Physical limits of acuity and hyperacuity

Wilson S. Geisler

Department of Psychology, University of Texas, Austin, Texas 78712

Received October 26, 1983; accepted March 13, 1984

An ideal detector is derived for the discrimination of arbitrary stimuli in the two-alternative forced-choice paradigm. The ideal detector's performance is assumed to be limited only by quantal fluctuations, the optics of the eye, and the size and spacing of the receptors in the retinal mosaic. Detailed predictions are presented for two-point acuity and hyperacuity tasks. The ideal detector's two-point resolution, over a wide range of luminances, is approximately 10 times worse than its two-point vernier acuity or separation discrimination. Furthermore, two-point resolution is shown to vary in proportion to the $-1/4$ power of spot intensity, but vernier acuity and separation discrimination vary in proportion to the $-1/2$ power of spot intensity. It is shown that this ideal detector can be implemented by the use of appropriately shaped receptive fields. The derivation provides a simple way to determine the shapes of these optimal receptive fields for arbitrary stimuli. The sensitivities of real (human) and ideal detectors are compared.

Under many circumstances, the eye is capable of resolving changes in position that are nearly an order of magnitude smaller than the 0.5-min diameter of a foveal photoreceptor. For example, observers can reliably detect an 8–10-sec instantaneous displacement of a short line or bar,[1] a 2–4-sec misalignment in a vernier-acuity task,[2–4] a 6-sec difference in the separation of two lines,[1] and a 2–4-sec difference in the separation of lines in a stereoacuity task.[5,6] Westheimer coined the term hyperacuity to describe the high levels of performance observed in the above tasks relative to those obtained in more conventional spatial-acuity tasks, such as grating resolution, which yields thresholds of the order of 30–60 sec (the approximate diameter of the inner segment of a foveal cone).

The exquisite sensitivities obtained in hyperacuity tasks are amazingly resistant to changes in the spatial configuration of the stimuli.[1,7] Thus hyperacuity is a robust phenomenon of some generality. Furthermore, many of the hyperacuity tasks undoubtedly are measuring the resolution limits of fundamental visual processes. For example, highly sensitive displacement (small-movement) detection and separation discrimination are probably of prime importance for the mechanisms that extract, from the optical flow, the orientation and motion of the observer and the orientation and relative distances of surfaces in the environment.[8–10] Similarly, excellent stereoacuity must be crucial for the extraction of information about surface orientation and depth from stereopsis.[11–13] Since stereo and optical-flow information decline rapidly with distance from the observer (the inverse-square law), it follows that every additional second of resolution that can be obtained will increase significantly the volume of space around the observer from which reliable distance and surface-orientation information can be extracted. It would not be surprising to find that the evolutionary process had exerted some of its greatest pressure to produce the mechanisms underlying hyperacuity tasks.

The extraordinary sensitivity obtained in hyperacuity judgments leads one to ask, What are the physical limits of sensitivity in the hyperacuity tasks? The primary purpose of this paper is to provide an answer to this question. In particular, we will consider the performance of an ideal detector (or machine) that has complete knowledge of the number of quanta caught in each photoreceptor in the retinal lattice. In other words, we assume that the machine has access to all the information that could possibly be transmitted from the photoreceptors.

Ideal detectors of this sort provide a useful yardstick against which the performance of real observers is measured. If the ideal detector's performance significantly outdistances that of the real observer, then the neural processes must be discarding considerable information. In this case, the obvious research strategy is to determine first what information the real observers are using and then consider how this information is extracted and used. On the other hand, if the real observer and the ideal detector are nearly equal in performance, then neural processes must be using nearly all the available information, just as the ideal detector does. This, of course, greatly constrains potential models of the real observer.

The present ideal detector is useful in another important way. Since it uses all the information available in the receptors, its performance is a measure of the information content of the stimuli being considered. In other words, the ideal detector's performance provides us with a better understanding of the stimuli used in spatial-discrimination tasks.

The performance of the present ideal detector will be derived for the two-alternative forced-choice paradigm. In this paradigm, the detector must decide, on each trial, which of two possible stimuli (α or β) is presented. To begin with, we will assume that the detector has complete knowledge of the possible stimuli. In other words, we assume that the detector knows the exact intensity distributions, durations, and positions of the stimuli and that it knows that the stream of photons absorbed in each receptor is described by a Poisson process. It is also assumed that the Poisson randomness of light is the only source of noise present.

The distinction between what we refer to here as ideal-detector models and other types of models deserves some comment, since the distinction has not been made clear in

0740-3232/84/070775-08$02.00

either the audition or the vision literature. Generally speaking, the purpose of a typical model is to account for the actual behavior of the sensory system under consideration; thus we might call it a real-observer model. On the other hand, the purpose of an ideal-detector model is to determine the behavior of a hypothetical sensory system that has only those properties that we know exist in the real system and that we understand in quantitative detail. This means that real-observer models should contain the properties of the most complete ideal-detector model available plus additional properties that are hypothesized to account for the differences between real and ideal performance. Thus ideal-detector and real-observer models work hand in hand; ideal-detector models show what aspects of performance can be explained by known factors, and real-observer models try to explain other aspects of performance by including additional hypothesized factors. As more properties of a system are understood in quantitative detail, ideal-detector and real-observer models should converge to the same model.

It should be clear from the above discussion that there may be many ideal-detector models for the same sensory tasks. The relative merits of one ideal detector over another must be judged by comparing relative accuracy and completeness. The best ideal detector is the one that accurately includes all the factors that are well understood at that time.

DERIVATION OF THE IDEAL DETECTOR

Point-Spread Function

The line-spread functions reported by Campbell and Gubisch[14] were used to calculate the distribution of light on the retina. For purposes of calculation the curves in their Fig. 10 were approximated by the sum of two Gaussian functions:

$$h(x) = \frac{a_1}{2s_1}\exp[-0.5(x/s_1)^2] + \frac{a_2}{2s_2}\exp[-0.5(x/s_2)^2].$$

For example, their data and the predicted curve for a 2-mm pupil size are shown in Fig. 1. The point-spread function was then determined by assuming that the optics of the eye are isotropic.

Receptor Lattice

It is assumed that the receptors are positioned at the nodes of a hexagonal array,[15] such that the internodal distance equals the diameter of the receptors. This provides the tightest possible packing of photoreceptors (see Fig. 2). The diameter of the receptors is assumed to be 0.6 min of arc (Table 4 of Ref. 15). The total size of the receptor lattice, which varied from 400 to 10 000 receptors, depending on the stimuli, was always picked, so any increases in lattice size produced no further improvements in performance.

Poisson Distribution

In a fixed time period, the number of photons (Z_i) absorbed in the ith photoreceptor of the lattice is described by the Poisson density:

$$p(z) = \frac{\alpha_i{}^z \exp(-\alpha_i)}{z!}, \qquad z = 0, 1, 2, \ldots, \qquad (1)$$

where α_i is the mean number of quanta absorbed.

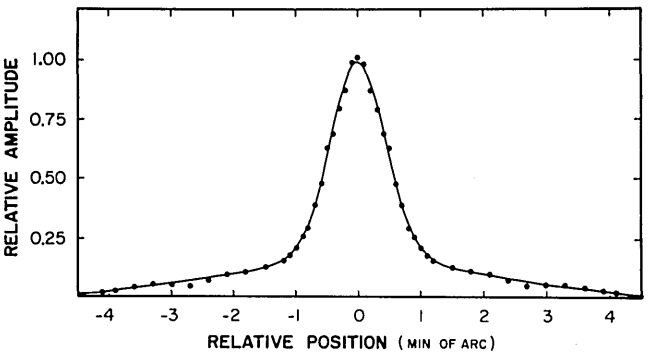

Fig. 1. Line-spread function of the eye for a 2-mm pupil. The symbols are data from Campbell and Gubisch.[14] The smooth curve is the weighted sum of two Gaussian functions; $s_1 = 0.443$ min, $s_2 = 2.035$ min, $a_1 = 0.684$, $a_2 = 0.587$. The actual point-spread function was obtained from the smooth curve by assuming that the eye is isotropic and by normalizing the total volume under the function to 1.0.

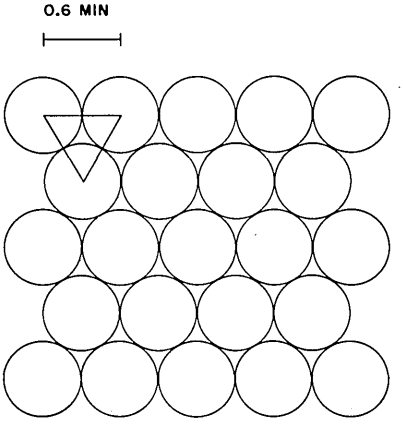

Fig. 2. Hexagonal receptor lattice used to generate predictions. A hexagonal array is characterized by the fact that the center-to-center distances for any triplet of neighboring elements forms an equilateral triangle. The actual array sizes were as large as 10 000 elements.

Other Optical Factors

The value of α_i is dependent in a straightforward manner on a number of optical factors in addition to the point-spread function. If we use the conversion values given by Wysecki and Stiles,[16] the mean number of photons effectively absorbed by the ith receptor, whose position we denote by the coordinates (x_i, y_i), is given by

$$\alpha_i = ADSTE_{555}347.8l(x_i, y_i) * h(x_i, y_i), \qquad (2)$$

where A is the cross-sectional area of the receptor in square minutes, D is the duration of the stimulus in seconds, S is the pupil area in square millimeters, T is the transmittance of the ocular media, E_{555} is the quantum efficiency of the photoreceptors at 555 nm, $l(x, y)$ is the luminance distribution of the stimulus in candelas per square meter, and $h(x, y)$ is the point-spread function. (Note that $*$ represents the operation of convolution.) In all the predictions presented here, the following values for these parameters were used: $A = 0.28$ min^2, $D = 0.2$ sec, $S = 3.1416$ mm^2 (2-mm pupil), $T = 0.68$, and $E_{555} = 0.5$. Since all the predicted curves are plotted in log–log coordinates, any changes in these parameters would leave the shapes and relative positions of the curves unchanged.

Equation (2) does not take into account the Stiles–Crawford effect, because the effect is negligible for the 2-mm pupil size assumed here. However, for larger pupil sizes, it should be included, especially when deriving predictions for stimuli falling on the peripheral retina.

Likelihood Density and d'

It can be shown[17] that an optimal detector performing in a two-alternative forced-choice task bases its decision on the likelihood ratio or on a monotonic transformation of it. When the density functions of the likelihood ratio (or its monotonic transform) are approximately equal-variance normal, then the detector's sensitivity is given by the quantity d', the difference in the means of the two distributions divided by the common standard deviation. In Appendix A, we show (for the present case) that

$$d' = \frac{\sum_{i=1}^{n} (\beta_i - \alpha_i)\ln(\beta_i/\alpha_i)}{\left[0.5 \sum_{i=1}^{n} (\beta_i + \alpha_i)\ln^2(\beta_i/\alpha_i)\right]^{1/2}}, \quad (3)$$

where α_i and β_i are the mean numbers of effectively absorbed quanta in receptor i to the stimuli α and β, respectively. In calculating the threshold predictions, we assumed that the threshold was reached when $d' = 1.36$. This value of d' will produce exactly 75% correct responses if the detector is unbiased and if the two alternatives are presented randomly, each with probability 0.5. Note that the above derivation is for the single-interval forced-choice paradigm (i.e., a yes–no task). The predicted value of d' for the two-interval forced-choice paradigm differs from Eq. (3) by a factor of $\sqrt{2}$.[17]

PREDICTIONS

There are many experimental conditions for which it might be interesting and useful to generate the predictions of the present ideal detector. Indeed, it is fairly easy to implement the present model on a computer so that predictions can be generated for almost any forced-choice discrimination experiment. However, in this paper, predictions are shown only for stimuli consisting of (at most) two small spots presented foveally against a dark background. It was decided first to examine this rather restricted domain for two reasons. First, even with two-spot stimuli, one can construct simple intensity-discrimination, resolution, separation-discrimination, and vernier-acuity tasks. It seems reasonable first to derive predictions for these simple tasks. Second, from an information-processing view, the most surprising hyperacuity and acuity thresholds reported in the literature have been for conditions using small spots. For example, one of Westheimer and McKee's[18] subjects had a vernier acuity for small spots of about 6 sec of arc. The dots were only 30 sec long and a "fraction of a minute" wide. They had a luminance of about 100 cd/m^2 and were presented for a duration of 0.2 sec. Westheimer and McKee also showed that vernier acuity is not improved significantly by using lines up to 8 min long. The present ideal detector's performance improves substantially with line length; thus real and ideal performances probably come the closest for the small-spot tasks.

Predictions were derived for the small-spot conditions shown in Fig. 3. In each case, the solid circles show the spatial arrangement for stimulus α and the open circles for stimulus β. The bar graphs show the intensity profiles of the stimuli in the direction of θ.

Intensity Discrimination

As a check on the computer implementation of the theory, it was decided to generate first predictions for the intensity-discrimination experiment shown in Fig. 3A. Stimulus α was a single spot and stimulus β a single spot of a different intensity falling at the same point. Since stimuli α and β are identical except for intensity, all α_i and β_i in Eq. (3) have the same ratio, and Eq. (3) reduces to the familiar equation

$$d' = \Delta N/\sqrt{N}, \quad (4)$$

where ΔN is the average difference in the number of effectively absorbed quanta from stimuli α and β and N is the mean number of quanta per stimulus. It follows immediately that ΔN at threshold should equal $1.36\sqrt{N}$.[19] This is the square-root law.[20,21] If the program correctly calculates all the quantities in Eq. (3), then it should yield the above relation. The result, which is correct, is shown in Fig. 4. Note that the result in Fig. 4 also holds for arbitrary stimuli that differ only in intensity. In other words, in the simple inten-

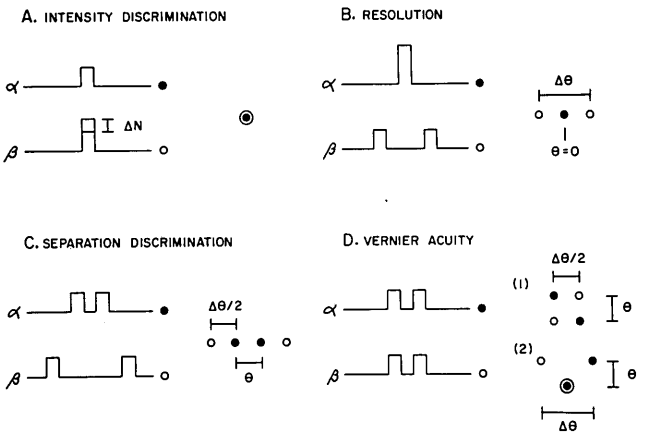

Fig. 3. Acuity and hyperacuity tasks using point sources. The filled circles show the spatial arrangement for stimulus α, and the open circles stimulus β in the forced-choice paradigm. The bar graphs show the intensity profiles of the stimuli in the direction of the base separation θ.

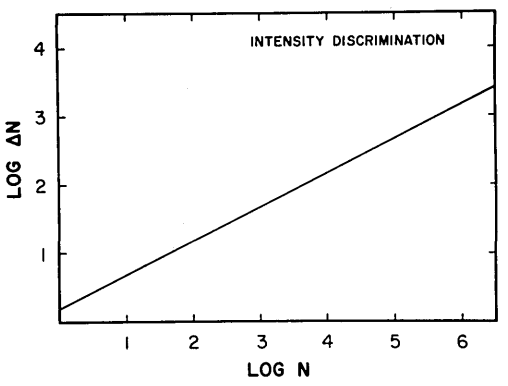

Fig. 4. Mean difference in effectively absorbed quanta at threshold (75% correct) as a function of the average total effectively absorbed quanta per stimulus.

sity-discrimination task, the ideal detector does not care how the quanta are distributed around the receptor lattice.

Resolution

As mentioned above, resolution tasks employing humans typically yield thresholds around an order of magnitude larger than those of the hyperacuity tasks. In the hope that the ideal detector might shed some light on the difference between the two classes of experiment, predictions were obtained for the two-point resolution task shown in Fig. 3B. For stimulus α, two point sources are exactly superimposed (i.e., the separation angle θ is zero); for stimulus α, they are separated by an angle $\Delta\theta$. Figure 5 shows the value of ΔN needed for 75% correct performance as a function of the mean number of effectively absorbed quanta per point source. If the test spots are 30 sec × 30 sec in size (approximately the size used by Westheimer and McKee[18]), then their luminance (in candelas per square meter) is given by the upper scale. The lowermost curve in the figure is the prediction for a perfect point source. The other curves show the predictions for point sources that have been blurred (low-pass filtered) by Gaussian weighting functions with various standard deviations. The corresponding low-pass cutoff frequencies associated with these standard deviations are given in the figure caption. The cutoff frequency is defined as the spatial frequency that would be attenuated by 90% (i.e., the 10-dB down point).

The resolution threshold varies with intensity, but, surprisingly, only to the $-1/4$ power of intensity. From 10 to 100 cd/m^2, resolution is only slightly better than the diameter of a photoreceptor, but at higher luminances it becomes considerably better. Blurring the point sources increases the resolution threshold but does not interact with the effect of intensity.

For the resolution task and the others described below, the stimuli were moved around the receptor lattice to determine if positioning had any effect on the performance of the ideal detector. It did not for any of the present conditions. Apparently, the radius of the point-spread function is large enough, relative to the receptor diameter and the spaces between receptors, to wash out any positional effects. This may not be true in the peripheral visual field, in which receptor density drops faster than the quality of the optics.[22]

Separation Discrimination

Next, consider the performance of the ideal detector in the separation-discrimination task shown in Fig. 3C. On each trial, the detector must decide whether the two point sources are separated by angle θ or by $\Delta\theta + \theta$. Westheimer[1] has shown that for appropriate values of θ, human observers perform as well in this task as in the vernier tasks (Fig. 3D). Interestingly, two-point resolution and separation discrimination can be thought of as lying along the continuum created by changing the base separation θ. Figure 6 shows the ideal detector's performance for point sources of around 100 cd/m^2 as the base separation θ is increased. Increasing θ produces about a sixfold increase in sensitivity. Once the separation exceeds about 5 min of arc, sensitivity levels out. The magnitude of the effect of separation depends on the luminance of the stimuli. Figure 7 shows how the separation threshold changes with luminance when the base separation θ exceeds 5 min. Unlike resolution ($\theta = 0.0$), the separation threshold varies in proportion to the $-1/2$ power of intensity. Thus the difference in the ideal detector's performance in the resolution

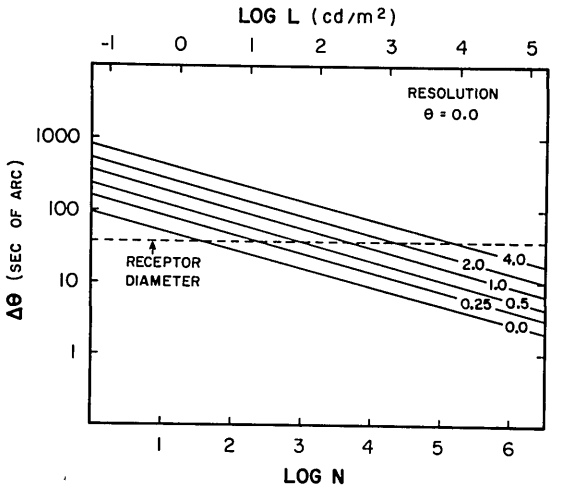

Fig. 5. Two-point resolution threshold as a function of the average total number of effectively absorbed quanta per point source. Different curves show the effect of blurring the point sources. The numbers on the curves indicate the standard deviations of the Gaussian intensity distribution. The cutoff frequencies associated with these curves are (from top to bottom) 5, 10, 20, 41, 82, and ∞ cycles/deg. (All curves were obtained with the point-spread function derived from the line-spread function in Fig. 2.)

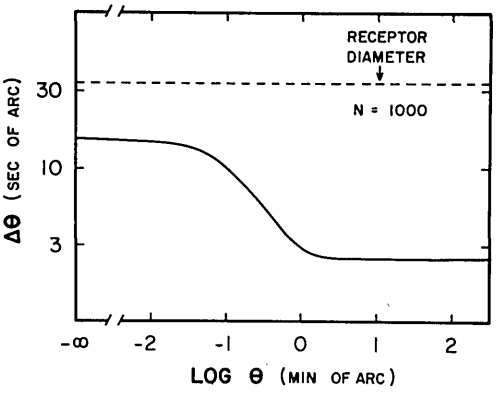

Fig. 6. Separation-discrimination threshold for point sources as a function of base separation. The mean total effective quanta per point source equals 1000. When the base separation is 0.0 ($-\infty$ on the log scale) the task reduces to two-point resolution.

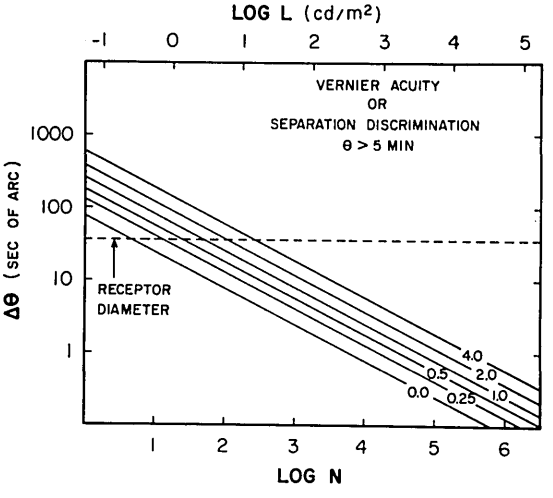

Fig. 7. Separation discrimination or vernier threshold for base separations greater than 5 min of arc as a function of the average total number of effectively absorbed quanta per point source. The different curves show the effect of blurring the point sources (see caption for Fig. 5).

and separation-discrimination tasks increases with luminance. Figure 7 also shows the effect of low-pass filtering the point sources. Blurring the point sources has the same effect on separation discrimination as it does on resolution, namely, threshold increases but there is no change in the slope.

Vernier Acuity

To the ideal detector, the two-point vernier acuity task is almost equivalent to the separation-discrimination task. In fact, if the separation angle θ exceeds 5 min, the predictions of the model for the vernier task in Fig. 3D(1) are exactly the same as those in Fig. 7. The reason they are the same is that once the separation exceeds 5 min or so, the two point sources making up a stimulus will not interact with each other. Thus one of the solid–open-circle pairs in Fig. 3C can be moved to produce the arrangement in Fig. 3D(1) without affecting performance. If θ is less than 5 min, then performance is different for the task in Fig. 3D(1). For example, it is obvious that the threshold approaches infinity instead of a constant as θ approaches 0.0.

The ideal detector performs almost the same in the vernier task shown in Fig. 3D(2) as in the task shown in Fig. 3D(1). Again, if the separation angle is greater than about 5 min, the predicted curves are identical to those in Fig. 7, but they are all shifted downward by 0.3 log unit. This is because only one of the point sources moves from trial to trial. The two point sources that do not move may, therefore, be ignored. This, in effect, reduces the noise without reducing the signal. Unlike the other vernier task, as θ is decreased the threshold increases and then levels off to a constant value.

DISCUSSION

Comparison of Real and Ideal Detectors

Figures 4–7 incorporate estimates of the physical limits of small-spot acuity imposed by quantum fluctuations, the optics of the eye, and receptor spacing. There have not been nearly enough carefully controlled small-spot experiments reported in the literature to permit detailed comparison of real and ideal performances. However, the available data indicate that, under some circumstances, humans approach ideal performance. Westheimer and McKee[18] obtained vernier acuities of around 6 sec in a task like that in Fig. 3D(2). They did not report the pupil size of their subjects or the exact width of the point sources; we only know that it was a "fraction of a minute" wide. Thus the mean number of quanta absorbed per test spot (N) cannot be accurately estimated. If the pupil diameter were 2 mm, then the appropriate ideal detector would have an acuity of around 2 sec. If the pupil size were larger (as is likely), the value of N would increase, resulting in improved acuity for the ideal observer. However, this improvement would be partially offset by the increased radius of the point-spread function[14] and by the Stiles–Crawford effect. Under the conditions of Westheimer and McKee's experiment, the human observer (subject SM) seems to be only 3–6 times less sensitive than the ideal observer.

The ideal detector shows the same, rather large, difference between resolution tasks and hyperacuity tasks that is found in human observers.[7] Over the midphotopic range, two-point resolution is 6–10 times worse than separation discrimination or vernier acuity (Figs. 5–7). This is about the same order of magnitude found in human observers. Thus the difference in performance found in human observers may be accounted for primarily by differences in the stimulus at the receptors.

Ideal and real observers are similar in another way. Both achieve the best performance in hyperacuity tasks when the separation angle θ is 3–5 min.[7] However, the real observer's performance, unlike that of the ideal detector, deteriorates at larger separations. This decline may be due to an increase in position uncertainty or to the absence of detectors able to completely integrate information over larger distances.

A large number of studies have examined the effects of luminance on visual-acuity tasks[23–25] (see Ref. 25 for a review of the older literature), but relatively few have dealt with small-spot stimuli. Campbell and Legge[26] found that increasing luminance produced only slight decreases in displacement thresholds for small spots. However, under their conditions, observers did not perform so well as in other studies and, therefore, did not come close to the performance of the ideal detector. Westheimer and McKee[4] measured vernier acuity for thin lines (15 sec × 6.4 min) as a function of luminance. They found that the threshold decreased with luminance until the midphotopic range was reached; then it leveled out. Qualitatively similar results are found in resolution tasks.[24]

Constant acuity as a function of luminance is consistent with Weber's law and hence is not too surprising for the mid- to high-photoic range. Weber's law is, of course, not the behavior of an ideal detector limited by quantum fluctuations. Perhaps there is a range of luminances below the Weber region for which acuity (in hyperacuity and resolution tasks) is limited by quantal fluctuations. Carefully controlled small-spot experiments will be needed to determine if this is the case. One interesting prediction is that the slopes of the two-point resolution and separation-discrimination curves should differ by a factor of 2.

The effects of low-pass filtering on two-point resolution and separation discrimination are shown in Figs. 5 and 7. The effect of filtering is similar for the two tasks and not so strong as one might expect if high-spatial-frequency information were critical in these tasks. For example, changing the cutoff frequency from 20 to 10 cycles/deg reduces resolution by a factor of only 1.6 and separation discrimination by a factor of only 1.4. In two-line stereoacuity, Westheimer and McKee[6] observed a factor of 1.3 for these same two low-pass filters.

On the Differences between Real and Ideal Performance

Although there are some similarities in the performance of real observers and ideal detectors, it is clear from the above discussion that there are other factors limiting the real observer's performance. Below is a brief discussion of some factors that are almost certainly involved, although at this time their relative contributions to the real observer's performance are uncertain.

Accommodation

Variability or systematic errors in accommodation would result in reduced performance equivalent to low-pass filtering the stimuli (see Figs. 5 and 7). Since this would not change the shapes of the predicted curves, it could not account for many of the differences between real and ideal performance.

Internal Noise

No real processing system is free of intrinsic noise. Two possible types of noise, additive[27] and multiplicative,[28] have been suggested to help account for absolute- and increment-threshold data. The effect of either type of intrinsic noise on visual thresholds critically depends on where the noise arises along the visual pathway. For example, additive noise that arises early in the visual system, such as spontaneous thermal breakdown of photopigment,[27] has its major effect at low light levels. This dark-light noise yields a threshold-versus-intensity function (tvi curve) like the one in Fig. 4, except that it asymptotically approaches a constant threshold at low light (log N) levels. Dark-light noise is easy to incorporate into the present theory and may account for some of the discrepancies between real and ideal performance.[29] If we let x_0 be the mean number of quantumlike events per receptor resulting from internal noise during stimulus presentation, then it follows directly from the derivation in Appendix A that

$$d' = \frac{\sum_{i=1}^{n} (\beta_i - \alpha_i)\ln[(\beta_i + x_0)/(\alpha_i + x_0)]}{\left\{0.5 \sum_{i=1}^{n} (\beta_i + \alpha_i + 2x_0)\ln^2[(\beta_i + x_0)/(\alpha_i + x_0)]\right\}^{1/2}}.$$

Additive noise that arises further along the visual pathway could easily affect sensitivity at low light levels much like dark light, but, because of nonlinearities in the visual system, it may have a different effect at high light levels (see below).

Nonlinearities

If all significant sources of intrinsic noise arise before some monotonic nonlinearity in the visual system, then, in many cases, the nonlinearity will have no effect on the threshold. This is true for the same reason that, in signal-detection theory, every monotonic transformation of the decision variable (the likelihood ratio) leads to the same optimum performance level.[17] On the other hand, if significant noise arises after the nonlinearity, then the nonlinearity can strongly affect sensitivity.

For example, suppose that the visual system contains a compressive nonlinearity[30,31] and a small level of postcompression noise. The response compression will reduce the signal and precompression noise (e.g., quantal fluctuations) equally but leave the postcompression noise unaffected. Thus the postcompression noise becomes more dominant as the system is driven into response compression. In other words, at high light levels, a little postcompression noise may determine sensitivity, even though at low light levels the postcompression noise may be overshadowed by quantal fluctuations. This effect (which would occur with almost any sort of gain-control mechanism) may contribute to the smooth transition from quantal-fluctuation-limited sensitivity at low light levels to Weber's law at high light levels.

Receptive-Field Shape

The ideal detector is able to integrate information across the receptors in an optimal fashion. Examination of Eq. (5) in Appendix A, shows that one can build optimal receptive fields by using simple linear summation and inhibition. The coefficients of the receptive-field weighting function are just the values $\ln(\beta_i/\alpha_i)$ {or $\ln[(\beta_i + x_0)/(\alpha_i + x_0)]$ if there is dark light present in the photoreceptors}. The dashed curve in Fig. 8A shows the cross section of the optimal receptive field for the vernier-acuity task in Fig. 3D(2) when the distance between the point sources exceeds 5 min of arc. In other words, this is the ideal receptive field for deciding if a single point source appears to the left or to the right of some reference point (position 0). The dashed curve shows the optimal receptive field if there is absolutely no intrinsic noise in the visual system. However, a little dark–light noise (1 quantumlike event per receptor every 20 sec) leads to the optimal receptive field given by the solid curve in Fig. 8A. The entire receptive-field profile is given in Fig. 8B.

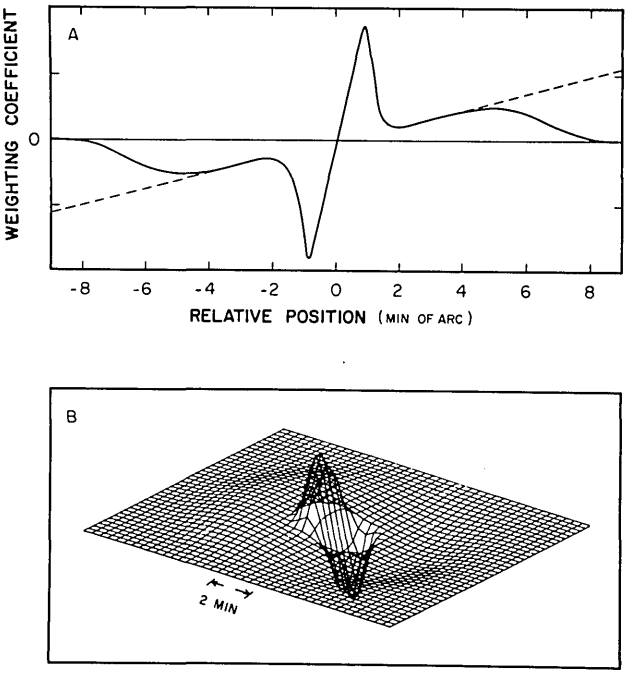

Fig. 8. Ideal receptive field for localizing a single point source on a dark background. A, cross section of the receptive field assuming no intrinsic noise in the visual system (dashed curve) and a dark–light noise level of 1 quantumlike event per receptor every 20 sec (solid curve). B, entire ideal receptive field for the dark–light noise level in A.

It may seem nonintuitive that multiple-lobed receptive fields are optimal for discriminating the location of a single point source. The multiple lobes are due to the shallow side skirts of the human point-spread function (see Fig. 1). It is easy to show, by examining the coefficient terms $\ln[(\beta_i + x_0)/(\alpha_i + x_0)]$, that if the point-spread function were described by a single Gaussian function instead of the sum of two Gaussian functions, then the extra side lobes would not be present in the optimal receptive field.

The optimal receptive fields for two-point resolution and separation discrimination also have multiple side lobes, but, instead of having an excitatory region on one side and an inhibitory region on the other, they consist of an excitatory or inhibitory center region flanked by regions of the opposite polarity.

The reason that human performance approaches ideal performance in the small-spot tasks may be because the neurons early in the visual pathway have receptive-field profiles that are nearly optimal for these tasks. However, it

seems certain that real receptive fields can only approximate the optimal receptive fields for certain stimuli. Thus the performance of real observers must deviate greatly from ideal under many circumstances. This may explain, for example, why human performance deteriorates in vernier and separation-discrimination tasks when the base separation θ exceeds 5 min of arc.

Positional Uncertainty

In most experimental situations, small eye movements must cause the stimuli to fall on different receptors from trial to trial. This and perhaps other factors, such as nonoptimal fixation cues, are likely to produce considerable uncertainty about stimulus position on the receptor lattice. We have tried to ascertain the effect of positional uncertainty on the present ideal detector. Although one can easily write an expression for the likelihood ratio, it is in an integral form from which we could not determine the appropriate probability densities. Elaborate Monte Carlo simulations may be necessary to understand the effects of positional uncertainty.

As shown above, the present ideal detector can be implemented in any given two-alternative forced-choice task by monitoring the output of a single appropriately shaped liner receptive field. However, if the receptive-field shapes in the real visual system are not optimal or if the neural units extracting the information are noisy, then it would be useful to integrate or look across the output of more than one unit. Furthermore, if there is substantial positional uncertainty, then it is absolutely essential to monitor a number of receptive fields scattered throughout the uncertainty region. Several authors[32–34] have presented some ideas about how this might be done. Their hypotheses rely on the precise localization of stimulus features. Although the visual system undoubtedly contains reasonably good localization mechanisms, it seems unnecessary (and unparsimonious) to suppose that feature localization has to be accurate to within a few seconds of arc. In the typical forced-choice hyperacuity tasks, the visual system may detect changes in the pattern of neural unit responses within the uncertainty area to the two different stimuli without determining the actual location of stimulus features. Localization mechanisms are undoubtedly involved in hyperacuity tasks in which stimuli are separated by relatively large distances over which receptive fields cannot integrate the necessary information; however, the performance of real observers is then out of the hyperacuity range.

CONCLUSION

The ideal detector derived here provides a baseline against which to gauge the performance of human observers. A rigorous comparison of real and ideal performance will require that resolution, separation-discrimination, and vernier-acuity experiments, such as those in Fig. 3, be carried out under the same conditions and with the same subjects.[35] Nonetheless, on the basis of the available data, it appears that real observers approach ideal performance under some conditions.

Although there are some similarities between the real and ideal performance, there are also many important differences. To understand these differences, it may prove useful to derive the performance of nonoptimal detectors obtained by incorporating into the theory some or all of the factors described above. The easiest to consider would be internal noise (in particular, dark light) and receptive-field shape. For example, it may be useful to compare the shapes of the optimal receptive fields with the receptive-field shapes implied by electrophysiological measurements. Conversely, it may also be useful to see how much the shapes of the ideal receptive fields can be distorted and still predict performance levels better than that of human observers.

The most difficult and probably the most important factor to incorporate into the present ideal-detector theory of spatial vision is stimulus uncertainty. That is, we need to develop ideal detectors for conditions in which the absolute positions of the stimuli are uncertain from trial to trial. (These detectors would parallel, in the auditory literature, ideal observers in which the signal is specified exactly except for phase.[17]) Human observers always operate under some degree of positional uncertainty; thus, to help understand human performance, it is crucial to understand how to extract information optimally under such circumstances.

APPENDIX A. DERIVATION OF d' FOR AN IDEAL DETECTOR

Let Z_i be the number of quanta absorbed on a single trial in receptor i. Then $\langle Z_1, \ldots, Z_n \rangle$ is a sample of independent Poisson-distributed random variables. On each trial, one of two stimuli, call them α and β, is presented at random. Let α_i and β_i be the mean numbers of quanta absorbed in receptor i for stimuli α and β.

The ideal detector decides which of the two stimuli was presented by computing the likelihood ratio (or some monotonic transformation of it). If the likelihood ratio falls on one side of some criterion value, the detector picks stimulus α; if the ratio falls on the other side, stimulus β is chosen. (Optimum placement of the criterion depends on what aspect of performance the detector is trying to optimize.)

Now, let $p(z_1, \ldots, z_n|\alpha)$ and $p(z_1, \ldots, z_n|\beta)$ be the joint probability densities of $\langle Z_1, \ldots, Z_n \rangle$ for stimuli α and β, respectively. Then the likelihood ratio L is given by

$$L = p(Z_1, \ldots, Z_n|\beta)/p(Z_1, \ldots, Z_n|\alpha)$$

$$= \prod_{i=1}^{n} p(Z_i|\beta)/\prod_{i=1}^{n} p(Z_i|\alpha),$$

where

$$p(Z_i|\alpha) = \frac{\alpha_i^{Z_i} \exp(-\alpha_i)}{Z_i!},$$

$$p(Z_i|\beta) = \frac{\beta_i^{Z_i} \exp(-\beta_i)}{Z_i!}.$$

As is often the case, it is easier to work with some monotonic transformation of the likelihood ratio. By letting Z be the natural logarithm of L minus the constant $C = \Sigma_{i=1}^{n} \alpha_i - \beta_i$, we have

$$Z = \sum_{i=1}^{n} Z_i \ln(\beta_i/\alpha_i). \tag{5}$$

To compute the detector's sensitivity, we need to know the density function for Z given α and β. From Eq. (5) we see that Z is the sum of scaled Poisson random variables. It is difficult

to determine the shape of this density in general; however, it is easy to obtain the means and variances:

$$E(Z|\alpha) = \sum_{i=1}^{n} \alpha_i \ln(\beta_i/\alpha_i),$$

$$E(Z|\beta) = \sum_{i=1}^{n} \beta_i \ln(\beta_i/\alpha_i),$$

$$\mathrm{Var}(Z|\alpha) = \sum_{i=1}^{n} \alpha_i \ln^2(\beta_i/\alpha_i),$$

$$\mathrm{Var}(Z|\beta) = \sum_{i=1}^{n} \beta_i \ln^2(\beta_i/\alpha_i).$$

Finally, by approximating Z with a normal density, we have

$$d' = \frac{\sum_{i=1}^{n} (\beta_i - \alpha_i)\ln(\beta_i/\alpha_i)}{\left[0.5 \sum_{i=1}^{n} (\alpha_i + \beta_i)\ln^2(\beta_i/\alpha_i)\right]^{1/2}}.$$

ACKNOWLEDGMENTS

This research was supported by the National Institutes of Health grant EY02688. I thank Martin S. Banks and Dennis McFadden for helpful comments.

REFERENCES

1. G. Westheimer, "The spatial sense of the eye," Invest. Ophthalmol. **18,** 893–912 (1979).
2. T. Y. Baker and G. B. Bryan, "Errors of observation," in *Proceedings of the Optical Convention* (Hodder and Stroughton, London, 1912).
3. R. N. Berry, "Quantitative relations among vernier, real depth, and stereoscopic depth acuities," J. Exp. Psychol. **38,** 708–721 (1948).
4. G. Westheimer and S. P. McKee, "Spatial configurations for visual hyperacuity," Vision Res. **17,** 941–947 (1977).
5. H. J. Howard, "A test for the judgment of distance," Am. J. Opthalmol **2,** 656–675 (1919).
6. G. Westheimer and S. P. McKee, "Stereoscopic acuity with defocused and spatially filtered retinal images," J. Opt. Soc. Am. **70,** 772–778 (1980).
7. G. Westheimer, "Visual hyperacuity," in *Progress In Sensory Physiology*, Vol. 1, H. Autrum, D. Ottoson, E. R. Perl, and R. F. Schmidt, eds. (Springer-Verlag, New York, 1981), pp. 1–30.
8. J. J. Gibson, *The Senses Considered As Perceptual Systems* (Houghton-Mifflin, Boston, 1966).
9. C. H. Longuet-Higgins and K. Prazdny, "The interpretation of a moving retinal image," Proc. R. Soc. London Ser. B **208,** 385–397 (1980).
10. W. F. Clocksin, "Perception of surface slant and edge labels from optical flow: a computational approach," Perception **9,** 252–267 (1980).
11. D. T. Marr and T. Poggio, "A computational theory of human stereo vision," Proc. R. Soc. London Ser. B **204,** 301–328 (1979).
12. J. Mayhew, "The interpretation of stereo-disparity information: the computation of surface orientation and depth," Perception **11,** 387–403 (1982).
13. C. H. Longuet-Higgins, "The role of the vertical dimension in stereoscopic vision," Perception **11,** 377–386 (1982).
14. F. W. Campbell and R. W. Gubisch, "Optical quality of the human eye," J. Physiol. (London) **186,** 558–578 (1966).
15. W. H. Miller, "Interocular filters," in *Handbook of Sensory Physiology*, H. Autrum, R. Jung, W. R. Loewenstein, H. L. Tueber, and D. M. Mackay, eds. (Springer-Verlag, Berlin, 1979), Vol. VII/6a, pp. 70–135.
16. G. Wyszecki and W. S. Stiles, *Color Science* (Wiley, New York, 1967).
17. D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966; Krieger, Melbourne, Fla., 1974).
18. G. Westheimer and S. P. McKee, "Integration regions for visual hyperacuity," Vision Res. **17,** 89–93 (1977).
19. W. P. Tanner and R. C. Clark-Jones, "The ideal sensor system as approached through statistical decision theory and the theory of signal detectability," in *Vision Research Problems*, A. Morrow and E. P. Horne, eds., NAS–NRC Publ. No. 712 (National Academy of Sciences, Washington, D.C., 1960).
20. A. Rose, "The relative sensitivities of television pick-up tubes, photographic film, and the human eye," Proc. IRE **30,** 293–300 (1942).
21. H. de Vries, "The quantum character of light and its bearing upon the threshold of vision, the differential sensitivity and acuity of the eye," Physica **10,** 553–564 (1943).
22. J. Yellot, "Spectral consequences of photoreceptor sampling in the rhesus retina," Science **22,** 382–385 (1983).
23. S. Shlaer, "The relation between visual acuity and illumination," J. Gen. Physiol. **21,** 165–188 (1937).
24. F. L. Van Nes and M. A. Bouman, "Spatial modulation transfer in the human eye," J. Opt. Soc. Am. **57,** 401–406 (1967).
25. L. A. Riggs, "Visual acuity," in *Vision and Visual Perception*, C. H. Graham, ed. (Wiley, New York, 1966), pp. 321–349.
26. F. W. Campbell and C. E. Legge, "Displacement detection in human vision," Vision Res. **21,** 205–214 (1981).
27. H. B. Barlow, "Increment thresholds at low intensities considered as signal/noise discriminations," J. Physiol. (London) **136,** 469–488 (1957).
28. M. C. Teich, P. R. Prucnal, G. Vannucci, M. E. Breton, and W. J. McGill, "Multiplication noise in the human visual system at threshold. 1. Quantum fluctuations and minimum detectable energy," J. Opt. Soc. Am. **72,** 419–431 (1982).
29. H. B. Barlow, "Retinal and central factors in human vision limited by noise," in *Photoreception in Vertebrates*, H. B. Barlow and P. Fatt, eds. (Academic, New York, 1977), Chap. 19, pp. 337–358.
30. W. S. Geisler, "Effects of bleaches and backgrounds on the flash response of the cone system," J. Physiol. (London) **312,** 413–434 (1981).
31. W. S. Geisler, "Mechanisms of visual sensitivity: backgrounds and early dark adaptation," Vision Res. **23,** 1423–1432.
32. H. B. Barlow, "Reconstructing the visual image in space and time," Nature **79,** 189–190 (1979).
33. H. B. Barlow, "Critical limiting factors in the design of the eye and visual cortex," Proc. R. Soc. London Ser. B **212,** 1–34 (1981).
34. R. J. Watt and M. J. Morgan, "Mechanisms responsible for the assessment of visual location: theory and evidence," Vision Res. **23,** 97–109 (1983).
35. Figures 5 and 7 show, for the present ideal detector, that two-point resolution follows an inverse fourth-root law and that two-point separation discrimination and vernier acuity follow an inverse square-root law. It will be proved in a subsequent paper that these relations are not true just for point sources but hold for essentially arbitrarily shaped stimuli. Thus, for any stimulus shape, the measurement of resolution and some hyperacuity as a function of luminance will provide a strong, parameter-free test of the quantum-fluctuations hypothesis.