio of real to ideal contrast sensitivity was reduced to a factor of approximately 4, over the 5–40 c/deg range, by using fewer than 7 cycles and by adding a pedestal contrast. However, the important point here is that these neural mechanisms must have (or sum to have) a uniform effect because the shapes of the functions are already predicted by losses of information occurring prior to neural processing and, in the case of fixed-size gratings, by spatial summation across cycles.

Increment detection. The classic intensity-discrimination experiment requires the observer to discriminate the intensity differences between two regions. This fundamental experiment has been extensively used in the study of light and dark adaptation. It is particularly well suited for measuring visual sensitivity in local regions of the visual field. This is important because of the severe spatial inhomogeneity of the visual system (e.g., see the receptor distributions in Figure 8).

A common stimulus arrangement consists of a circular region and a surrounding region. In the simple increment- or decrement-threshold version of this experiment, the center and surround luminances are the same for Stimulus a (background alone), and the center is incremented (or decremented) for Stimulus b (background plus increment). For green (530-nm) light and a small center region, the mean absorptions in the photoreceptors for the two patterns is shown in Figures 16A and 16B. Again, the G receptors appear brighter because they absorb a greater fraction of the photons. The ideal receptive field for making the discrimination consists of a single excitatory region that covers the area of the increment (Figure 16C).

Aside from wavelength composition, there are several important stimulus dimensions in the increment-threshold experiment: the intensity, size, shape, and duration of the background, and the size, shape, and duration of the increment. Figure 17 shows ideal- and real-observer performance as a function of several of these dimensions. In each case, the predictions of the ideal observer have been shifted by the relative-efficiency method for the purposes of comparing shapes. The thresholds for the ideal discriminator depend on stimulus duration and wavelength composition. However, the only effects of duration and wavelength are to translate the predicted curves vertically (in Figure 17B and 17C) or vertically and horizontally by equal amounts (in Figure 17A).

The solid curves in Figure 17A show ideal-observer performance as a function of background intensity for a foveally presented increment of 3.5′ and 50′ diameter. These threshold-versus-intensity (t.v.i.) curves represent the effect of all the preneural factors, plus a small arbitrary amount of receptor dark noise (see Equation 5). As previously mentioned, the dark-noise level in cones is somewhat uncertain at this time. The dotted line shows the ideal observer's performance if there is no dark noise. Notice that once the background intensity sufficiently exceeds the dark-noise level, the curves follow the familiar square-root relation.

The symbols in Figure 17A are the foveal t.v.i. functions reported in Geisler (1978) for white light and a 2-mm pupil. They are similar to those reported by Barlow (1958b). As Barlow pointed out, when the increment diameter is very small, there is a modest (1–2 log unit) intensity range over which the square-root law holds, but at higher intensities, Weber's law holds. For increment fields with larger diameters, there is a much smaller square-root region. (The rod system displays a greater square-root region for small spots.) Fourier transforms show that small-diameter increments have a greater high-frequency content than large-diameter increments. Therefore, the increment detection results are qualitatively consistent with the grating detection results that show a smaller square-root range for low spatial frequencies.

Figure 17B shows foveal increment threshold as a function of background diameter for the ideal discriminator along with Westheimer's (1967b) measurements with white light for human observers. Ideal and human observers both show an initial increase in threshold as background diameter is increased, but only the human observer shows a subsequent decrease in threshold (the sensitization effect). Thus, Figure 17B confirms Westheimer's conclusion that the sensitization effect must be of neural origin. Ever since Westheimer's work, the background-summation experiment has been used extensively to map local spatial interactions in the visual system (MacLeod, 1978). Like the low-frequency roll-off of the CSF, the sensitization effect has been attributed to the spatially antagonistic receptive fields of retinal and cortical neurons (Buss, Hayhoe, & Stromeyer, 1982; MacLeod, 1978; Westheimer, 1967b). Figure 17B suggests that the initial rise in foveal threshold with background diameter may be partly attributable to preneural factors.

The foveal increment threshold (in energy units) of the ideal discriminator as a function of increment area is given by the solid curve and solid circles in Figure 17C. The model predicts that for small areas, threshold is independent of area (Ricco's law) and for larger areas, threshold is proportional to the square root of area (Piper's law). The prediction of Ricco's law is the result of the optical point-spread function and the receptor aperture. When the area of the increment shrinks below a certain point, the distribution of absorbed photons across the receptors becomes fixed—further decreases in area only reduce the total number of absorbed photons. Piper's law is predicted by quantal fluctuations (Barlow, 1958a).

The open circles in Figure 17C are measurements by Davila and Geisler (1987) for green background and increment fields (peak spectral emission of 550 nm) and a 3-mm pupil. The close agreement between real and ideal performance implies that in the fovea, Ricco's area is almost completely attributable to the optical point-spread function and the receptor aperture. This apparent agreement was verified by the following analysis. A minimization program (STEPIT; Chandler, 1969) was used to obtain (simultaneously) least squares fits of straight lines to the horizontal and rising portions of the real data and the ideal predictions. The intersection of the lines was then taken to be the size of Ricco's area. The average Ricco's area for 3 subjects was 5.0 min^2 and that of the ideal observer was 5.4 min^2. These areas are within experimental error of being equal. (This psychophysical estimate of Ricco's area confirms the much earlier study of Lamar, Hecht, Shlaer, & Hendley, 1947.)

Like the grating detection experiments, these results suggest that the center mechanisms of many neurons in the retina and cortex are driven by a single cone or row of cones. If the neural summation produced by center mechanisms were as large as that produced by the optics, then the combined effect of the preneural and neural summation would produce necessarily a

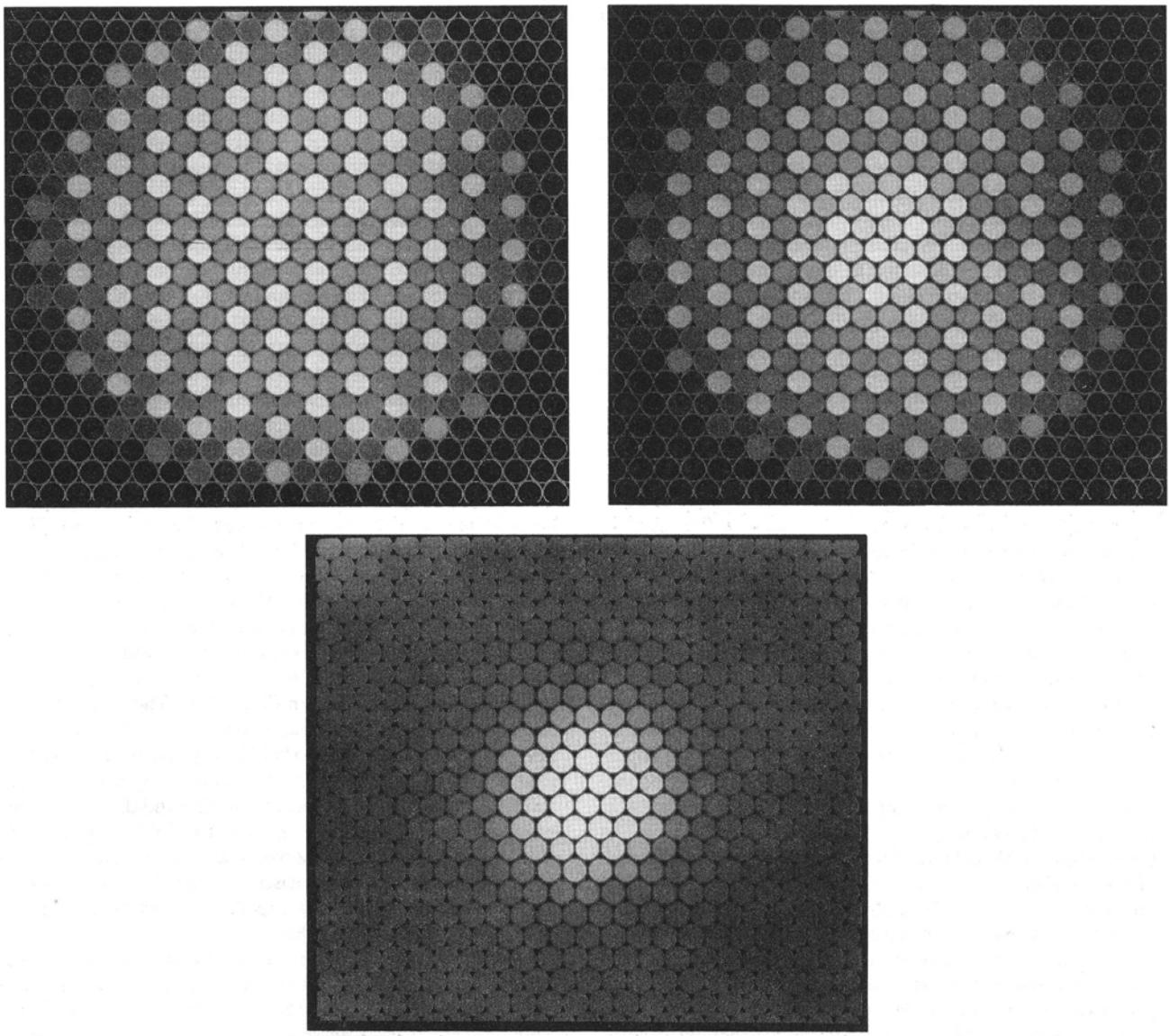
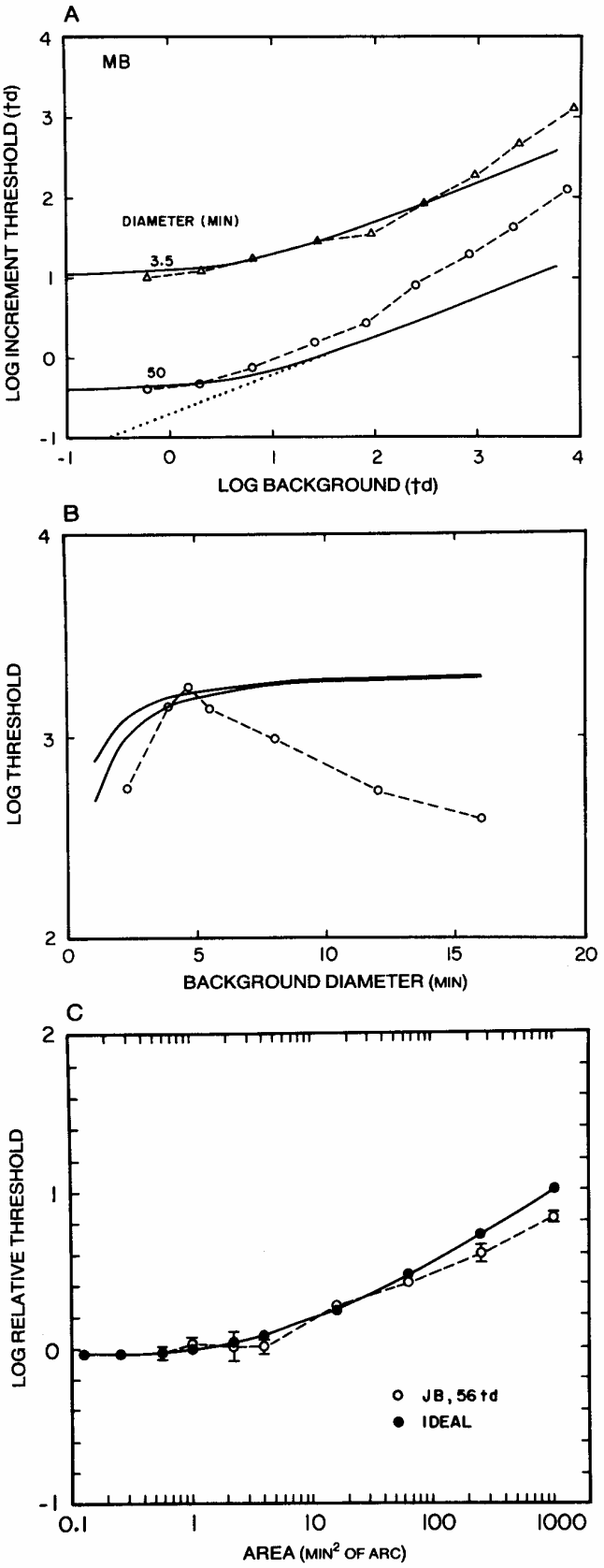


Figure 16. Increment detection. (A. Average pattern of photon absorptions in the receptors produced by a green circular background. B. Average pattern of absorptions produced by the background plus a green increment field of 3′ diameter. C. Ideal receptive field for discriminating the patterns in A and B.)

Ricco's area of at least 10 min^2 (assuming that neural noise would prevent any undoing of the summation effects). Thus, any neural summation must be a small fraction of that produced by the optics (cf. Figure 3).

It may seem puzzling at first that there is reasonable agreement between real and ideal performance in the Ricco and contrast-sensitivity experiments, but not in the background-summation experiment. However, this may be explained by the fact that performance in the background-summation experiment should be strongly influenced by the nonlinear response properties of visual neurons. The powerful nonlinearities involved in the background-summation experiment have been most clearly demonstrated by Hayhoe and her colleagues (e.g., Buss et al., 1982). Their results suggest that the center/surround receptive fields in the early visual system require a substantial amount of light on the surround areas in order to prevent severe response saturation and hence infinite thresholds. On the other hand, these severe nonlinearites should not come into play in the Ricco and contrast-sensitivity experiments because the experiments are run against a large uniform background and because the test stimuli (at threshold) consist of relatively small deviations from the background intensity level. (For a discussion of why small perturbation experiments can be unaffected by certain types of simple nonlinearities, see Cornsweet, 1970, pp. 324–328.)

Contrast and increment discrimination. So far, we have con-



sidered only the detection of target patterns against a uniform field of light. Although simple detection is essential in even the most basic visual systems, useful visual pattern recognition also requires that the organism be able to discriminate between the various intensity differences existing in the visual scene. The contrast- and increment-detection paradigms described previously can be modified to study this kind of discrimination.

An important version of the contrast-discrimination experiment requires the observer to discriminate changes in the contrast of a sine-wave grating (Bradley & Ohzawa, 1986; Legge, 1981; Pelli, 1985; Nachmias & Sansbury, 1974). The two obvious variables to manipulate are spatial frequency and the base contrast of the two gratings to be discriminated. The data in Figure 18A are the contrast-discrimination thresholds for Gaussian-damped gratings of 4, 8, 16, and 24 c/deg as a function of base contrast (Geisler & Davila, 1987). The Gaussian damping was such that only about 2.5 cycles of the grating patch were visible (i.e., the bandwidth of the stimuli was 1.0 octave). The Gaussian damping covered twice this extent in the direction parallel to the gratings. The stimuli were presented for 300 ms on a Joyce CRT (cathode-ray tube) with a white (P4) phosphor at 150 cd/m^2, in a two-alternative forced-choice procedure. The stimuli were viewed through a 3-mm artificial pupil.

Note that at all spatial frequencies, threshold initially declines as base contrast is increased and then begins to increase. The minimum of the curves occurs at about the point at which the base contrast is at detection threshold.

The solid curves show the performance of the ideal discriminator for the same stimuli. As usual, the relative-efficiency method was used to translate the predicted curves; however, the relative positions of the curves were not altered. As can be seen, the ideal discriminator's contrast threshold is unaffected by the grating contrast. In other words, the information available at the photoreceptors for contrast discrimination is independent of the base contrast. There are two obvious conclusions one can draw from this. First, the effects of base contrast on threshold are neither the result of preneural factors (excluding uncer-

Figure 17. Comparison of increment detection for real and ideal observers. (In all cases, the predictions of the ideal observer have been scaled by the relative-efficiency method for the purpose of comparing shapes. A. Increment threshold as a function of steady background intensity for two increment diameters. Ideal-observer predictions include an arbitrary level of receptor dark noise; approximately 80 events/s/receptor. Dotted line shows one of the predicted curves without dark noise. Ideal predictions were translated vertically by 1.1 log units; neutral density value = 2.2 log units. B. Increment threshold as a function of background diameter for a 1′ test spot. Upper solid line shows the ideal performance when the entire stimulus pattern is set against a dim background; lower solid line shows the ideal performance when it is set against a dark background. Data are from Westheimer, 1967b. Ideal predictions have been translated vertically by 1.8 log units; neutral density value = 3.2 log units. C. Threshold energy as a function of increment field area for square increment fields. Ricco's area, the horizontal portion of the curves, is estimated to be nearly the same for both real and ideal observers. Ideal predictions were translated vertically by 1.0 log units; neutral density value = 2.0 log units.)

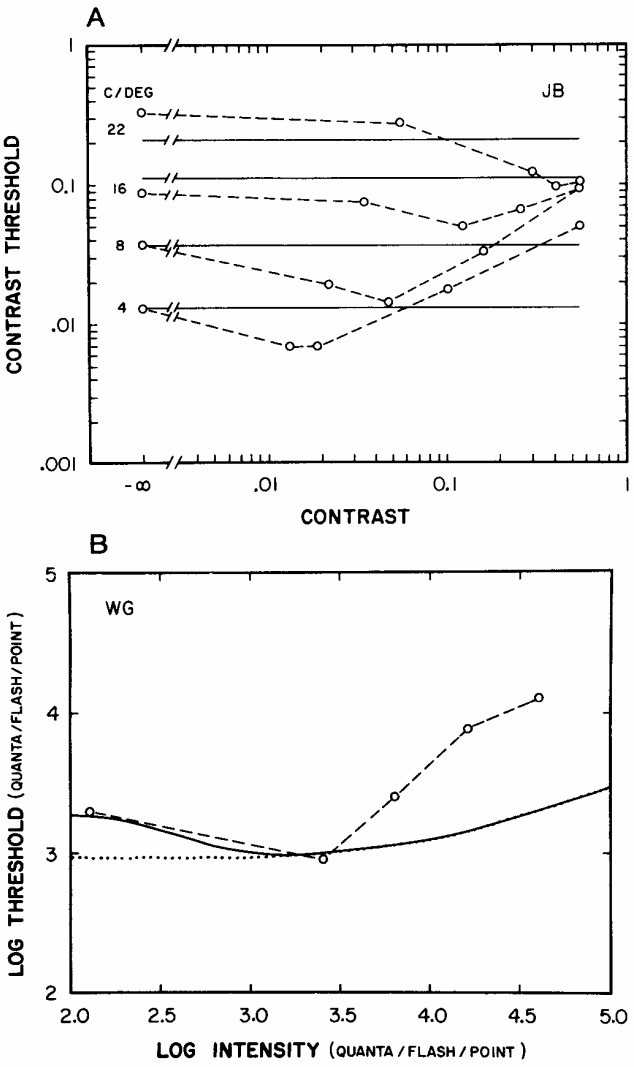


Figure 18. Contrast and increment discrimination for real and ideal observers. (Predictions of the ideal observer have been scaled by the relative efficiency method for the purpose of comparing shapes. A. Contrast discrimination threshold as a function of base contrast for 1.0-octave Gabor patches at 5 spatial frequencies. Ideal predictions were translated vertically by 1.5 log units; neutral density value = 3.0 log units. B. Increment discrimination threshold for point-source stimuli as a function of point-source base intensity measured against a 21-td steady background. Dotted line is the prediction of the ideal observer with an no position uncertainty. Solid line is the ideal-observer prediction with—in addition—an uncertainty region of 20 min^2. Ideal predictions were translated vertically by 0.8 log units; neutral density value = 1.6 log units.)

tainty due to eye movements) nor of changes in the discrimination information available in the stimuli. Second, the relative efficiency of the human observer to the ideal discriminator is best when the base contrast is equal to detection threshold. The ideal-observer analysis shows that all of the effects of base contrast on threshold must be explained by neural mechanisms.

The increment-discrimination paradigm yields results similar to those for contrast discrimination. The data points in Figure 18B show increment discrimination for point sources on a uniform background as a function of pedestal point-source intensity. These data, from Geisler and Davila (1985), confirm the well-known result (Leshowitz, Taub, & Raab, 1968; Nachmias & Kocher, 1970; Whittle & Swanston, 1974) that increment-discrimination threshold decreases until the pedestal is near its detection threshold and then rises according to Weber's law at higher pedestal intensities. The dotted curve shows the performance of the ideal observer, translated appropriately for the purposes of comparing shapes. Again, it is seen that most of the effects of pedestal intensity cannot be explained by preneural factors. Notice that the ideal observer does predict an increase in threshold with pedestal level; however, the prediction is that threshold should follow the de Vries-Rose law, not Weber's law.

We now briefly consider some of the neural hypotheses that have been proposed for the contrast- and increment-discrimination functions. It is quite possible that the initial decrease in threshold, the so-called *dipper* or *pedestal* effect, may be of different origin from the subsequent increases in threshold; thus, we will consider explanations for the two parts of the curve separately.

There have been several proposed explanations for the pedestal effect. One is the uncertainty hypothesis (Cohn & Lasley, 1974; Foley & Legge, 1981; Nachmias & Kocher, 1970; Pelli, 1985; Tanner, 1961). The hypothesis is that the observer might have some uncertainty about one or more aspects of the signal (e.g., position, spatial frequency, or size). Because of this uncertainty, the observer must look for the signal throughout the uncertainty regions. When the base contrast is zero, there is a substantial probability that some part of the noise existing within the uncertainty regions will be similar enough to the signal to be misinterpreted. But when the base contrast is itself high enough to be detectable, then the observer can base his or her judgment on the maximum response within the uncertainty regions, which will always occur at the correct place. Pelli (1985) has shown in detail how quantitative models incorporating uncertainty effects can predict appropriate pedestal effects. The predictions of the uncertainty hypothesis can also be derived within the framework of the present ideal-observer analysis by considering the ideal discriminator for situations in which there is real stimulus uncertainty. In this case, Equation 7 gives the maximum likelihood decision rule. Monte Carlo simulations with Equation 7 show that the ideal observer predicts a substantial pedestal effect even for small uncertainty regions. For example, the solid curve in Figure 18B shows ideal-observer performance for an uncertainty region of 20 min^2. This is the amount of uncertainty that would be introduced into the retinal image if the random eye movements that occur under steady fixation are not accurately registered by the visual system (Geisler & Davila, 1985). Thus, it appears that uncertainty could account for the pedestal effect.

Also, it is apparent that uncertainty may account for some of the differences between real and ideal performance in the simple contrast-detection experiment, as described earlier. However, the size of the observed pedestal effect implies that uncertainty cannot account for more than a small part of the differ-

ence (about 0.5 log units) between real and ideal contrast sensitivity. The predicted pedestal effect (like the observed effect) is about the same size at all spatial frequencies. Thus, as mentioned earlier, position uncertainty uniformly scales the predicted contrast-sensitivity function.

The pedestal effect could also be explained by the existence of appropriate combinations of nonlinear transduction and internal noise, or by an internal-threshold mechanism. The nonlinear-transduction hypothesis requires that the contrast signals first pass through a positively accelerating nonlinearity and then undergo degradation by postnonlinearity noise (Legge & Foley, 1980; Nachmias & Sansbury, 1974; Wilson, 1980). Because of the nonlinearity, increasing the pedestal contrast increases the size of the response given to the contrast increment. Because some of the noise occurs after the nonlinearity, there is an improvement in the signal-to-noise ratio and hence a decrease in threshold. The pedestal effect could also be explained if the contrast signals had to exceed some internal threshold in order to be transmitted further along the visual pathway (Blackwell, 1963; Foley & Legge, 1981). Obviously, any base contrast that pushes the contrast signals toward or above this internal threshold will cause a decrease in contrast threshold.

Two types of mechanism have been proposed to explain the monotonic increase in contrast threshold that occurs when the pedestal exceeds detection threshold. One is a compressive nonlinearity followed by the addition of internal noise. This explanation is not necessarily inconsistent with the accelerating nonlinearity suggested to explain the pedestal effect. Indeed, visual neurons often display an accelerating nonlinear response at low contrasts and a highly compressive nonlinearity at high contrasts (Albrecht & Hamilton, 1982; Werblin & Copenhagen, 1974). Another, possible mechanism is a monotonic increase in the internal noise with stimulus contrast. Cortical neurons often display increases in response variability with increasing response (Tolhurst, Movshon, & Dean, 1983). It is possible, of course, that both mechanisms contribute to the psychophysical results.

Subthreshold summation and contrast masking. The subthreshold-summation paradigm (Graham, Robson, & Nachmias, 1978; Sachs, Nachmias, & Robson, 1971; Watson, 1982) and the contrast-masking paradigm (Legge & Foley, 1980; Stromeyer & Julesz, 1972; Wilson, McFarlane, & Phillips, 1983) have been used extensively to study the frequency selectivity of spatial mechanisms within the visual system. In recent years, they have provided the major source of data used in the development of multiple spatial-frequency channel models of spatial vision (Graham et al., 1978; Watson, 1983; Wilson et al., 1983). In these models, each retinal location is processed in parallel by multiple channels (e.g., classes of receptive field), each tuned to a different range of spatial frequencies and orientations. Usually, the channel outputs are assumed to be statistically independent.

In the typical subthreshold-summation experiment (e.g., Watson, 1982), thresholds for compound gratings consisting of a pair of spatial frequencies are compared with the thresholds for the component gratings presented alone. When the component gratings are close in spatial frequency, their signals add perfectly in the compound grating. Thus, the compound grating reaches threshold when each component grating is set to one half its individual threshold. Experiments show that when the gratings differ sufficiently in frequency (e.g., by a factor of 2), each component must then be set well above one half (up to around 80%) its individual threshold value in order for the compound grating to reach threshold. This effect has been taken to imply that stimuli containing sufficiently different spatial frequencies are processed by separate channels (e.g., classes of receptive field) tuned to different spatial-frequency ranges.

However, the ideal-observer model also predicts much (but not all) of this latter effect without including neural mechanisms. In particular, it predicts that when the gratings in the compound are of sufficiently different frequency (e.g., differ by a factor of 2), the components must be set to around 70% of their individual threshold contrasts, as opposed to 50% when they are close in frequency.

The reason that the ideal-observer model makes this prediction is similar to the reason that it is predicted by probability summation across space within a single spatial-frequency channel (Graham et al., 1978; Quick, Hammerly, & Reichert, 1976). When the sine waves are sufficiently different in spatial frequency, the bands of frequency components in the Poisson noise that mask the sine waves' detectability are statistically independent; whereas, when the sine waves are similar in frequency, the masking frequency components are *not* statistically independent. However, the important point here is that the predicted effect is attributable to the information content of the stimuli at the cornea. In other words, the information available to perform the detection is physically lower at the cornea when the frequencies are dissimilar. Obviously, this information difference in the stimuli needs to be considered when trying to draw inferences about the spatial-frequency tuning of neural mechanisms from subthreshold-summation data.

The contrast-masking experiment does not suffer from the same problem. In the typical version of this paradigm (e.g., Wilson et al., 1983), the observer must detect a target sine-wave grating patch added to a high-contrast masking grating of some other spatial frequency. In this case, the predictions derived from the ideal observer show *no* effect of masker spatial frequency or orientation for much the same reason that no effect of base contrast is predicted in the contrast–discrimination experiment. On the other hand, the data show that target threshold decreases monotonically as the masker frequency is moved away from the target frequency in either direction. These effects must be due to neural mechanisms and may reflect the fair degree of frequency selectivity and the wide range of center frequencies displayed by neurons in striate cortex.

Location and Shape Discrimination

In the location-discrimination task, the observer must discriminate changes in the relative positions of two or more objects whose shapes, intensities, and wavelength compositions remain fixed. For present purposes, these objects are regarded as two dimensional (i.e., as intensity functions in retinal coordinates); thus, the location changes can consist only of vertical or horizontal translation or rotation.

Location-discrimination performance displays a number of interesting properties. Perhaps the most spectacular is that under the right conditions, humans can discriminate changes in

relative location that are only a fraction of the diameter of a single cone. For example, in the vernier-discrimination task, observers can detect a change of only a few seconds of arc in the relative location of two nearly abutting lines (Wülfing, 1892). This appears remarkable both because the diameter of a foveal cone is much larger than this value (around 30″ of arc) and because the optics of the eye spread out the light over many receptors. Furthermore, this high degree of sensitivity holds even when the lines are reduced to points (Ludvigh, 1953; Sullivan, Oatley, & Sutherland, 1972; Westheimer & McKee, 1977b). Westheimer (1975) has coined the term *hyperacuity* to describe location-discrimination thresholds that are significantly smaller than the diameter of a single cone.

Hyperacuity performance is not unique to the vernier task. Westheimer and McKee (1977b) and others (e.g., see Westheimer, 1981) have shown that hyperacuity performance can be observed for a wide range of stimulus shapes (e.g., lines, dots, chevrons, small blobs, and sine-wave gratings) and tasks (e.g., bisection discrimination, separation discrimination, displacement discrimination, and stereo discrimination). The bisection task, in which the observer must decide whether one object is centered between two others, can yield thresholds near 1″ of arc (Klein & Levi, 1985).

The achievement of hyperacuity performance in location-discrimination tasks has been a long-standing puzzle in vision science. The fact that such performance levels can be achieved with a rather wide range of stimulus configurations has led a number of investigators to suggest that special cortical mechanisms (e.g., interpolation networks) exist for extracting relative location information (Barlow, 1979, 1981; Crick, Marr, & Poggio, 1981; Hirsch & Hylton, 1982; Westheimer, 1981). However, ideal-observer analysis removes much of the mystery by showing that many of the location-discrimination phenomena coincide nicely with variations in the information content of the stimuli at the photoreceptors (Geisler, 1984; Geisler & Davila, 1985, and below). Thus, it is quite possible that most location-discrimination phenomena are attributable to the same mechanisms thought to underly other basic visual discrimination phenomena.

To illustrate the basic properties of location discrimination, I will concentrate as much as possible on the separation-discrimination task (especially on the two-line and two-point separation tasks). In the separation task, the observer must discriminate changes in the distance between two identical objects. Stimulus a is a pair of objects at some base separation; Stimulus b is a pair with some incremental change in the separation. Not all of the basic variables affecting location discrimination have been examined in a separation task, so other localization tasks will be considered as needed.

Base separation. The effect of base separation in the two-line separation-discrimination task for lines under 1 min in width is shown in Figure 19A. These data from Westheimer and McKee (1977b) illustrate a typical result for simultaneously presented objects; namely, hyperacuity is observed only when the objects are separated neither too much nor too little. The base separation apparently must be within the range of approximately 1–10 min of arc, with maximum sensitivity being reached for base separations of 2–5 min. Note that when the base separation is zero (a case not shown in Figure 19), the separation task reduces to the classical task of deciding whether there is one or two line segments (Helmholtz, 1866/1925; Wilcox & Purdy, 1933). It is well-known that threshold in the two-line (or two-point) resolution task is typically 45–60 s, a value that appears to be consistent with the data in Figure 19A.

It is important to note that the separation thresholds reported by Westheimer and McKee (1977b) were doubled before plotting them in Figure 19. This was done (a) to facilitate comparison of data obtained with their procedure and that obtained with standard two-alternative forced-choice procedures and (b) to allow proper comparison of performance in hyperacuity and resolution tasks.[7]

It is intuitively reasonable that location discrimination should deteriorate at large base separations; however, it is not obvious why location discrimination at separations of 1–10 min should be better than two-line resolution (0-min separation). Indeed, the observed values of two-line resolution threshold are often viewed as intuitively reasonable. Because the receptor diameter is 30 s, a 60-s separation of the lines should center them on two rows of receptors that are separated by exactly one row of receptors. It would appear then that two-line resolution matches nicely the sampling grain of the receptor lattice. Therefore, the puzzle has been the much smaller thresholds observed at base separations of 1–10 min. It is this puzzle that has led to the hypothesis of special mechanisms for hyperacuity.

Consider how the ideal observer performs in the separation-discrimination task. Figure 20A and 20B show the pattern of photon absorptions produced by two short, green line segments 30 s in width that are separated by 3 and 3.2 min, respectively. The ideal receptive field for making the discrimination is shown

[7] Westheimer and McKee (1977b) used a somewhat nonstandard psychophysical procedure in which one of seven separations of the stimuli were presented on each trial. The subject had to judge whether the separation was greater or less than the middle of the range (the fourth separation). Westheimer and McKee took *threshold* to be the change in separation from the midrange separation required for 75% correct. This definition of threshold might lead one to believe that the subject could reliably discriminate the midrange stimulus from the threshold stimulus in a simple forced-choice situation. However, this is highly unlikely on theoretical grounds. Signal detection analysis and simulations with the ideal observer show that the optimal strategy for the subject is to regard the Westheimer and McKee task as a standard single-interval, two-alternative task that has only two of the stimuli that bracket the midrange. With this strategy, the ideal-observer's threshold is the same in the simple two-stimulus task and in the Westheimer and McKee task. Yet, the ideal-observer cannot discriminate, at 75% correct, the midrange stimulus from the threshold separation. Therefore, it is likely that the better definition of threshold in the Westheimer and McKee task is the difference in separation between the lower and upper 75%-correct points.

Dennis McFadden and I have confirmed this argument psychophysically. We measured vernier threshold in the Westheimer and McKee (1977b) task and in a single-interval yes/no task. We found that thresholds were approximately a factor of 2 larger in the yes/no task. Hence, the Westheimer and McKee definition of threshold underestimates location-discrimination thresholds by a factor of 2. Many of the studies of location discrimination that are described here have used the procedure of Westheimer and McKee. When appropriate, the thresholds have been doubled.

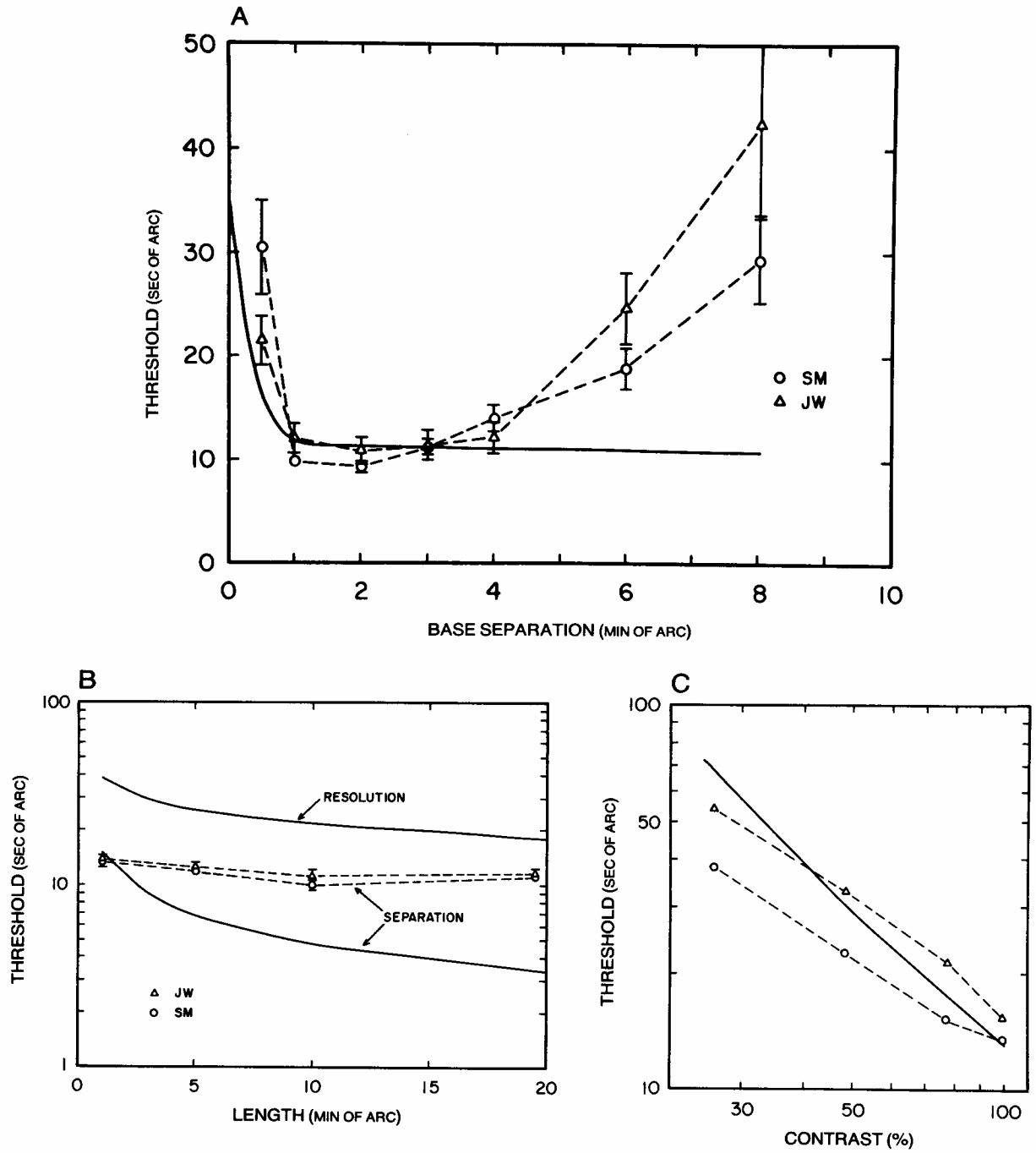


Figure 19. Location discrimination for real and ideal observers. (Predictions of the ideal observer have been scaled by the relative-efficiency method for the purpose of comparing shapes. In this case, scaling does not correspond to vertical translation in log coordinates. A. Separation thresholds for thin, 12.8′-long lines [presented for 200 ms] as a function of base separation. Data for 2 subjects are from Westheimer & McKee, 1977b. Ideal predictions were scaled using a neutral density value of 3.6 log units. B. Two-line separation and resolution thresholds as a function of line length. Separation discrimination data for 2 subjects are from Westheimer & McKee, 1977b. Ideal predictions were scaled using a neutral density value of 2.15 log units. C. Vernier acuity for 1′-wide dark lines presented for 750 ms on a 17.5 cd/m^2 background as a function of contrast [100% contrast is a completely dark line]. Data for 2 observers are from Wilson, 1986. Ideal predictions were scaled using a neutral density value of 3.3 log units.)

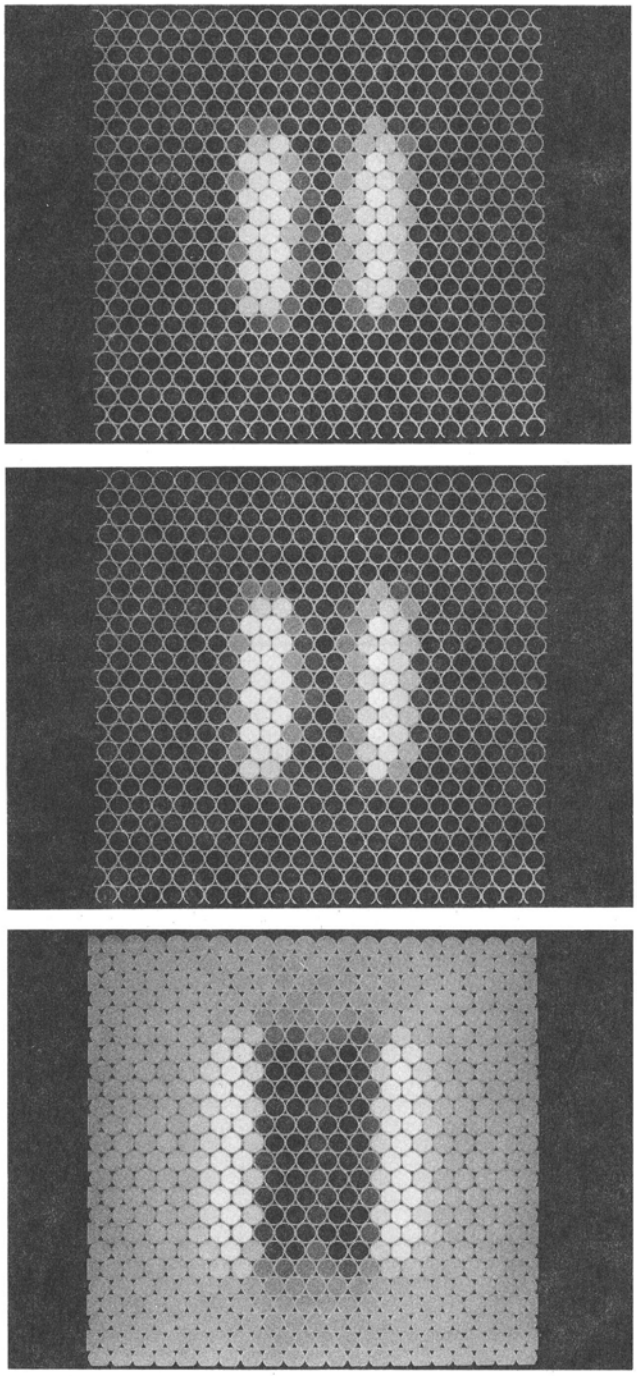


Figure 20. Two-line separation discrimination. (A. Average pattern of photon absorptions in the receptors produced by two thin [30″-wide] green lines separated by 3′. B. Average pattern of absorptions produced when the lines are separated by 3.2′. C. Ideal receptive field for discriminating the patterns in A and B.)

in Figure 20C. The absorption patterns and ideal-receptive field for two-line resolution are shown in Figure 21. The performance of the ideal observer is given by the solid curve in Figure 19A, again scaled by the relative-efficiency method for the purpose of comparing shapes.

Like the human observer, the ideal observer is much less sensitive when the base separation is zero (resolution) and achieves maximal sensitivity when the base separation reaches 2–5 min. Thus, the present ideal-observer analysis shows that the discrimination information available at the photoreceptors in the two-line separation (hyperacuity) task is many times better than that in the two-line resolution (acuity) task. This suggests that the same neural mechanisms, operating at a constant level of efficiency, could account for both results—no special mechanisms would seem to be required. (In the discussion section, I will show that a relatively sparse array of receptive fields could underlie hyperacuity performance without the need for neural interpolation processes.)

Unlike the human observer, the ideal observer's performance does not deteriorate at larger base separations. This occurs because the ideal observer integrates information perfectly over arbitrary distances. Apparently, the human observer is unable to integrate information accurately over such large distances.

Stimulus size. Westheimer and McKee (1977b) also measured the effect of line length on separation discrimination for a base separation of 3 min. The results for 2 subjects are shown in Figure 19B. As can be seen, there is only a slight improvement in sensitivity with line length—two points are localized about as well as two lines. Consistent with this result is the fact that two-point resolution is about as good as two-line resolution (e.g., see Riggs, 1965). This result also illustrates that the human visual system does not effectively integrate information over large distances.

Of course, the ideal observer does integrate information perfectly. The lower solid curve in Figure 19B shows the effect of line length in the separation task for the ideal observer. Ideal-observer sensitivity increases with the square root of line length (unless the line length is very small), whereas real sensitivity increases much more gradually. The upper solid line in Figure 19B shows that two-line resolution is predicted to increase more slowly (with the fourth root) of line length.

Stimulus intensity and contrast. Geisler and Davila (1985) measured the effects of point source and background intensity on two-point resolution and two-point separation discrimination. The stimuli were presented for 100 ms and viewed through a 3-mm artificial pupil. The points were less than 30 s in diameter and were created on a CRT screen with a green (P15) phosphor. The open symbols in Figure 22A show the results for the separation task, with a base separation of 3 min. The circles were obtained against a dark background, the triangles against a uniform background of 21 td (photopic trolands). As point-source intensity is increased, threshold is seen to decline rapidly at first and then more gradually, reaching an asymptotic value of around 11 s. The solid symbols are the results for the resolution task (base separation of zero). Again, threshold declines rapidly at first, but quickly reaches a plateau of around 48 s. These asymptotic values for resolution and separation discrimination are consistent with the data of Westheimer and McKee (1977b) shown in Figure 19A. Another interesting aspect of the data is that at low point-source intensities, resolution threshold is lower than separation threshold (the curves cross each other).

It was rather surprising to find that the ideal observer shows

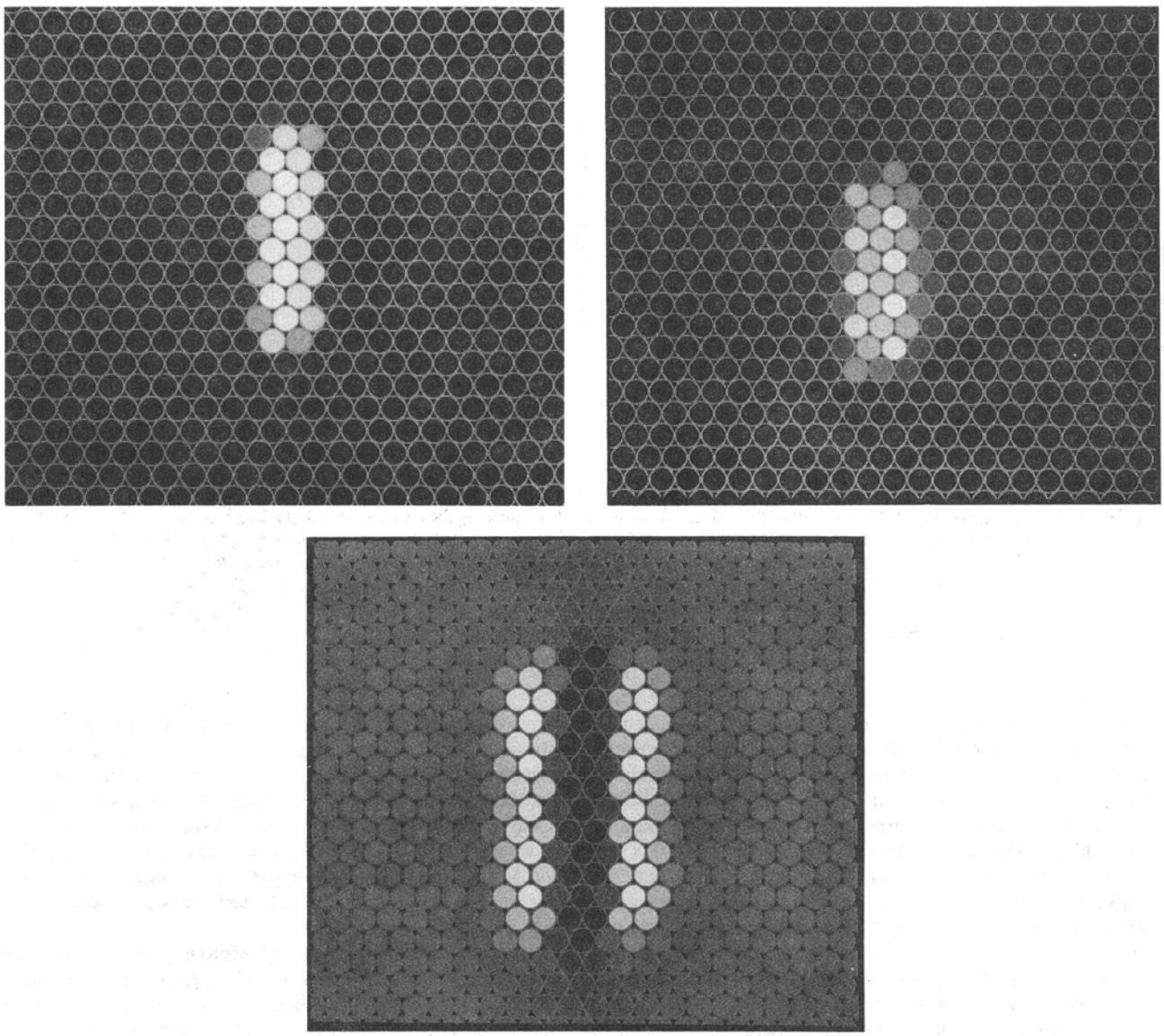


Figure 21. Two-line resolution. (A. Average pattern of photon absorptions in the receptors produced by two superimposed thin, 30″-wide green lines. B. Average pattern of absorptions produced when the lines are separated by 50″. C. Ideal receptive field for discriminating the patterns in A and B.)

a different effect of line length in the resolution task than in the separation task. As shown in Figure 22B, the ideal observer also shows a different effect of intensity in the two tasks. In particular, the ideal observer's resolution threshold decreases more gradually with increasing intensity than does its separation threshold. At high intensities, resolution threshold decreases with the fourth root of intensity, but separation threshold decreases with the square root of intensity. In addition, at low intensities the resolution and separation curves cross each other so that resolution threshold is lower than separation threshold. These predictions hold qualitatively in the psychophysical data, although the quantitative fit is not precise. Specifically, the observed slopes are shallower than predicted, corresponding to a deviation from the predictions toward Weber's law (Geisler & Davila, 1985). As was mentioned earlier, similar deviations from the ideal-observer predictions toward Weber's law are observed in intensity-discrimination experiments (e.g., see Figures 17A and 18B).

Note that the predictions in Figure 22B have not been translated; thus, the reader can compare the *absolute* sensitivities of the human and ideal observers. The predictions shown in Figure 22B were generated with a small level of dark noise included, which has the effect of causing the predicted curves for the dark-background condition to rise steeply toward infinity at low point-source intensities. The dark noise has no noticeable effect on the predictions for the 21-td background conditions.

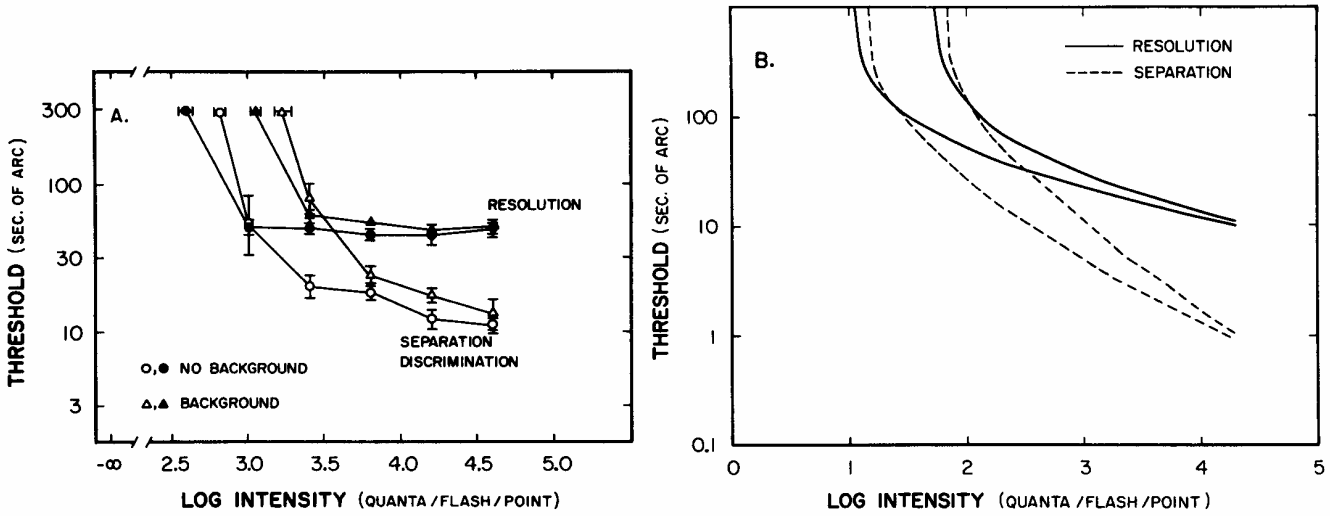


Figure 22. Two-point resolution and separation discrimination as a function of point-source energy. (A. Circles are thresholds obtained in the dark-adapted eye. Triangles are thresholds obtained against a steady background of 21 td. Open symbols are separation thresholds; closed symbols are resolution thresholds. B. Predictions of the ideal observer for exactly the same conditions as in A. Dark-adapted predictions were made assuming an arbitrary amount of receptor dark noise; 23 events/s/receptor. Dark noise has no significant effect on predictions for the 21-td background predictions. Predictions of the ideal observer have not been scaled.)

Geisler and Davila (1985) showed that the surprising behavior of the ideal observer as a function of intensity is not dependent on the particular parameters or assumptions made about the preneural mechanisms, but is explained by more fundamental properties of the information contained in the stimuli. In particular, they showed by mathematical analysis that the separation task is a "first-order task"—the information for performing the task is contained in the first derivative of the stimulus shape at the retina. Whereas the resolution task is a "second-order task," the information is contained in the second derivative of the stimulus shape. They also showed that the square-root and fourth-root relations hold almost independently of stimulus shape, as long as the intensity of the targets is sufficiently high relative to the background intensity. Thus, these relations can be added to the de Vries-Rose law as general properties of quantum-noise-limited detectors and discriminators:

$\Delta N \propto N^{1/2}$ intensity discrimination

$\Delta\Theta \propto N^{-1/2}$ location discrimination (separation > 2′)

$\Delta\Theta \propto N^{-1/4}$ resolution (separation = 0′)

It should be kept in mind, however, that under many circumstances, the target intensities do not differ sufficiently from the background for these relations to hold. For example, in Figure 22B, the square-root and fourth-root relations are only approximated at the highest point-source intensities. Another example of this is the ideal-observer model's prediction for the experiment of Wilson (1986). Wilson measured vernier threshold for dark, 1′ × 8′ lines on a background of fixed luminance, as a function of the intensity difference between the target and background. In Figure 19C, Wilson's data and the performance of the ideal observer are compared. Again, the ideal-observer's sensitivity has been reduced (by the relative-efficiency method) for the purposes of comparing shapes. As can be seen, both the ideal observer and the real observer display a power–law relation, but with an exponent significantly more negative than −0.5. The receptor-absorption patterns and ideal receptive field for vernier discrimination with dark lines is shown in Figure 23.

Image blur. Westheimer and McKee (1980), Westheimer (1979), and Watt and Morgan (1983) have examined the effects of blur on location discrimination. For example, Westheimer and McKee measured the effects of Gaussian blur on three-line stereo discrimination. Their data and the performance of the ideal observer are shown in Figure 24. There are two points to make about the effects of blur. First, the ideal-observer analysis predicts that there should only be a modest effect of blur on the localization of line stimuli. In other words, the information for performing the task is not greatly reduced by moderate amounts of blurring. But it does predict a considerably larger effect than that found by Westheimer and McKee. For localization of edges in a vernier task, Watt and Morgan found a somewhat larger effect of Gaussian blur, although again, not as large as that predicted by the ideal-observer analysis. Curiously, this result suggests that the relative efficiency of the human observer *increases* as blur increases within this range. This might be explained by the fact that relative efficiency decreases at high intensities for high-contrast stimuli (cf. Figure 20). Blurring the stimuli reduces the peak intensity and, thus, may counteract the effects of the blur on relative efficiency.

The second, rather obvious point is that the effects of blur

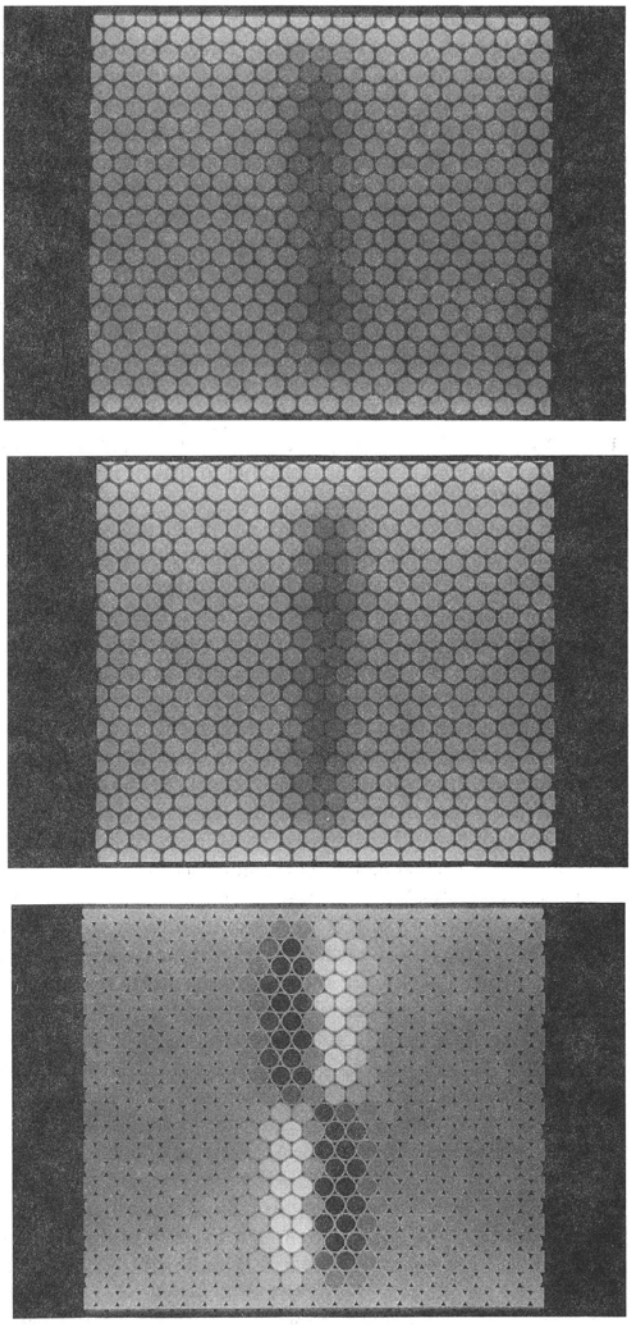


Figure 23. Two-line vernier acuity. (A. Average pattern of photon absorptions in the receptors produced by two 1′-wide, completely dark lines against a gray background at an offset of 12″. B. Average pattern of absorptions produced when the lines are at an opposite offset of 12″. C. Ideal receptive field for discriminating the patterns in A and B.)

predicted by the ideal observer depend on the spatial frequency content of the stimuli. For example, grating acuity, which depends on a very narrow range of high spatial frequencies, is more severely affected by blur than is two-line acuity. Of course, the effects of blur also depend on the form of the blurring function. Thus, there is no general statement that can be made about the effects of blur on the information available at the receptors for discrimination—the ideal observer model must be evaluated to check each case individually.

Context. Another variable that has been examined in the localization paradigms is the context of visual objects within which the task is performed. Westheimer and Hauske (1975) showed that vernier acuity is elevated by flanking lines that are 2–7 min distant from the vernier junction. Similarly, R. A. Williams and Essock (1986) showed that vernier thresholds are most elevated by rectangular background fields that are 3–4 min wide. These effects are superficially similar to those found for increment detection as a function of background diameter (see Figure 17B). Like the increment-detection case, the prediction of the ideal-observer analysis is an increase in vernier threshold with background size, but it does not predict the correct shape for the increase nor the subsequent decrease. Thus, most of these effects can be explained neither by preneural factors nor by changes in the information content of the stimuli; they must be due to neural mechanisms.

Chromatic Discriminations

In all of the experiments analyzed previously, the wavelength distributions (spectral content) of the stimuli were held fixed and were generally broad band enough to ensure that the R and G cones were stimulated about equally well. Under these circumstances, the shapes of the predicted curves depend little on the absorptance spectra of the photopigments or on the distribution and relative numbers of the three classes of cone. I will now consider a few examples involving chromatic discriminations for which the effects of the absorptance spectra and cone distributions play a more important role.

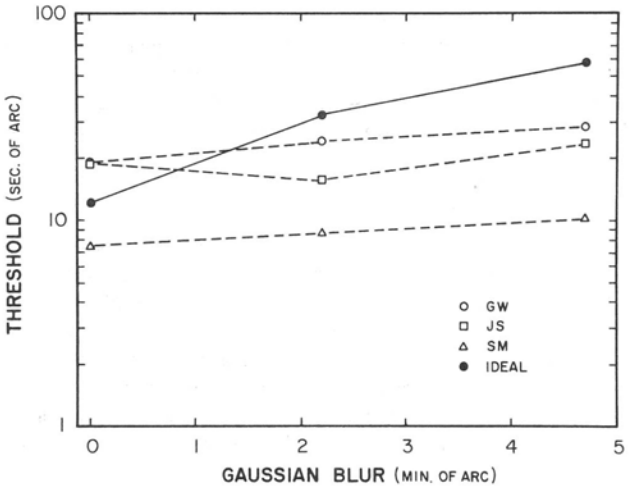


Figure 24. Comparison of real and ideal localization discrimination as a function of Gaussian blur. (Data are stereo discrimination thresholds for thin, 4.5′-long lines presented for 500 ms. Ideal-observer predictions are for monocular localization of a single line of the same dimensions [neutral density value = 3.8 log units]. Data for 3 subjects are from Westheimer & McKee, 1980.)

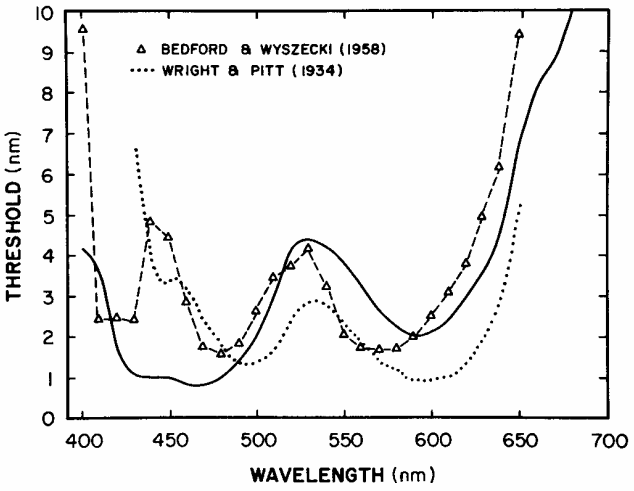


Figure 25. Comparison of ideal and real wavelength-discrimination functions. (Solid line is the ideal-observer prediction translated by the relative-efficiency method for the purpose of comparing shapes. No meaningful neutral density value can be given because the stimulus presentation was continuous. Data are averages across subjects from Bedford & Wyszecki, 1958, and Wright & Pitt, 1934.)

Wavelength discrimination. In the typical wavelength-discrimination experiment, the observer must distinguish between two regions each uniformly illuminated with monochromatic light. The luminances of the regions are held constant, whereas the wavelength of one of the regions is varied. Results obtained with relatively large foveal stimuli (1 and 2° bipartite fields) are shown in Figure 25—the dotted line gives the average data of Wright and Pitt (1934) and the open triangles those of Bedford and Wyszecki (1958). Recall that the derivation and computer implementation of the current ideal-observer model assumes that the stimuli are created by mixing three primaries. Thus, our computer implementation cannot easily generate predictions for chromatic discriminations of monochromatic lights. However, for the case of discrimination of uniform regions, the sensitivity of the ideal observer reduces approximately to the following equation:

$$d'^2 = n_r \Delta R^2/R + n_g \Delta G^2/G + n_b \Delta B^2/B, \qquad (10)$$

where R, G, and B are the mean number of photons absorbed (per receptor) in the red, green, and blue sensitive cones; ΔR, ΔG, and ΔB are the differences in the mean numbers of absorbed photons; and n_r, n_g, and n_b are the numbers of the three cone types lying within the stimulus regions. This equation for ideal-observer performance was originally described by Vos and Walraven (1972), although they did not report its predictions. I have used Equation 10 to generate the solid curve drawn in Figure 25.

The small irregularities seen in the predicted curve should be assigned little significance. They occur because the wavelength-discrimination thresholds predicted for monochromatic lights are very sensitive to small perturbations in the assumed cone-absorption spectra and ocular-transmittance function (Figures 4–6). In fact, because of what is apparently noise in the Estevez absorptance spectra in Figure 6, the predictions in Figure 25 are actually based on the very similar, but smoother spectra of Smith and Pokorny (1972, 1975) as tabled in Wyszecki and Stiles (1982). The predictions with the Estevez spectra appear similar but with larger irregular bumps. The predicted thresholds for the other chromatic discrimination tasks discussed later are not nearly so sensitive to slight measurement errors in the absorption spectra and transmittance function; hence, the Estevez primaries were used.

Figure 25 shows that the ideal observer placed at the level of photon absorption in the receptors displays many (but not all) of the qualitative features of the wavelength-discrimination data. It predicts relatively poor discrimination in the middle wavelengths and at the ends of the spectrum, and best performance in the yellow and blue-green regions of the spectrum. Thus, it appears that these aspects of the wavelength-discrimination data may be largely attributable to loss of information at the receptors.

There are two aspects of the data not predicted by the ideal observer. First, the ideal observer's sensitivity does not decline quickly enough at the extreme long and short wavelengths. Second, most human wavelength-discrimination functions display a second (but smaller) bump in the short wavelength end of the spectrum. There is some hint of this in the predicted curves, but it is not nearly large enough. Some of these discrepancies could be explained by errors in the measurement of the cone-absorption spectra, transmittance function, or in the wavelength-discrimination function at the short-wavelength end of the visible spectrum where the measurements are least reliable. However, it is more likely that they represent losses of information by mechanisms further along the visual pathway.

One possible mechanism that has been widely suggested would be an encoding of receptor outputs into opponent and luminance channels. A wealth of psychophysical data suggests that the receptor outputs are encoded by the visual system into a red/green (e.g., R–G) channel, a blue/yellow (e.g., B–R) channel, and a luminance (e.g., R + G) channel (for a review, see Boynton, 1979; for an example model that predicts a wide range of discrimination data, see Guth, Massof, & Benzschawel, 1980). There is also ample physiological evidence for opponent channels in macaque monkey (de Monasterio & Gouras, 1975; Derrington, Krauskof, & Lennie, 1984; De Valois, 1965). However, an encoding of receptor signals into opponent and achromatic channels by itself is not enough to explain the present discrepancies. For example, Vos and Walraven's (1972) article reported the predictions of a model with simple opponent and achromatic stages following photon absorption. Nonetheless, their predicted curve is quite similar to that predicted by Equation 10. It is smoother, but this is undoubtedly because they further smoothed the absorptance and transmittance functions.

The reason that the two predicted curves appear very similar is because Vos and Walraven's (1972) hypothesized postreceptor stages removed little additional discrimination information. The problem is this: If a hypothesized processing stage can (in principle) be inverted to restore the original input, then it cannot remove any information for discrimination. This is almost true for their model, in which the only additional mechanism for information loss is gain parameters on the opponent and achromatic outputs. In fact, the information loss produced

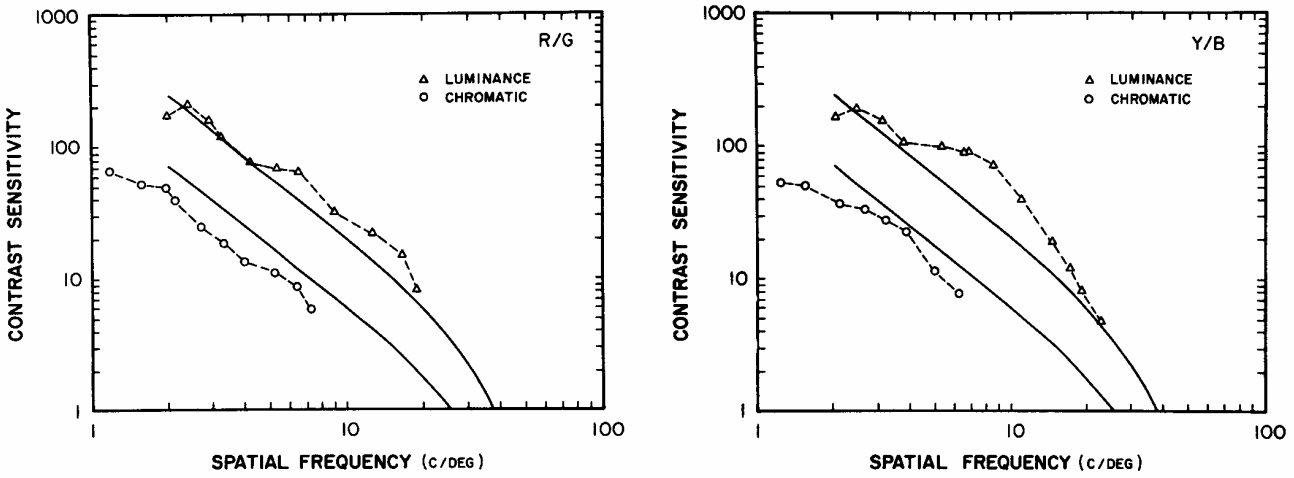


Figure 26. Luminance and chromatic contrast sensitivity at high spatial frequencies for real and ideal observers. (A. Circles are contrast sensitivity for isoluminant red [602 nm]/green [526 nm] gratings, and triangles for green [526 nm] monochromatic gratings. Predictions of the ideal have been translated by the relative-efficiency method for the purpose of comparing shapes. No meaningful neutral density value can be given because number of cycles in the grating were not held fixed. B. Similar data and predictions for isoluminant yellow [577 nm]/blue [470 nm] gratings and monochromatic yellow [577 nm] gratings. Data are from Mullen, 1985.)

by these gain factors could be produced by gain factors on the receptor outputs without ever introducing postreceptor processing.

The differences between ideal and real wavelength discrimination are most likely due to mechanisms such as internal noise and/or nonlinearities operating in conjunction with opponent mechanisms (Guth et al., 1980).

Chromatic contrast sensitivity. Just as the contrast-sensitivity functions of luminance sine-wave gratings are fundamental for understanding the spatial processing of luminance information, measurements of contrast sensitivity for chromatic sine-wave gratings are fundamental for understanding the spatial processing of chromatic information. Within the limits of trichromacy, an arbitrary image can be represented as a sum of sine-wave gratings in three primary colors that may vary in amplitude, frequency, orientation, and phase. Thus, under conditions in which the visual systems behaves in an approximately linear fashion, knowledge of how the visual system responds to chromatic and luminance gratings provides sufficient data to make predictions of its responses to arbitrary stimuli.

Of particular interest in the linear-systems analysis of spatiochromatic processing are measurements of contrast-sensitivity functions for isoluminant chromatic gratings. An isoluminant grating varies only in wavelength composition across space; running a light meter across it produces a constant reading. Isoluminant gratings can be produced by adding together, in opposite phase, two gratings of different wavelength compositions that have the same mean luminance and contrast. (For present purposes, luminance is defined in accordance with the Commission Internationale de l'Eclairage [CIE] standard observer; see Wyszecki & Stiles, 1982; and chromatic contrast is defined as the contrast of the component gratings; see Footnote 5.)

The solid curves in Figure 26A show the predicted luminance and isoluminance CSFs for the spatial configuration of the Banks et al. (1987) study that used gratings that had a spatial extent of 7 cycles in the vertical and horizontal direction. The luminance predictions were for green (526 nm) gratings, the isoluminance predictions for gratings obtained by adding together red (602 nm) and green (526 nm) gratings. Note that when the chromatic contrast was 0.0, the stimulus was a uniform red-plus-green field (it appeared uniform yellow). Figure 26B shows similar predictions for yellow (577 nm) luminance gratings and blue/yellow (470 nm/577 nm) isoluminance gratings. These particular wavelengths are the ones used in the experiments reported by Mullen (1985). Mullen's data for high spatial frequencies are also plotted in the figures. The predicted curves have been translated vertically together, by the relative-efficiency method, for the purpose of comparing shapes.

Mullen's (1985) data were chosen because she took considerable care to control for luminance artifacts produced by chromatic aberration in the eye.[8] Unfortunately, Mullen's (1985) data are not entirely appropriate for comparison with the Banks et al. (1987) study because the number of cycles in the gratings was not held constant across spatial frequency. On the other hand, it is likely that her obtained thresholds would not have been much different had she kept the number of cycles fixed at 7 (or indeed at any number greater than 4; Howell & Hess, 1978). This is supported by the fact that the shape of her luminance CSF agrees quite well with that of Banks et al. In particular, the predicted luminance curves in Figure 26 are almost identical in shape to those fitted to the luminance CSFs in Figure 14; they appear different because of the different scale.

[8] It was assumed that the correction for chromatic aberration was perfect and, hence, that the point-spread functions were identical for all three primaries.

An obvious hypothesis for the poorer human contrast sensitivity with isoluminant gratings is that the chromatically opponent channels produce more spatial summation (e.g., they have receptive fields with larger center regions) than the luminance channel. In particular, because an isoluminant grating would, by definition, produce spatial modulations of output only in the opponent channels, the difference in the CSF shape would be produced if the spatial summation were much larger in the opponent channels. However, it appears safe to conclude from the ideal-observer predictions in Figure 26 that most (although not all) of the difference between the luminance and chromatic CSFs at high spatial frequencies are attributable to preneural mechanisms.

Notice also that the ideal-observer analysis correctly predicts the fact that the difference between chromatic and luminance contrast sensitivity is about the same for the red/green and blue/yellow gratings (at least for these particular chromaticities). This occurs because the red and green cones are doing most of the work, even for the blue/yellow gratings. However, in both cases, the observed difference between the luminance and chromatic CSF is somewhat larger than predicted; this effect may be due to neural mechanisms.

It would also be interesting to compare real and ideal CSFs for stimuli modulated along the tritanopic confusion line, although this case is not considered here. Because of the low density of blue cones and the low transmittance of the ocular media at short wavelengths, the ideal-observer model will predict strongly reduced contrast sensitivity under these conditions.

Mullen (1985) also explored the region between pure luminance and pure chromatic contrast sensitivity. This was done by measuring contrast sensitivity as a function of the luminance ratio of red to green and blue to yellow, while holding total luminance fixed. Figure 27A shows the predictions and data for grating threshold as a function of red/green luminance ratio for high spatial frequencies.[9] Figure 27B shows similar predictions and data as a function of the blue/yellow ratio. Note that the ideal observer does a reasonable job of predicting the effect of luminance ratio, even correctly predicting a shift in the point at which the red/green and blue/yellow functions reach their minimum.

Two factors in the ideal-observer analysis contribute to the predicted difference between luminance and chromatic CSFs. First, and most important, is the high degree of spectral overlap of the middle- and long-wavelength photo pigments and the relatively small number of blue cones. The spectral overlap of the photopigments greatly reduces the effective retinal contrast of the grating. This is illustrated in Figure 28. Figure 28A shows the pattern of photon absorptions produced by a red/green grating of zero contrast (a uniform yellow field). Figure 28B shows the pattern produced by a red/green, 20-c/deg grating at a chromatic contrast of 1.0 (the highest possible). Figure 26 should be compared with Figure 12B, which shows the pattern of photon absorptions produced by a luminance grating of 20 c/deg at a contrast of 1.0. Obviously, there is a drastic difference in the effective contrast at the receptors.

The second, but less important, factor contributing to the difference between luminance and chromatic CSFs is the reduction in receptor-sampling efficiency near the isoluminant point. The best way to understand this is to consider red and green gratings for which the mean luminances produce identical absorptions in the red-sensitive cones. If these two gratings are added in opposite phase, the result is a grating that can only be seen by the green cones, which are a third as numerous as the red cones. This undersampling will produce some information loss at high spatial frequencies. The effect will increase, the closer the mean luminances are to producing the same rate of absorptions in the red (or green) cones.

Two conclusions can be drawn from the ideal-observer analysis. First, the difference between the high-frequency cutoff of luminance and chromatic CSFs should be attributed largely to the overlap of the photopigment absorption spectra. The obvious alternative hypothesis, that the difference is due to differences in spatial summation within the chromatic and luminance channels, appears to be ruled out. Second, the application of ideal-observer analysis to the studies of Banks et al. (1987) and Mullen (1985) suggests that for spatial frequencies above a few cycles per degree, the efficiency of visual processing beyond the level of the photopigments is nearly constant for both luminance and chromatic discrimination (for gratings with a fixed number of cycles). That is, much of the variation in the threshold across all of these conditions may be accounted for by losses of information prior to photoreceptor output. This is consistent with the physiological evidence that the chromatic and luminance information travel initially along the same neural channels (de Monasterio & Gouras, 1975; Derrington et al., 1984; De Valois, 1965).

As with the luminance CSF, some important aspects of the isoluminance CSF are not explained by preneural factors. For instance, the ideal observer is many times more sensitive than the human observer. Possible explanations are the same as the ones discussed earlier in connection with the luminance CSFs.

Inspection of Figure 26 shows that the difference between the luminance and chromatic CSFs is somewhat larger than that predicted by the preneural factors alone. This effect might be explained in part by differences in the gain, nonlinearities, or neural noise between the opponent and luminance channels.

The final important aspect of the data not explained by the ideal-observer analysis occurs at low spatial frequencies, which were excluded from Figure 26. At low spatial frequencies, the chromatic CSF approaches a constant value, whereas the ideal-observer predicts that it should continue to increase. There are of course a number of possible explanations for this leveling-off effect: a shortage of large low-frequency units, relatively greater noise and/or reduced gain in low-frequency units, or a failure at low frequencies to integrate information over a fixed number of cycles.

Recall also that the luminance CSF does not just level off but decreases at low spatial frequencies; in fact, it crosses the chromatic CSF (see Figure 15). This difference between chromatic and luminance CSFs could easily result if cortical luminance and opponent units were formed by combining the re-

[9] The data in Figure 27 differ slightly from those originally reported in Mullen (1985) in that they have been scaled to correct for the loss of contrast in the display that occurred at high spatial frequencies. Mullen (personal communication, November 4, 1987) had applied this correction to the data that are replotted in Figure 26, but not to those in Figure 27.

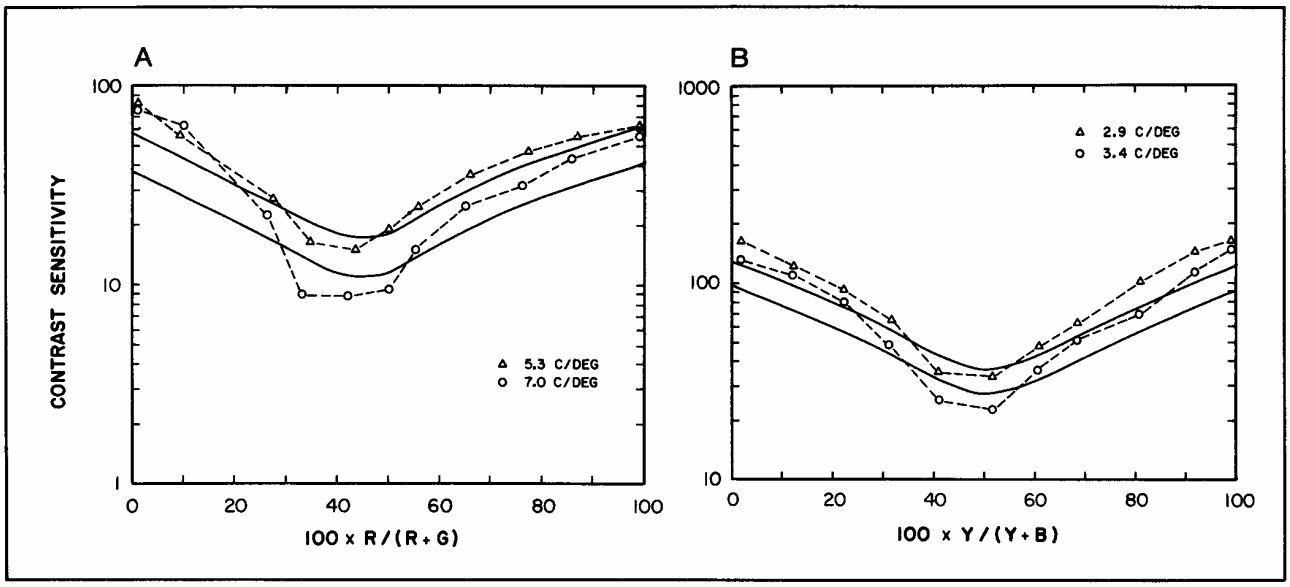


Figure 27. Chromatic contrast sensitivity for real and ideal observers as a function of the ratio of (A) red-to-green and (B) yellow-to-blue luminance while mean luminance was held constant. (Stimulus parameters were the same as for Figure 26B. Ideal-observer predictions were translated by the same amount as in Figure 24A. Data are from Mullen, 1985.)

sponses of LGN units as hypothesized by D'Zmura and Lennie (1986). In particular, because the chromatic units are *not* spatially opponent, they would produce no low-frequency roll-off.

Chromatic vernier acuity. As a final example, consider position discrimination with chromatic stimuli. Morgan and Aiba (1985) have recently measured vernier acuity for green targets in a red, 41-cd/m^2 background, as the intensity of the green target was varied from below to above the point of equal luminance with the background. In the vernier task, the observer must discriminate the relative position of lines or bars that are positioned end-to-end. Figure 29 shows the data for two observers and the prediction of the model. For purposes of comparing predicted and observed shapes, the predicted curve has been translated by the relative-efficiency method. The predictions were generated using Morgan and Aiba's 1-mm pupil size and our best estimate of the spectral distributions of their CRT phosphors. The model shows that the factor of 3 increase in vernier acuity that occurs near the equal luminance point could be explained by loss of information at the receptors. Again, this loss is due mainly to the overlap of the absorption spectra of the red- and green-cone photopigments. At higher luminances of the green vernier targets, the ideal observer would continue to increase in sensitivity, whereas the human observers would surely have leveled off. This same discrepancy between real and ideal performance also occurs in hyperacuity and acuity tasks with pure luminance stimuli (see Figure 22) and may be due to the same mechanisms responsible for producing the transition from square root to Weber's law in intensity discrimination.

Discussion

A fundamental problem in visual science is to link physical and physiological mechanisms to behavioral performance. For complex stimuli, it has been difficult to assess the contributions to performance of even the relatively simple preneural factors of photon noise, preretinal optics, receptor optics, the receptor-lattice arrangement, and the receptor-absorption spectra. The sequential ideal-observer analysis presented here offers a rigorous means of assessing the contributions of such factors in two-alternative discrimination experiments with arbitrary two-dimensional stimuli. Furthermore, the analysis can, in principle, be generalized to other classification tasks. By factoring out the contributions of the preneural factors, vision scientists can, within this approach, also assess the likely contribution of higher level neural mechanisms.

The many examples presented here illustrate the usefulness of ideal-observer analysis in the study of visual discrimination. In particular, it was found that a wide range of phenomena are predicted by the variations in information content of the stimuli at the level of photon absorptions in the receptor photopigments. Those effects most completely predicted by the analysis are those involving detection of stimuli containing relatively high spatial frequencies measured under steady adaptation conditions—luminance contrast sensitivity (Figure 14), the foveal Ricco's area (Figure 17C), and chromatic contrast sensitivity (Figures 26 and 27). The relative importance of stimulus-information content and preneural mechanisms here may be explained because these stimulus conditions minimize the influence of nonlinear visual mechanisms.

Under some other conditions, neural mechanisms are clearly responsible for major effects seen in the data, yet stimulus-information content and preneural mechanisms also appear to be of fundamental importance—vernier acuity as a function of contrast (Figure 19C), resolution versus separation threshold as a function of luminance (Figure 22), wavelength discrimination

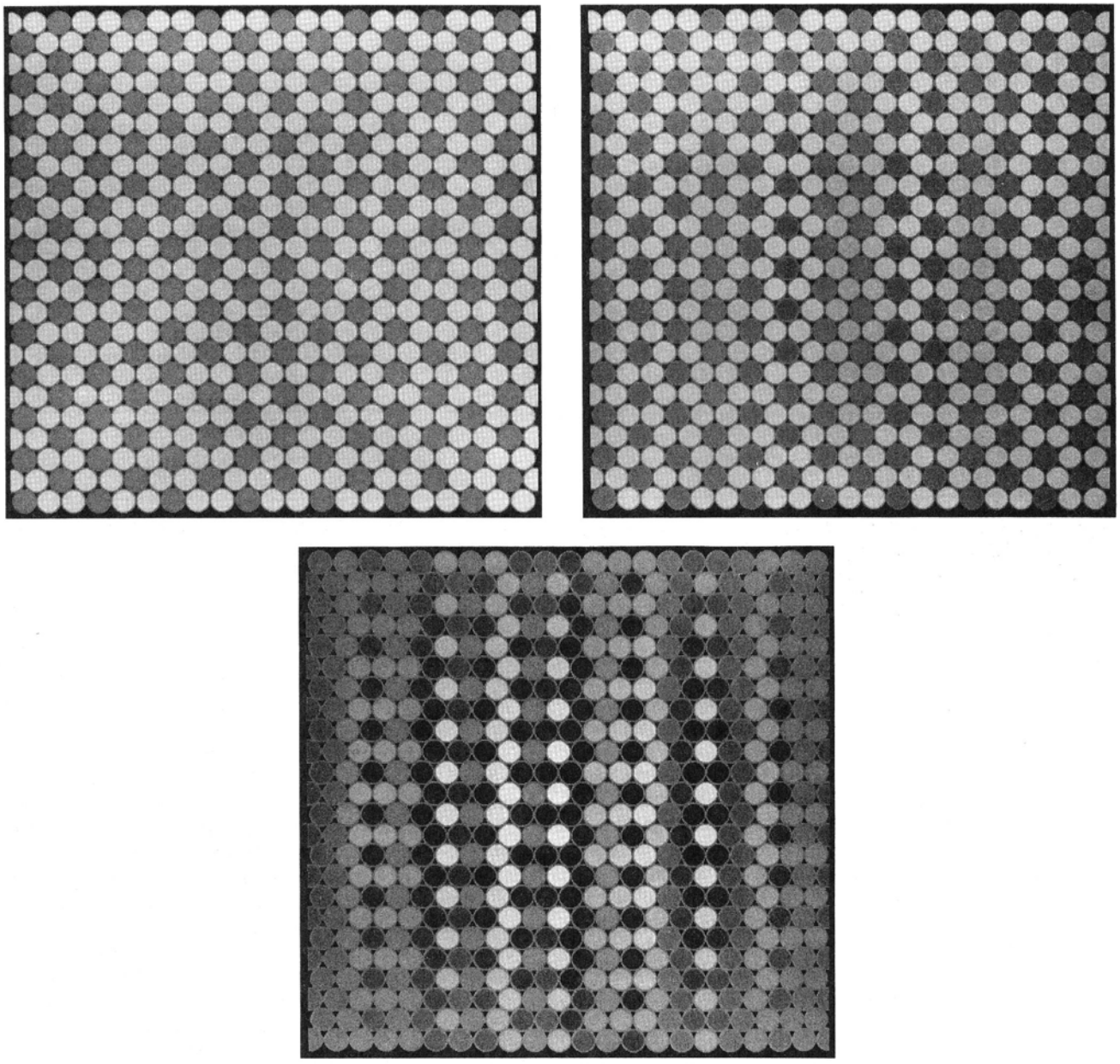


Figure 28. Chromatic contrast detection of isoluminant sine-wave gratings. (A. Average pattern of photon absorptions produced by the sum of uniform red [602 nm] and green [526 nm] fields. B. Average pattern of photon absorptions in the receptors produced by a 100% contrast isoluminant red/green sine-wave grating at 20 c/deg [cycles per degree]. Compare with Figure 12B. C. Ideal receptive field for discriminating the patterns in A and B.)

(Figure 23), and chromatic vernier acuity (Figure 29). The most dramatic predictions concern the relation between resolution and separation (hyperacuity) thresholds. Ideal-observer analysis shows that the physical information for localizing separated objects or features (as in the two-point separation task) is much superior to that for localizing objects that are placed close together (as in the two-point resolution task). Information content and preneural mechanisms may be the major factors responsible for the dramatic differences in performance observed in acuity and hyperacuity tasks.

There are, of course, many discrimination phenomena not predicted by the ideal-observer analysis—luminance (and chromatic) contrast sensitivity at low spatial frequencies (Figure 15), Weber's law for detection threshold as a function of background luminance (Figures 15, 17A), contrast discrimination, and masking (Figure 18), and the effects of base separation, size,

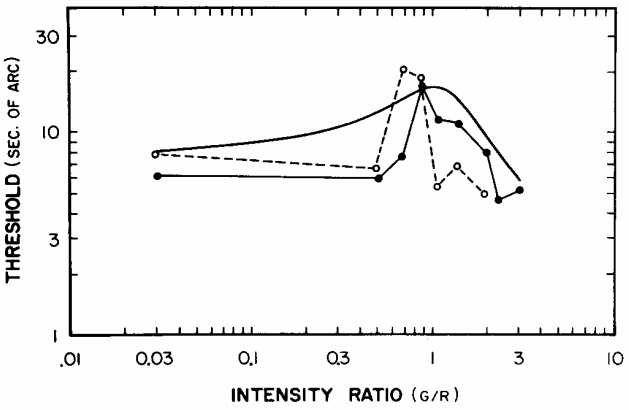


Figure 29. Two-line vernier acuity for green bars in a red background of fixed intensity (R) for various green-bar intensities (G). (Predictions of the ideal observer [solid curve] have been translated by the relative-efficiency method. No meaningful neutral density value can be given because the stimuli remained on until the subject responded. Data for the 2 subjects are from Morgan & Aiba, 1985.)

and blur on location discrimination (Figure 19A and 19B, and Figure 24). The ideal-observer analysis is also informative in such cases because it tells us which phenomena must be explained by neural mechanisms. For example, recall that in the contrast-masking experiment the ideal-observer's performance was essentially unaffected by the contrast or spatial frequency of the masking grating. This helped to validate the contrast-masking experiment as a method for estimating the bandwidths of neural spatial-frequency selective mechanisms. On the other hand, for subthreshold summation experiments, ideal-observer analysis predicts substantial effects of the sort that have been interpreted as evidence for spatial-frequency-selective mechanisms. This leads one to question the validity of using subthreshold summation experiments to estimate spatial-mechanism bandwidths.

It is my belief that an ideal-observer analysis like that developed here should be used as routinely as a light meter in designing and interpreting discrimination experiments. If the physiological mechanisms included in the analysis are sufficiently accurate and complete (which seems likely), then the analysis provides a precise measure of the information contained in the discrimination stimuli at the level of photon absorption in the receptors. In other words, the performance of ideal observer serves as an information meter for the stimuli. Real performance measured using appropriate psychophysical techniques serves as a meter of the information transmitted through the whole system. The difference in the meter readings measures the aspects of performance that must be explained by neural factors. The ideal-observer model should also serve as an appropriate starting point for developing models containing hypothesized neural mechanisms.

As more quantitative information about the neural mechanisms becomes available, it may be possible to develop ideal-observer models that include higher-level mechanisms. As more and more mechanisms are included, the discrepancy between ideal and real observers should diminish.

Relative Importance of the Various Preneural Mechanisms

In this article, the primary concern has been with the combined effects of all of the preneural mechanisms. However, it is worthwhile to consider briefly the relative information loss produced by the individual mechanisms (Figure 1) and, hence, their importance in the overall predictions.

As I have mentioned earlier, eye movements (specifically uncertainty about eye position) may cause some loss of sensitivity in detection tasks with stimuli containing mostly high-spatial-frequency information. Monte Carlo simulations with Equation 7 show that these effects are not large, however, being on the order of a factor of 2. Furthermore, they show that eye movements have a negligible effect in discrimination experiments with suprathreshold targets. This is because position uncertainty is effectively removed when the target stimuli become clearly visible.

The optics, as represented by the optical point-spread function (Figure 3), do produce a substantial amount of information loss. For example, as shown in Figure 13B, it is the most important factor contributing to the shape of the contrast-sensitivity function at high spatial frequencies. The transmittance of the ocular media also produces substantial information loss especially in the short-wavelength end of the spectrum.

In the fovea, the structure of the receptor lattice produces relatively little loss of information over what has already been lost in the optics. This was shown in several ways. First, when an actual foveal lattice (Figure 7 top) was used in place of the perfectly regular lattice of Figure 9A, there was a negligible change in the performance of the ideal observer in all of the tasks we have examined. Second, there was also only a very small improvement in ideal performance if finer lattices (with correspondingly smaller receptors) were used. This is because the foveal lattice is well matched to the optics of the eye. However, peripheral lattices (e.g., Figure 7, middle and bottom) produce a decrease in ideal-observer performance. Third, attempts to find a substantial effect of the approximately triangular arrangement of the foveal lattice failed. Hirsch and Hylton (1984) measured spatial frequency-discrimination thresholds for sine-wave gratings as a function of orientation and found that threshold modulated with a 60° periodicity. They noted that this modulation was correlated with the approximately triangular arrangement of the foveal receptor lattice. (Note that in the perfectly triangular lattice of Figure 9, the receptor density along a line varies with a period of 60° as the line is rotated.) The solid line in the bottom of Figure 30 shows the performance of the ideal observer in this experiment for a spatial frequency of 26 c/deg. This frequency was picked because it exactly divides the lattice spacing and should produce the largest possible effect. As can be seen, there is essentially no effect. The upper curves in Figure 30 show the predictions when the triangular lattice is exploded to 3 times the normal spacing, while keeping the receptor diameter fixed. In this unnatural situation, there is indeed a 60° periodicity. However, the effect depends on the position of the grating with respect to the lattice. When the two are in-phase (upper solid curve), the threshold decreases every 60°; when they are out-of-phase (upper dashed curve), it increases every 60°. Thus, even in this unnatural case, small eye

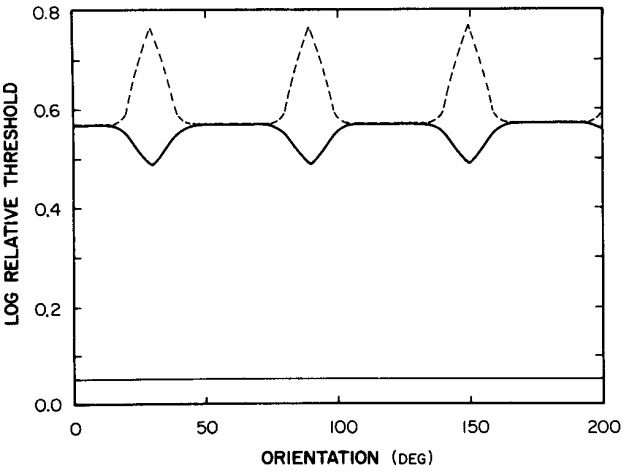


Figure 30. Spatial frequency discrimination of the ideal observer for 26 c/deg (cycles per degree) sine-wave gratings as a function of grating orientation. (Solid line at the bottom is the prediction for a normal foveal receptor lattice. Upper curves are predictions for a lattice exploded to 3 times the normal spacing while maintaining the normal receptor diameter.)

movements would largely wipe out the effect. Any bumps in threshold as a function of orientation will have to be explained by factors other than lattice structure. It should be added that perceptual effects of the 60° symmetry of the foveal lattice have been observed with high-contrast, high-frequency gratings (110 c/deg) produced by laser interferometry (Williams, 1988). However, these effects do not occur under normal viewing conditions.

Receptor optics also produce relatively minor information loss. As shown in Figure 13, the receptor aperture does account for some of the high-frequency roll-off in contrast sensitivity at high spatial frequencies.

The directional selectivity of the cones (the Stiles-Crawford effect) also has some effect. However, it is very small in the central fovea with small pupil diameters (the conditions of most interest here). Its primary effect is to attenuate retinal intensities and to effectively reduce the pupil diameter by a small amount (Stiles-Crawford apodization). Thus, it does not strongly affect the shapes of threshold curves predicted by the ideal observer.

The photopigment spectral sensitivities produce substantial information loss under some circumstances. They play the crucial role in the predicted differences between luminance and chromatic contrast sensitivity and in the predictions for wavelength discrimination and chromatic vernier acuity. They do not result in much information loss in the experiments that used stimuli constructed from monochrome, broad-band light.

Ideal Receptive Fields

In a given task, the maximum-likelihood decision rule of the SDE ideal observer, when placed at the level of the photoreceptors, can be implemented by a single ideal receptive field (RF). Because the weights assigned to the receptors depend only on the ratio of the mean quantum absorptions produced by the two stimuli (see Equation 8), the shapes of the ideal RFs are unaffected by scaling the stimulus intensities. In other words, the shapes of the ideal RFs are not affected by changing the overall intensity of a visual display or the illumination of a scene nor are they affected when the sensitivity of the ideal observer is reduced in the relative-efficiency method of comparing real and ideal performance. The shape of the ideal RF for a particular task is, of course, strongly dependent on the shapes and spectral compositions of the stimuli in the discrimination task.

The neurons in the early levels of the visual system have essentially hard-wired receptive fields of particular shapes. Thus, it may be worth pursuing the hypothesis that the discrimination stimuli for which relative efficiency is high are those whose associated ideal RFs approximate the real RFs found in the macaque (and presumably human) visual system. Barlow (1978) and Watson et al. (1983) have made a similar suggestion with respect to quantum efficiency.

In fact, there seems to be a crude correspondence between high relative (or quantum) efficiency and real RF shape. For example, Figures 19A and 19B show that relative efficiency is greatest in two-line separation discrimination when the lines are short and close together; this is also true for two-line vernier discrimination. Under these conditions, the ideal RFs are of similar spatial extent in all directions, a property that holds for measured RFs up to the level of striate cortex. When the lines are long, the ideal observer still integrates the information perfectly; hence, it uses long ideal RFs that do not have counterparts at the early levels of the visual system.

A similar story holds for contrast detection of sine-wave grating patches. Recall that in grating detection, the visual system appears to integrate effectively over only a small, fixed number of cycles for all but the lowest spatial frequencies (Howell & Hess, 1978). Furthermore, the integration area seems to be of the same spatial extent in all directions. Watson et al. (1983) found that quantum efficiency was highest for grating patches of 4 c/deg that contain around 3 cycles. Crowell et al. (1988) found that relative efficiency was uniformly optimum for patches from 5 to 40 c/deg that contain around 2 cycles. The ideal RFs for such stimuli are like the one shown in Figure 12C. Simple cells in macaque striate cortex have similar receptive fields that also extend over 2–3 cycles, although the number of cycles tends to increase slightly with optimum spatial frequency (De Valois et al., 1982; Movshon et al., 1978).

There are, however, some potential weaknesses in this general line of argument. To begin with, even if the visual system lacks a receptive field that is near optimal for a particular discrimination, it does not imply that any discrimination information is lost. Sampling-theory analysis (Geisler & Hamilton, 1986) shows that the visual cortex could (ignoring the effects of noise) represent faithfully all of the information extracted by the photoreceptors. If a near-ideal RF does not exist in the retina or striate cortex, the discrimination information may still be contained in a *set* of units. However, it might reasonably be argued that distributed information in the cortex is not extracted as efficiently as that concentrated in one or a few striate neurons.

Even if one supposes that hard-wired receptive fields provide the most efficient form of information extraction, there is another potential problem. The ideal observer places its ideal RF

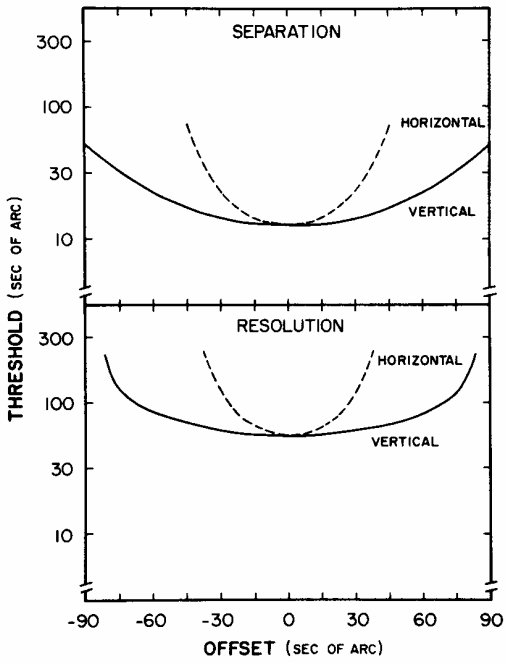


Figure 31. Performance of the ideal receptive fields for two-point separation discrimination and two-point resolution as a function of the offset of the stimuli from optimal alignment with the receptive field. (Dashed lines are for offsets in the horizontal direction; solid lines are for offsets in the vertical direction; thresholds are for changes in the separation of the points in the horizontal direction.)

in exactly the correct position to extract all of the discrimination information. Because there are only a finite number of real RFs, there may be a good chance that on many of the trials in an experiment no appropriate receptive field would be in the right position. If so, the visual system may always have to integrate partial information from many receptive fields in order to achieve its high level of performance. This raises the question of how well positioned an ideal RF has to be in order to achieve the performance level of a human observer. To examine this question, the performance of the ideal observer in the two-point separation and resolution tasks (Figure 22) was calculated as a function of the offset of the stimuli from the optimal position in the ideal RF (e.g., Figure 11C). The results are shown in Figure 31 for offsets of the stimuli in the vertical and horizontal directions. Figure 31 shows that the ideal RF does quite well (would exceed human performance) as long as the stimuli are within approximately ±30 s of the optimal position. (Note that the predictions have been scaled by the relative-efficiency method to match human performance at 0 offset.) The density of small receptive fields in the retina and striate cortex is high enough (Barlow, 1981) to make it plausible that a real RF would be close enough to the "right" position to serve as a near-optimal RF in some tasks. Thus, it is possible that real RF shape may prove useful for predicting relative efficiency, or vice versa, although it is too early to be certain.

Extensions of Ideal-Observer Analysis

The applications of sequential ideal-observer analysis presented here illustrate its usefulness in analyzing the effects of physiological mechanisms and stimulus information content on human discrimination performance. The present analysis is reasonably convincing because it is based on fairly solid knowledge of the physics of light, the optics of the eye, the receptor lattice, and the receptor photopigments. An obvious question is whether this kind of rigorous analysis can be pushed further up the visual pathway. Recent measurements of macaque receptor responses and noise properties (e.g., Baylor et al., 1984; Nunn et al., 1984) may be sufficiently precise and complete to allow inclusion of neural receptor mechanisms in a sequential ideal-observer analysis. This might allow one to compute—to a close approximation—how much information is lost in the photoreceptors for arbitrary discrimination tasks. I am currently exploring this possibility. Unfortunately, accurate knowledge of neuron response and noise properties will necessarily be very slow in coming, so that it may be decades or longer for the next retinal layer to be accurately characterized. The major hindrances to the development of an accurate characterization of the postreceptor layers are the complexity of the responses of the individual neurons and the wide range of the response and noise properties observed across neurons at the same level (e.g., see Rodieck, 1973).

However, it may be possible to carry out a limited ideal-observer analysis on postreceptor stages by restricting the analysis to specific discrimination tasks. One approach that I am currently exploring involves developing and testing methods for applying ideal observers to single-neuron responses (Geisler, 1988). The idea is to measure the repeated response of a single neuron to two stimuli, a and b. Each presentation of a or b produces a spike train. Some fraction of these measured spike trains to each stimulus is used to build a minimodel of the neuron's response to that particular stimulus. These minimodels are then used to design an ideal discriminator that is applied to the remaining fraction of the measured spike trains from each stimulus. If the two minimodels are sufficiently accurate, the performance of the ideal discriminator will be an accurate estimate of the amount of information transmitted by that neuron for the specific discrimination being tested. Computer simulations with synthetic neurons suggest that relatively few trials are necessary to construct a discriminator that is nearly ideal. If a number of neurons at a given neural stage are analyzed in the same discrimination task, it may be possible to estimate the total amount of information for that particular discrimination that is transmitted by that stage.

It should also be possible to apply sequential ideal-observer analysis to other visual systems, those of other species and perhaps to abnormal human visual systems.[10] Once the optics, receptor-lattice structure, and photopigment-absorption spectra have been accurately measured, it is a relatively straightforward matter to incorporate them into the present computer implementation.

[10] Banks and Bennett (1988) have applied the present ideal-observer analysis to the visual system of human neonates.

References

Albrecht, D. G., & Hamilton, D. B. (1982). Striate cortex of monkey and cat: Contrast response function. *Journal of Neurophysiology, 48,* 217–237.

Andrews, D. P., Butcher, A. K., & Buckley, B. R. (1973). Acuities for spatial arrangement in line figures: Human and ideal observers compared. *Vision Research, 13,* 599–620.

Arnulf, A., & Dupuy, O. (1960). La transmission des contrastes par le système optique de l'oeil et les seuils de contrastes rétinines [The transmission of contrast by the optical system of the eye and the contrast threshold]. *Comptes Rendus Hebdomadaires des Séances de l'Academie des Sciences, Paris, 250,* 2757–2759.

Banks, M. S., & Bennett, P. J. (1988). Optical and photoreceptor immaturities limit the spatial and chromatic vision of human neonates. *Journal of the Optical Society of America A: Optics and Image Science, 5,* 2059–2079.

Banks, M. S., Geisler, W. S., & Bennett, P. J. (1987). The physical limits of grating visibility. *Vision Research, 27,* 1915–1924.

Barlow, H. B. (1956). Retinal noise and absolute threshold. *Journal of the Optical Society of America, 46,* 634–639.

Barlow, H. B. (1957). Increment thresholds at low intensities considered as signal/noise discriminations. *Journal of Physiology* (London), *136,* 469–488.

Barlow, H. B. (1958a). Intrinsic noise of cones. In *National physical laboratory symposium on visual problems of colour* (Vol. 2, pp. 617–630). London: H. M. Stationery Office.

Barlow, H. B. (1958b). Temporal and spatial summation in human vision at different background intensities. *Journal of Physiology* (London), *141,* 337–350.

Barlow, H. B. (1962). Measurements of the quantum efficiency of discrimination in human scotopic vision. *Journal of Physiology* (London), *160,* 169–188.

Barlow, H. B. (1977). Retinal and central factors in human vision limited by noise. In H. B. Barlow & P. Fatt (Eds.), *Vertebrate photoreception* (pp. 337–358). London: Academic Press.

Barlow, H. B. (1978). The efficiency of detecting changes of density in random dot patterns. *Vision Research, 18,* 637–650.

Barlow, H. B. (1979). Reconstructing the visual image in space and time. *Nature, London, 279,* 189–190.

Barlow, H. B. (1981). Critical limiting factors in the design of the eye and visual cortex. *Proceedings of the Royal Society of London* (B), *212,* 1–34.

Baylor, D. A., Nunn, B. J., & Schnapf, J. L. (1984). The photocurrent, noise and spectral sensitivity of rods of the monkey, *Macaca fascicularis. Journal of Physiology* (London), *357,* 575–607.

Bedford, R. E., & Wyszecki, G. (1957). Axial chromatic aberration of the human eye. *Journal of the Optical Society of America, 47,* 564–565.

Bedford, R. E., & Wyszecki, G. W. (1958). Wavelength discrimination for point sources. *Journal of the Optical Society of America, 48,* 129–135.

Blackwell, H. R. (1963). Neural theories of simple visual discriminations. *Journal of the Optical Society of America, 53,* 129–160.

Borwein, B., Borwein, D., Medeiros, J., & McGowan, J. W. (1980). The ultrastructure of monkey foveal photoreceptors, with special reference to the structure, shape, size, and spacing of the foveal cones. *American Journal of Anatomy, 159,* 125–146.

Bossomaier, T. R. J., Snyder, A. W., & Hughes, A. (1985). Irregularity and aliasing: Solution? *Vision Research, 25,* 145–147.

Bowmaker, J. K., Dartnall, H. J. A., Lythgoe, J. N., & Mollon, J. D. (1978). The visual pigments of rods and cones in the rhesus monkey, *Macaca mulatta. Journal of Physiology* (London), *274,* 329–348.

Boycott, B. B., & Dowling, J. E. (1969). Organization of the primate retina: Light microscopy. *Philosophical Transactions of the Royal Society of London* (B), *255,* 109–184.

Boynton, R. M. (1979). *Human color vision.* New York: Holt, Rinehart & Winston.

Bracewell, R. N. (1978). *The Fourier transform and its applications* (2nd ed.). New York: McGraw-Hill.

Bradley, A., & Ohzawa, I. (1986). A comparison of contrast detection and discrimination. *Vision Research, 26,* 991–997.

Brindley, G. S. (1970). *Physiology of the retina and the visual pathway* (2nd ed.). London: Edward Arnold.

Brown, P. K., & Wald, G. (1964). Visual pigments in single rods and cones of the human retina. *Science, 144,* 45–52.

Burgess, A. E., Wagner, R. F., Jennings, R. J., & Barlow, H. B. (1981). Efficiency of human visual signal discrimination. *Science, 214,* 93–94.

Buss, C. M., Hayhoe, M. M., & Stromeyer, C. F. III. (1982). Lateral interactions in the control of visual sensitivity. *Vision Research, 22,* 693–709.

Byram, G. M. (1944). The physical and photochemical basis of visual resolving power: Part 2. Visual acuity and the photochemistry of the retina. *Journal of the Optical Society of America, 34,* 718–738.

Campbell, F. W., & Green, D. G. (1965). Optical and retinal factors affecting visual resolution. *Journal of Physiology* (London), *181,* 576–593.

Campbell, F. W., & Gubisch, R. W. (1966). Optical quality of the human eye. *Journal of Physiology* (London), *186,* 558–578.

Campbell, F. W., Robson, J. G., & Westheimer, G. (1959). Fluctuations of accommodation under steady viewing conditions. *Journal of Physiology* (London), *145,* 579–594.

Chandler, J. P. (1969). STEPIT—Finds local minima of a smooth function of several parameters. *Behavioral Science, 14,* 81–82.

Charman, W. N., & Jennings, J. A. M. (1976). Objective measurements of the longitudinal chromatic aberration of the human eye. *Vision Research, 16,* 999–1005.

Charman, W. N., & Tucker, J. (1977). Dependence of accommodation response on the spatial frequency spectrum of the observed object. *Vision Research, 17,* 129–139.

Cohn, T. E. (1976). Quantum fluctuation limit in foveal vision. *Vision Research, 16,* 573–579.

Cohn, T. E., & Lasley, D. J. (1974). Detectability of a luminance increment: Effect of spatial uncertainty. *Journal of the Optical Society of America, 64,* 1715–1719.

Cohn, T. E., & Lasley, D. J. (1986). Visual sensitivity. *Annual Review of Psychology, 37,* 495–521.

Cohn, T. E., & Wardlaw, J. C. (1985). Effect of large spatial uncertainty on foveal luminance increment detectability. *Journal of the Optical Society of America A: Optics and Image Science, 2,* 820–825.

Cornsweet, T. N. (1970). *Visual perception.* New York: Academic Press.

Crick, F. H. C., Marr, D. C., & Poggio, T. (1981). An information-processing approach to understanding the visual cortex. In F. O. Schmidt & F. G. Worden (Eds.), *The organization of the cerebral cortex* (pp. 505–533). Cambridge, MA: MIT Press.

Crowell, J., Banks, M. S., Anderson, S., & Geisler, W. S. (1988). Physical limits of grating visibility: Fovea and periphery. *Supplement to Investigative Ophthalmology and Visual Science, 29,* 139.

Dartnall, H. J. A., Bowmaker, J. K., & Mollon, J. D. (1983). Human visual pigments: Microspectrophotometric results from the eyes of seven persons. *Proceedings of the Royal Society of London* (B), *220,* 115–130.

Davila, K. D., & Geisler, W. S. (1987). Pre-neural factors predict Ricco's area in the fovea. *Supplement to Investigative Ophthalmology and Visual Science, 28,* 357.

de Groot, S. G., & Gebhard, S. W. (1952). Pupil size as determined by adapting luminance. *Journal of the Optical Society of America, 42,* 492–495.

De Monasterio, F. M., & Gouras, P. (1975). Functional properties of ganglion cells of the rhesus monkey retina. *Journal of Physiology* (London), *251,* 167–195.

De Monasterio, F. M., McCrane, E. P., Newlander, J. K., & Schein, S. J. (1985). Density profile of blue-sensitive cones along the horizontal meridian of macaque retina. *Investigative Ophthalmology and Visual Science, 26,* 289–302.

DeMott, D. W., & Boynton, R. M. (1958). Retinal distribution of entoptic stray light. *Journal of the Optical Society of America, 48,* 13–22.

Derrington, A. M., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology* (London), *357,* 241–265.

Derrington, A. M., & Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *Journal of Physiology* (London), *357,* 219–240.

De Valois, R. L. (1965). Analysis and coding of color vision in the primate visual system. In L. Frisch (Ed.), *Cold Spring Harbor Symposia on Quantitative Biology: Vol.* 30. *Sensory receptors* (pp. 567–579). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory of Quantitative Biology.

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research, 22,* 545–559.

De Valois, R. L., Morgan, H., & Snodderly, D. M. (1974). Psychophysical studies of monkey vision: 3. Spatial luminance contrast sensitivity tests of macaque and human observers. *Vision Research, 14,* 75–81.

de Vries, H. (1943). The quantum character of light and its bearing upon threshold of vision, the differential sensitivity and visual acuity of the eye. *Physica, 10,* 553–564.

Ditchburn, R. W. (1973). *Eye-movements and visual perception.* Oxford, England: Clarendon Press.

Ditchburn, R. W., & Ginsborg, B. L. (1952). Vision with a stabilized retinal image. *Nature, 170,* 36–37.

D'Zmura, M., & Lennie, P. (1986). Mechanisms of color constancy. *Journal of the Optical Society of America A: Optics and Image Science, 3,* 1662–1672.

Enoch, J. M., & Stiles, W. S. (1961). The colour change of monochromatic light with retinal angle of incidence. *Optica Acta, 8,* 329–358.

Estevez, O. (1979). *On the fundamental data-base of normal and dichromatic color vision.* Unpublished doctoral dissertation, University of Amsterdam, Amsterdam, the Netherlands.

Feller, W. (1971). *An introduction to probability theory and its applications* (Vol. 2, 2nd ed.). New York: Wiley.

Flamant, F. (1955). Étude de la répartition de lumière dans l'image rétinienne d'une fente [Study of the distribution of light in the retinal image of a line]. *Revue D'Optique Theorique Et Instumentale, 34,* 433–459.

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

Gaskill, J. D. (1978). *Linear systems, Fourier transforms, and optics.* New York: Wiley.

Geisler, W. S. (1978). Evidence for the equivalent-background hypothesis in cones. *Vision Research, 19,* 799–805.

Geisler, W. S. (1983). Mechanisms of visual sensitivity: Backgrounds and early dark adaptation. *Vision Research, 23,* 1423–1432.

Geisler, W. S. (1984). Physical limits of acuity and hyperacuity. *Journal of the Optical Society of America A: Optics and Image Science, 1,* 775–782.

Geisler, W. S. (1987a). Ideal-observer analysis of visual discrimination. In Committee on Vision (Eds.), *Frontiers of visual science: Proceedings of the 1985 symposium* (pp. 17–31). Washington DC: National Academy Press.

Geisler, W. S. (1987b). SDE software package. *Vision Research, 27*(2), I.

Geisler, W. S. (1988). Ideal-observer analysis of single neuron responses. *Supplement to Investigative Ophthalmology and Visual Science, 29,* 297.

Geisler, W. S., & Davila, K. D. (1985). Ideal discriminators in spatial vision: Two-point stimuli. *Journal of the Optical Society of America A: Optics and Image Science, 2,* 1483–1497.

Geisler, W. S., & Davila, K. D. (1987). Can receptive-field properties explain hyperacuity thresholds? *Supplement to Investigative Ophthalmology and Visual Science, 28,* 137.

Geisler, W. S., & Hamilton, D. B. (1986). Sampling theory analysis of spatial vision. *Journal of the Optical Society of America A: Optics and Image Science, 3,* 62–70.

Goodman, J. W. (1968). *Introduction to Fourier optics.* New York: McGraw-Hill.

Graham, N., Robson, J. G., & Nachmias, J. (1978). Grating summation in fovea and periphery. *Vision Research, 18,* 815–825.

Green, D. M., & Swets, J. A. (1974). *Signal detection theory and psychophysics.* New York: Krieger.

Guth, S. L., Massof, R. W., & Benzschawel, T. (1980). Vector model for normal and dichromatic color vision. *Journal of the Optical Society of America, 70,* 197–212.

Harwerth, R. S., Boltz, R. L., & Smith, E. L. III. (1980). Psychophysical evidence for sustained and transient channels in the monkey visual system. *Vision Research, 20,* 15–22.

Heath, G. G. (1956). The influence of visual acuity on accommodative responses of the eye. *American Journal of Optometry and Archives of American Academy of Optometry, 33,* 513–524.

Helmholtz, H. (1925). *Physiological optics* (3rd ed., Vols. 1–3; J. P. Southall, Ed. & trans.). Rochester, NY: Optical Society of America. (Original work published 1866)

Helstrom, C. W. (1964). The detection and resolution of optical signals. *IEEE Transactions on Information Theory,* IT-*10.* 275–287.

Hirsch, J., & Hylton, R. (1982). Limits of spatial-frequency discrimination as evidence of neural interpolation. *Journal of the Optical Society of America, 72,* 1367–1374.

Hirsch, J., & Hylton, R. (1984). Orientation dependence of visual hyperacuity contains a component with hexagonal symmetry. *Journal of the Optical Society of America A: Optics and Image Science, 1,* 300–308.

Howell, E. R., & Hess, R. F. (1978). The functional area for summation to threshold for sinusoidal gratings. *Vision Research, 18,* 369–374.

Howland, H. C., & Howland, B. (1977). A subjective method for the measurement of monochromatic aberrations of the eye. *Journal of the Optical Society of America, 67,* 1508–1518.

Hsia, Y. (1965). Photochemistry of vision. In C. H. Graham (Ed.), *Vision and visual perception* (pp. 132–153). New York: Wiley.

Ingling, C. R., Jr., & Martinez-Uriegas, E. (1983). Simple-opponent receptive fields are asymmetrical: G-cone centers predominate. *Journal of the Optical Society of America, 73,* 1527–1532.

Jennings, J. A. M., & Charman, W. N. (1981). Off-axis image quality in the human eye. *Vision Research, 21,* 445–455.

Kelly, D. H. (1972). Adaptation effects on spatio-temporal sine-wave thresholds. *Vision Research, 12,* 89–101.

Kelly, D. H. (1977). Visual contrast sensitivity. *Optica Acta, 24,* 107–129.

Kersten, D. (1984). Spatial summation in visual noise. *Vision Research, 24,* 1977–1990.

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

Koenderink, J. J., Bouman, M. A., Bueno de Mesquita, A. E., & Slappendel, S. (1978). Perimetry of contrast detection thresholds of moving spatial sine wave patterns: 3. The target extent as a sensitivity controlling parameter. *Journal of the Optical Society of America, 68,* 854–860.

Krauskopf, J., Cornsweet, T. N., & Riggs, L. A. (1960). Analysis of eye movements during monocular and binocular fixation. *Journal of the Optical Society of America, 50,* 572–578.

Lamar, E. S., Hecht, S., Shlaer, S., & Hendley, C. D. (1947). Size, shape, and contrast in detection of targets by daylight vision: 1. Data and analytical description. *Journal of the Optical Society of America, 37,* 531–545.

Legge, G. E. (1981). A power law for contrast discrimination. *Vision Research, 21,* 457–467.

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

Le Grand, Y. (1957). *Light, colour, and vision.* New York: Wiley.

Leshowitz, B., Taub, H. B., & Raab, D. H. (1968). Visual detection of signals in the presence of continuous and pulsed backgrounds. *Perception and Psychophysics, 4,* 207–213.

Ludvigh, E. (1953). Direction sense of the eye. *American Journal of Ophthalmology, 36,* 139–143.

MacLeod, D. I. A. (1978). Visual sensitivity. *Annual Review of Psychology, 29,* 613–645.

MacLeod, D. I. A., Williams, D. R., & Makous, W. (1985). Difference frequency gratings above the resolution limit. *Supplement to Investigative Ophthalmology and Vision Science, 26,* 11.

Maloney, L. T. (1988). Spatially irregular sampling in combination with rigid movements of the sampling array. *Supplement to Investigative Ophthalmology and Vision Science, 29,* 58.

Marc, R. E., & Sperling, H. G. (1977). Chromatic organization of primate cones. *Science, 196,* 454–456.

Marks, W. B., Dobelle, W. H., & MacNichol Jr., E. F. (1964). Visual pigments of single primate cones. *Science, 143,* 1181–1183.

Miller, W. H. (1979). Ocular optical filtering. In H. Autrum (Ed.), *Handbook of sensory physiology: Vol. VII/6A. Comparative physiology and evolution of vision in invertebrates A: Invertebrate photoreceptors* (pp. 69–143). New York: Springer-Verlag.

Miller, W. H., & Bernard, G. D. (1983). Averaging over the foveal receptor aperture curtails aliasing. *Vision Research, 23,* 1365–1369.

Morgan, M. W. (1944). Accommodation and its relationship to convergence. *American Journal of Optometry and Archives of American Academy of Optometry, 21,* 183–195.

Morgan, M. J., & Aiba, T. S. (1985). Positional acuity with chromatic stimuli. *Vision Research, 25,* 689–695.

Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *Journal of Physiology* (London), *283,* 53–77.

Mullen, K. T. (1985). The contrast sensitivity of human colour vision to red-green and blue–yellow chromatic gratings. *Journal of Physiology* (London), *359,* 381–400.

Nachmias, J. (1959). Two-dimensional motion of the retinal image during monocular fixation. *Journal of the Optical Society of America, 49,* 901–908.

Nachmias, J., & Kocher, E. C. (1970). Visual detection and discrimination of luminance increments. *Journal of the Optical Society of America, 60,* 382–389.

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

Navarro, R., Santamaría, J., & Bescós, J. (1985). Accommodation-dependent model of the human eye with aspherics. *Journal of the Optical Society of America A: Optics and Image Science, 2,* 1273–1281.

Nunn, B. J., Schnapf, J. L., & Baylor, D. A. (1984). Spectral sensitivity of single cones in the retina of *Macaca fascicularis. Nature, 309,* 264–266.

Ohzu, H., & Enoch, J. M. (1972). Optical modulation by the isolated human fovea. *Vision Research, 12,* 245–251.

Osterberg, G. (1935). Topography of the layer of rods and cones in the human retina. *Acta Ophthalmologica, Kobenhaven, and Supplementum* (Suppl. 6), 1–102.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America A: Optics and Image Science, 2,* 1508–1532.

Perry, V. H., & Cowey, A. (1985). The ganglion cell and cone distributions in the monkey's retina: Implications for the central magnification factors. *Vision Research, 25,* 1795–1810.

Peterson, W. W., Birdsall, T. G., & Fox, W. C. (1954). The theory of signal detectability. *Transactions of the Institute of Radio Engineers, Professional Group on Information Theory, 4,* 171–212.

Polyak, S. L. (1941). *The retina.* Chicago: University of Chicago Press.

Quick, R. F., Hammerly, J. R., & Reichert, T. A. (1976). The absence of measureable "critical band" at low suprathreshold contrasts. *Vision Research, 16,* 351–356.

Rattle, J. D. (1969). Effect of target size on monocular fixation. *Optica Acta, 16,* 183–192.

Rempt, F., Hoogerheide, J., & Hoogenboom, W. P. H. (1971). Peripheral retinoscopy and the skiagram. *Ophthalmologica, 162,* 1–10.

Riggs, L. A. (1965). Visual acuity. In C. H. Graham (Ed.), *Vision and visual perception* (pp. 321–349). New York: Wiley.

Riggs, L. A., Ratliff, F., Cornsweet, J. C., & Cornsweet, T. N. (1953). The disappearance of steadily fixated visual test objects. *Journal of the Optical Society of America, 43,* 495–501.

Robson, J. G., & Enroth-Cugell, C. (1978). Light distribution in the cat's retinal image. *Vision Research, 18,* 159–173.

Robson, J. G., & Graham, N. (1981). Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Research, 21,* 409–418.

Rodieck, R. W. (1973). *The vertebrate retina.* San Francisco: Freeman.

Röhler, R., Miller, U., & Aberl, M. (1969). Zur Messung der Modulationsübertragunsfunktion des lebenden menschlichen Auges im reflektierten Licht [Measurement of the modulation transfer function of the eye by means of reflected light]. *Vision Research, 9,* 407–428.

Rose, A. (1942). The relative sensitivities of television pickup tubes, photographic film, and the human eye. *Proceedings of the Institute of Radio Engineers, 30,* 293–300.

Rose, A. (1948). The sensitivity performance of the human eye on an absolute scale. *Journal of the Optical Society of America, 38,* 196–208.

Rosenfeld, A., & Kak, A. C. (1982). *Digital picture processing,* Orlando, FL: Academic Press.

Rushton, W. A. H., & Gubisch, R. W. (1966). Glare: Its measurement by cone thresholds and by the bleaching of cone pigments. *Journal of the Optical Society of America, 56,* 104–110.

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

Saleh, B. E. A. (1982). Optical information processing and the human visual system. In H. Stark (Ed.), *Applications of optical Fourier transforms* (pp. 431–463). New York: Academic.

Shapley, R., & Enroth-Cugell, C. (1984). Visual adaptation and retinal gain controls. *Progress in Retinal Research, 3,* 263–346.

Sidman, R. L. (1957). The structure and concentration of solids in photoreceptor cells studied by refractometry and interference microscopy. *Journal of Biophysical and Biochemical Cytology, 3,* 15–30.

Smith, V. C., & Pokorny, J. (1972). Spectral sensitivity of color-blind observers and the cone photopigments. *Vision Research, 12,* 2059–2071.

Smith, V. C., & Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research, 15,* 161–171.

Snyder, D. L. (1975). *Random point processes.* New York: Wiley.

Snyder, A. W., & Miller, W. H. (1977). Photoreceptor diameter and

spacing for highest resolving power. *Journal of the Optical Society of America, 67,* 696–698.

Snyder, A. W., & Pask, C. (1973). The Stiles-Crawford effect: Explanation and consequences. *Vision Research, 13,* 1115–1137.

Snyder, A. W., & Srinivasan, M. V. (1979). Human psychophysics: Functional interpretation for contrast sensitivity vs. spatial frequency curve. *Biological Cybernetics, 32,* 9–17.

Spring, K. H., & Stiles, W. S. (1948). Variation of pupil size with change in theangle at which the lightstimulus strikes the retina. *British Journal of Ophthalmology, 32,* 340–346.

Steinman, R. M. (1965). Effect of target size, luminance, and color on monocular fixation. *Journal of the Optical Society of America, 55,* 1158–1165.

Stiles, W. S. (1929). The scattering theory of the effect of glare on the brightness difference threshold. *Proceedings of the Royal Society of London* (B), *105,* 131–146.

Stiles, W. S., & Crawford, B. H. (1933). The luminous efficiency of rays entering the eye pupil at different points. *Proceedings of the Royal Society of London* (B), *112,* 428–450.

Stromeyer, C. F., III, & Julesz, B. (1972). Spatial frequency masking in vision: Critical bands and spread of masking. *Journal of the Optical Society of America, 62,* 1221–1232.

Sullivan, G. D., Oatley, K., & Sutherland, N. S. (1972). Vernier acuity as affected by target length and separation. *Perception and Psychophysics, 12,* 438–444.

Tanner, W. P., Jr. (1961). Physiological implications of psychophysical data. *Annals of the New York Academy of Sciences, 89,* 752–765.

Tanner, W. P., Jr., & Birdsall, T. G. (1958). Definitions of d' and η as psychophysical measures. *Journal of the Acoustical Society of America, 30,* 922–928.

Tanner, W. P., Jr., & Clark-Jones, R. C. (1960). The ideal sensor system as approached through statistical decision theory and the theory of signal detectability. In A. Morris & E. P. Horne (Eds.), *Visual search techniques* (NAS–NRC Publication No. 712, pp. 59–68). Washington, DC: National Academy of Sciences, Armed Forces-National Research Council Committee on Vision.

Teich, M. C., Prucnal, P. R., Vannucci, G., Breton, M. E., & McGill, W. J. (1982). Multiplication noise in the human visual system at threshold: 3. The role of non-Poisson quantum fluctuations. *Biological Cybernetics, 44,* 157–165.

Tolhurst, D. J., Movshon, J. A., & Dean, A. F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Research, 23,* 775–785.

van Meeteren, A. (1974). Calculations on the optical modulation transfer function of the human eye for white light. *Optica Acta, 21,* 395–412.

van Meeteren, A., & Barlow, H. B. (1981). The statistical efficiency for detecting sinusoidal modulation of average dot density in random figures. *Vision Research, 21,* 765–777.

Van Meter, D., & Middleton, D. (1954). Modern statistical approaches to reception in communication theory. *Transactions of the Institute of Radio Engineers, Professional Group on Information Theory, 4,* 119–141.

Van Nes, F. L., & Bouman, M. A. (1967). Spatial modulation transfer in the human eye. *Journal of the Optical Society of America, 57,* 401–406.

Van Trees, H. L. (1968). *Detection, estimation, and modulation theory.* New York: Wiley.

Vos, J. J. (1963). Contribution of the fundus oculi to entoptic scatter. *Journal of the Optical Society of America, 53,* 1449–1451.

Vos, J. J. (1984). Disability glare: A state of the art report. *Journal of the Commission Internationale de l'Eclairage, 3,* 39–53.

Vos, J. J., & Walraven, P. L. (1972). An analytical description of the line element in the zone-fluctuation model of color vision: 1. Basic concepts. *Vision Research, 12,* 1327–1344.

Wald, G., & Griffin, D. R. (1947). The change in refractive power of the human eye in dim and bright light. *Journal of the Optical Society of America, 37,* 321–336.

Walraven, P. L. (1974). A closer look at the tritanopic convergence point. *Vision Research, 14,* 1339–1343.

Walsh, G., Charman, W. N., & Howland, H. C. (1984). Objective technique for the determination of monochromatic aberrations of the human eye. *Journal of the Optical Society of America A: Optics and Image Science, 1,* 987–992.

Watson, A. B. (1982). Summation of grating patches indicates many types of detector at one retinal location. *Vision Research, 22,* 17–25.

Watson, A. B. (1983). Detection and recognition of simple spatial forms. In O. J. Braddick & A. C. Sleigh (Eds.), Physical and biological processing of images (pp. 110–114). Berlin: Springer-Verlag.

Watson, A. B., Barlow, H. B., & Robson, J. G. (1983). What does the eye see best? *Nature, 302,* 419–422.

Watt, R. J., & Morgan, M. J. (1983). The recognition and representation of edge blur: Evidence for spatial primitives in human vision. *Vision Research, 23,* 1465–1477.

Werblin, F. S., & Copenhagen, D. R. (1974). Control of retinal sensitivity: 3. Lateral interactions at the inner plexiform layer. *Journal of General Physiology, 63,* 88–110.

Westheimer, G. (1960). Modulation thresholds for sinusoidal light distributions on the retina. *Journal of Physiology* (London), *152,* 67–74.

Westheimer, G. (1967a). Dependence of the magnitude of the Stiles-Crawford effect on retinal location. *Journal of Physiology* (London), *192,* 309–315.

Westheimer, G. (1967b). Spatial interaction in human cone vision. *Journal of Physiology* (London), *190,* 139–154.

Westheimer, G. (1975). Visual acuity and hyperacuity. *Investigative Ophthalmology, 14,* 570–572.

Westheimer, G. (1979). The spatial sense of the eye. *Investigative Ophthalmology and Visual Science, 18,* 893–912.

Westheimer, G. (1981). Visual hyperacuity. In D. Ottoson (Ed.), *Progress in sensory physiology* (Vol. 1, pp. 1–30). Berlin: Springer-Verlag.

Westheimer, G., & Campbell, F. W. (1962). Light distribution in the image formed by the living human eye. *Journal of the Optical Society of America, 52,* 1040–1045.

Westheimer, G., & Hauske, G. (1975). Temporal and spatial interference with vernier acuity. *Vision Research, 15,* 1137–1141.

Westheimer, G., & McKee, S. P. (1977a). Integration regions for visual hyperacuity. *Vision Research, 17,* 89–93.

Westheimer, G., & McKee, S. P. (1977b). Spatial configurations for visual hyperacuity. *Vision Research, 17,* 941–947.

Westheimer, G., & McKee, S. P. (1980). Stereoscopic acuity with defocused and spatially filtered retinal images. *Journal of the Optical Society of America, 70,* 772–778.

Whittle, P., & Swanston, M. T. (1974). Luminance discrimination of separated flashes: The effect of background luminance and the shapes of t.v.i. curves. *Vision Research, 14,* 713–719.

Wilcox, W. W., & Purdy, D. M. (1933). Visual acuity and its physiological basis. *British Journal of Psychology, 23,* 233–261.

Williams, D. R. (1985a). Aliasing in human foveal vision. *Vision Research, 25,* 195–205.

Williams, D. R. (1985b). Visibility of interference fringes near the resolution limit. *Journal of the Optical Society of America A: Optics and Image Science, 2,* 1087–1093.

Williams, D. R. (1988). Topography of the foveal cone mosaic in the living human eye. *Vision Research, 28,* 433–454.

Williams, D. R., MacLeod, D. I. A., & Hayhoe, M. M. (1981). Punctate sensitivity of the blue-sensitive mechanism. *Vision Research, 21,* 1357–1375.

Williams, R. A., & Essock, E. A. (1986). Areas of spatial interaction for a hyperacuity stimulus. *Vision Research, 26,* 349–360.
Wilson, H. R. (1980). A transducer function for threshold and suprathreshold human vision. *Biological Cybernetics, 38,* 171–178.
Wilson, H. R. (1986). Responses of spatial mechanisms can explain hyperacuity. *Vision Research, 26,* 453–469.
Wilson, H. R., McFarlane, D. K., & Phillips, G. C. (1983). Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research, 23,* 873–882.
Wright, W. D., & Pitt, F. H. G. (1934). Hue-discrimination in normal colour vision. *Proceedings of the Physical Society* (London), *46,* 459–473.
Wülfing, E. A. (1892). Über den kleinsten Gesichtswinkel [The smallest detectable visual angle]. *Zeitschrift für Biologie, 29,* 199–202.
Wyszecki, G., & Stiles, W. S. (1982). *Color science* (2nd ed.). New York: Wiley.
Yamada, E. (1969). Some structural features of the fovea centralis in the human retina. *Archives D'Ophthalmologie, 82,* 151–159.
Yellott, J. I. (1984). Image sampling properties of photoreceptors: A reply to Miller and Bernard. *Vision Research, 24,* 281–283.

Appendix A

Linear-Systems Analysis

Consider stimuli generated on a monochrome display screen located at a fixed distance in front of the eye. Ignoring the temporal dimension, one can describe an arbitrary stimulus as a two-dimensional intensity distribution that is a function of the horizontal and vertical directions expressed in visual angles. The optical system forms a two-dimensional retinal-intensity distribution that can also be described as a function of these two variables. Thus, the eye transforms one two-dimensional intensity distribution into another.

Under restricted conditions, the human eye, like many other optical systems, can be accurately modeled as a two-dimensional, linear shift-invariant (LSI) system. A system is said to be linear if (a) its response to the sum of two arbitrary inputs equals the sum of its responses to the individual inputs, and (b) multiplying the input by an arbitrary factor multiplies the output by the same factor. A linear system is shift-invariant if shifting an arbitrary input function to a different position causes a corresponding shift in the output function with no change in its shape. Most optical systems are shift-invariant only over a limited region of the visual field. Thus, it is often necessary to divide the visual field into a number of regions so that over each region the optical system can be treated as an LSI system.

An LSI system is relatively easy to analyze and model. This is because its response to any input can be calculated from knowledge of its response to a single impulse (the impulse-response function). In the present case, the appropriate impulse would be a unit-energy point source. The impulse-response function to a unit-energy point source is called the point-spread function. To see how the point-spread function can be used to predict the output for an arbitrary input, note that an input image can be regarded as the sum of a large number of individual point sources that vary in position and energy. The point-spread function, together with the scaling and shift-invariance properties, allows one to determine the output-intensity distribution for each individual point source. The additivity property implies that the complete output is given by the sum of the outputs for each point source. Formally, these two steps are achieved by computing the convolution of the point-spread function with the input-intensity distribution. Thus, the LSI system can be completely characterized by its point-spread function.

The methods of Fourier analysis and synthesis often prove useful in measuring point-spread functions and in carrying out convolutions. The fundamental theorem in Fourier theory is that any function can be decomposed into a sum of sinusoidal functions of differing frequencies, amplitudes, and phases. This decomposition is called the Fourier transform. The result of a Fourier transform is a pair of functions: amplitude as a function of frequency (the amplitude spectrum) and phase as a function of frequency (the phase spectrum). In the case of two-dimensional images, the sinusoidal functions are sine-wave gratings; that is, they are patterns in which the intensity modulates sinusoidally over space. As shown in Figure A1, these two-dimensional sine waves have both a vertical and a horizontal spatial frequency. (A slice of any orientation through a two-dimensional grating has a sinusoidal profile.) Thus, both the amplitude and phase spectra are functions of vertical and horizontal spatial frequency. Note that (a) a vertically oriented grating has a spatial frequency of zero in the vertical direction, (b) a horizontally oriented grating has a frequency of zero in the horizontal direction, and (c) all other gratings have nonzero frequencies in both directions.

It is sometimes possible to determine the point-spread function of an optical system by direct measurement, but it is often easier to first measure the Fourier transform of the point-spread function, otherwise known as the *transfer function* (TF), which is a complex-valued function. The real-valued amplitude and phase spectra contained in the transfer function are called the *modulation-transfer function* (MTF) and the *phase-transfer function* (PTF), respectively. The point-spread function is obtained from the transfer function by applying an inverse Fourier transform.

The usual methods of measuring the transfer function are based on the following facts of Fourier theory. (a) The amplitude spectrum of the output of an LSI system is the product of the amplitude spectra of the input and the impulse-response function (in the present case, the point-spread function):

$$A_{out} = A_{in} \times MTF. \tag{1a}$$

(b) The phase spectrum of the output is the sum of the phase spectra of the input and the impulse-response function:

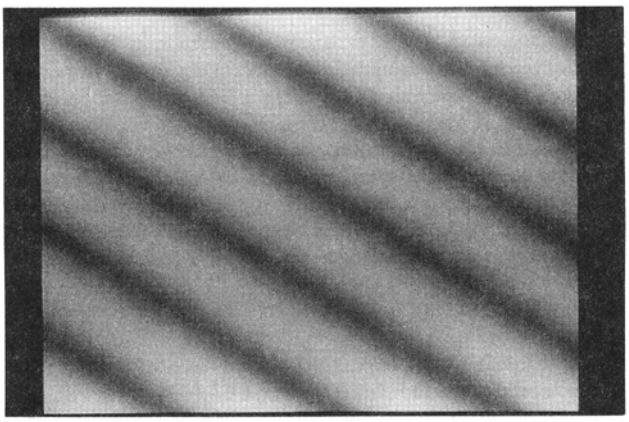


Figure A1. Example of a two-dimensional sinewave with spatial frequencies that are equal in the vertical and horizontal directions.

$$P_{out} = P_{in} + PTF. \quad (2a)$$

(c) A sine-wave input to an LSI system produces a sine-wave output that can differ from the input only in amplitude and phase. These properties imply that one can measure the transfer function of an optical system by presenting sine-wave gratings of all different spatial frequencies. If the amplitude and phase of all the gratings are recorded at both the input and the output, then the modulation-transfer function can be obtained by dividing output amplitude by input amplitude:

$$MTF = A_{out}/A_{in}.$$

The phase-transfer function can be obtained by subtracting input phase from output phase:

$$PTF = P_{out} - P_{in}.$$

Properties a and b are also sometimes useful for computing the convolution of the point-spread function with an input image (i.e., computing the output of a linear system). If one obtains the Fourier transform of the input, then the Fourier transform of the output can be easily obtained with Equations 1a and 2a. Finally, an inverse Fourier transformation yields the output image. This circuitous route through the Fourier domain is sometimes faster on a digital computer than directly calculating the convolutions. Such techniques were used in computing the performance of the present ideal-observer models (see Appendix B).

Appendix B

Computation Of Ideal-Observer Performance

This appendix provides some details on how the performance of the ideal observer was computed. Figure B1 illustrates the sequence of calculations.

Recall that in order to compute the retinal-intensity distribution, the input image must be convolved with the optical point-spread function and the receptor-aperture function (see Appendix A and the discussion of Equation 6 in the text). As discussed in Appendix A, convolutions can often be computed most quickly on a digital computer using Fourier transform techniques. Thus, the first step was to obtain the Fourier transform of the two stimuli in the discrimination task.

The most obvious way to obtain the Fourier transform is to apply a Fast Fourier Transform (FFT) routine. However, this has an undesirable consequence: Because the FFT is applied to discrete samples, it often introduces sampling artifacts (aliasing). These sampling artifacts are particularly large and troublesome for stimuli with sharp edges (discontinuities), a very common type of stimulus in psychophysical experiments. The common method of controlling for such artifacts in image processing is to remove the sharp edges by directly convolving the image with a small blurring function before applying the FFT (Rosenfeld & Kak, 1982). However, unless the sampling rate is very high (which is expensive in computation time), the size of the blurring function must be so large that the computations of the model would become inaccurate.

Therefore, a less general, but faster and more accurate, method was used. In particular, the program was constrained to deal only with stimuli that could be constructed by adding elementary parts with known Fourier transforms. These parts included rectangles, ellipses, two-dimensional Gaussian functions, and two-dimensional half-cosine functions of any size, aspect ratio, orientation, location, or color (mixture of primaries). In addition, any of these parts could be multiplied by a sine-wave function of any amplitude, phase, and frequency. Because the exact Fourier transform of each of these elementary parts is known (Gaskill, 1978), the exact Fourier transform of the stimuli could be obtained by summing the Fourier transforms of the parts composing the stimuli.

The second step in the computation was to multiply the Fourier transforms of the stimuli, in each primary, by the Fourier transform of the optical point-spread function (the OTF) associated with that primary. The point-spread function was normalized to have a volume of 1.0, so that it had no effect on the number of quanta reaching the retina. The OTFs were obtained by fitting the Campbell and Gubisch (1966) line-spread data with a sum of two Gaussian functions and (if needed) an exponential function, and taking Fourier transforms and converting them to two-dimensional OTFs assuming symmetric optics. Ocular transmittance was handled in the primary-to-cone transformation matrix described later. As discussed in Appendix A, multiplication in the Fourier domain corresponds to convolution in the image domain.

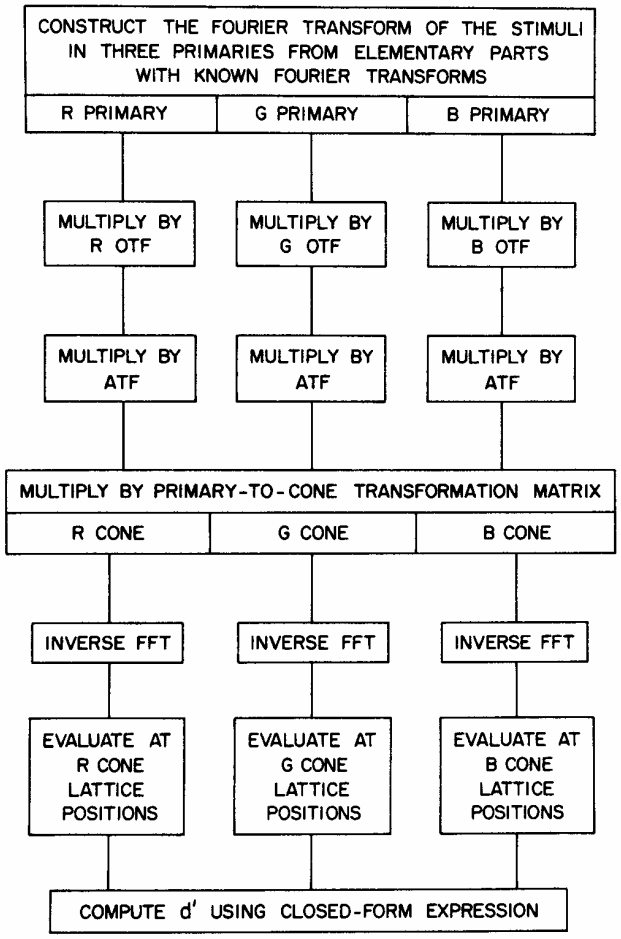


Figure B1. Flow diagram of ideal-observer calculations. (Calculations proceed from top to bottom.)

The third step was to multiply the result just discussed by the Fourier transform of the receptor-aperture function (ATF). The aperture function was also normalized to a volume of 1.0. The receptor-collection area was handled by the lattice-sampling stage described later.

The next step was to multiply the output for the three primaries by a 3×3 matrix that maps unit intensities of the primaries into the fractions of quantum absorptions in the cones. This matrix included both the effects of ocular transmittance and the absorptance spectra of the photopigments. Because Fourier transformation and matrix multiplication are both linear operations, it is legitimate to apply the primary-to-cone transformation matrix in the Fourier domain. Applying it at that stage saved storage by allowing all the inverse FFTs that followed to use the same complex data array.

The fifth step was to apply an inverse FFT in order to obtain the final images (expressed in units of cone absorption). At this point, the calculations went from a precise continuous representation to a discrete representation. However, the sampling problems associated with the inverse FFT were minor here because the optical point-spread function and aperture function effectively blurred the edges. In particular, the OTF and the ATF always limited the final image to contain frequencies below 60 c/deg. The Whittaker-Shannon sampling theorem implies that in this case, aliasing is avoided by keeping the sampling rate above 120 samples/deg. The effects of sampling were checked by varying the sampling rate within range of 120 to 360 samples/deg (1 to 3 times the Nyquist rate). For most stimuli, a sampling rate anywhere within this range was found to produce nearly identical results. It was necessary to stay near to top of this range only for stimuli composed almost entirely of very high spatial frequencies.

The sixth step was to evaluate the final image for each cone class at the lattice positions of the individual cones. The blur produced by the cone aperture was already incorporated into the calculations with the application of the ATF. Therefore, it was only necessary here to evaluate the image at the position of the center of each receptor and multiply by the area of the receptor.

The inverse FFT only produces the intensities of the images at the discrete locations of a square lattice. To obtain the image intensity at arbitrary receptor positions, it was necessary to interpolate the image intensities between the square lattice points. This was done by noting that each set of three neighboring image lattice points forms an isosceles triangle. Furthermore, the image intensity values above these points define a plane (in three space). Thus, the intensity at a given receptor position was obtained (to close approximation) by (a) finding the isosceles triangle in which the receptor center was located, (b) computing the plane through the associated image intensities, and (c) evaluating this plane at the position of the receptor's center. (Sinc function interpolation would have been slightly more accurate, but was not worth the increase in computation time.)

The final step was to compute d' using the closed-form expression given by Equation 9 in the text. For the position uncertainty cases, d' was computed by Monte Carlo simulation.

Thresholds were obtained by repeatedly executing the preceding calculations (while varying the appropriate stimulus parameter) until a d' of 1.36 (+/− 0.01) was found. Recall that a d' of 1.36 corresponds to 75% correct in a two-interval forced-choice task.

Received September 4, 1987
Revision received June 3, 1988
Accepted July 7, 1988 ■