UC Irvine

UC Irvine Previously Published Works

Title

Photon noise and constant-volume operators.

Permalink

<https://escholarship.org/uc/item/5p25w5kz>

Journal

Journal of the Optical Society of America. A, Optics and image science, 4(12)

ISSN

0740-3232

Author

Yellott, JI

Publication Date

1987-12-01

DOI

10.1364/josaa.4.002418

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Photon noise and constant-volume operators

John I. Yellott, Jr.

Department of Cognitive Sciences, University of California, Irvine, California 92717

Received April 3, 1987; accepted August 11, 1987

In an earlier paper [J. Opt. Soc. Am. A 2, 1769 (1985)] a class of nonlinear image processing operators was introduced in which each photoreceptor creates a nonnegative point-spread function whose center height is proportional to its quantum catch and whose volume is constant, so that the local spatial-summation area varies inversely with the local quantum catch. These constant-volume (CV) operators are designed to maximize spatial resolution in the presence of photon noise. In the previous paper it was shown that when CV operators are applied to deterministic images, they produce a surprising range of effects that are reminiscent of human vision, including Mach bands and Weber's-law behavior. In this paper the consequences of applying CV operators to images containing Poisson noise are analyzed. It is shown that a fixed-parameter CV operator can duplicate the global qualitative properties of spatial vision for retinal illuminances ranging from absolute threshold to 1000 Td. Although there are fundamental obstacles to modeling the exact quantitative properties of human spatial vision by CV operators, these operators seem likely to be useful in machine vision.

1. INTRODUCTION

This paper is a sequel to a recent paper by Cornsweet and Yellott,¹ which introduced a class of nonlinear image-processing operators based on the following idea: each point in the input image gives rise to a nonnegative point-spread function whose center height is proportional to the light intensity at that point and whose volume is constant, so that the area covered by the point spread, the local spatial-summation area, varies inversely with the input intensity at each point. The output image is the sum of the point-spread functions. (Figure 1 illustrates the case in which the point spread is Gaussian.) Analytically, the continuous version of this operation takes the general form

$$O[I](x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(u, v) S\{I(u, v)[(x - u)^2 + (y - v)^2]\} du dv, \quad (1.1)$$

where $I(u, v)$ is the value of an input image I at point (u, v) , $O[I](x, y)$ is the value of the output image $O[I]$ at point (x, y) , and S is a nonnegative real function (the *spread function*) that plays a role analogous to the impulse response of a linear operator.

Cornsweet and Yellott¹ referred to operators of form described by Eq. (1.1) as *intensity-dependent summation* (IDS) operators. Their paper began with a motivation of IDS operators in terms of their ability to deal efficiently with photon noise, in particular, to adjust automatically the size of the spatial-summation area according to the prevailing quantum catch, so as to maximize spatial resolution at every light level while maintaining a fixed reliability for spatial-contrast detection. Despite this motivation, however, the mathematical analysis of IDS operators in that paper deliberately ignored both the noisiness and the granularity of light: it treated the input image I as a deterministic function whose values range continuously over the nonnegative real numbers. In other words, the analysis derived the consequences of applying IDS operators to functions that represent the expected quantum catch at each point in an input

image rather than the catch itself. This restriction was motivated by mathematical expediency: it seemed sensible to begin the analysis of IDS operators with the simplest case and to defer an exact treatment of noisy images to a later paper.

In this paper the theory of IDS operators is extended to photon-noisy images, i.e., to the case in which the input values $I(u, v)$ are Poisson random variables, like the quantum catches of the photoreceptors. To highlight the similarities and the differences between operators of the form of Eq. (1.1) and linear operators, what we called IDS operators in the earlier paper are here renamed *constant-volume* (CV) operators. (From the standpoint of this new terminology, linear operators can be thought of as constant-area operators. This distinction is explained in Section 3.) Thus Eq. (1.1) is the general form of a spatially continuous CV operator, and Fig. 1 illustrates a spatially discrete CV operator whose spread function is Gaussian.

Cornsweet and Yellott's deterministic analysis focused on the consequences of applying CV operators to test images widely studied in psychophysics: edges, spots, and sinusoidal gratings. It showed that besides the expected effect of causing spatial summation, and thus, spatial resolution, to vary with input intensity, CV operators also automatically create a surprising range of additional effects that resemble important properties of human spatial vision, properties not usually associated with spatial-summation mechanisms. In particular, when CV operators are applied to deterministic images, they act as bandpass filters, so that Mach bands appear at edges; the peak amplitude of the Mach band at an edge separating regions of intensity L and $L + D$ depends only on the contrast ratio D/L , so that increment thresholds for spots could be expected to obey Weber's law. Moreover, the background intensity at which the detectability of any given spot begins to obey Weber's law depends on its size: the larger the spot, the sooner its increment threshold should begin to follow Weber's law. This is a property of human vision² that is hard to explain in terms of models for early visual processing in which the retinal image is first

Fig. 1. Discrete version of the Gaussian CV operator.

subjected to a pointwise nonlinear transformation (to model photoreceptor transduction and to predict Weber's-law behavior) and then convolved with a Mexican hat function [to model lateral inhibition and to predict bandpass-shaped contrast-sensitivity functions (CSF's)].

Overall, the main conclusion of the earlier paper was that while CV operators bear little resemblance to mechanisms commonly used to model retinal image processing (in particular, there is no lateral inhibition: the point-spread function is never negative³), they are remarkably adept at duplicating important features of human spatial vision. Consequently, it seemed worthwhile to extend their analysis to photon-noisy images, to determine how well a more realistic model of retinal image processing based on CV operators could match the quantitative properties of spatial vision.

In this paper we derive the statistical properties (pointwise means and variances) of the output images created by CV operators applied to photon-noisy versions of the input images dealt with in the earlier paper: edges, spots, and gratings. These results permit us to calculate the input image parameters (grating contrast, etc.) that would permit any test image to be discriminated from a uniform field with a given level of reliability and consequently to compare the predictions of a CV operator model with psychophysical data. These results also allow us to assess the potential usefulness of CV operators in artificial-image-processing applications, although that is not the primary focus of this paper.

Overview of Results

An important (and unexpected) result of this analysis is the discovery that when CV operators are applied to photon-noisy images, their consequences are qualitatively different at high and low light levels, a property that was not at all evident from the deterministic analysis of the earlier paper.¹ At high light levels (10 or more mean quanta/receptor) the expected output image for any noisy input image is essentially the same as the output image that would be produced by applying the same CV operator to the corresponding deterministic input image, i.e., the image whose intensity at each point is the mean quantum catch of the noisy image. Moreover, the variance at each output point is a constant that is independent of the input image (a constant that can be made arbitrarily small by a one-time setting of the scale parameter of the operator). Consequently, at high light levels the effects of CV operators on noisy images are essentially the same as their effects in the deterministic case: they still act as self-adjusting bandpass filters whose pass-

band varies with the retinal illuminance, and Weber's law still holds for any test spot once the background intensity reaches a critical level.

However, at very low light levels (when the mean quantum catch falls below 0.1 quantum/receptor), the intrinsic non-linearity of the CV operation has no opportunity to express itself, and every CV operator, no matter what its spread function, becomes effectively equivalent to some low-pass linear filter. Thus, at low light levels, CV operators applied to photon-noisy images have entirely different consequences from those that we would expect from the deterministic case: they no longer create Mach bands at edges or imply Weber's law. Instead, spot detectability obeys the deVries-Rose law.

From a psychophysical standpoint, this change from bandpass to low-pass filtering means that when photon noise is explicitly acknowledged, CV operators give a natural account of the fact that the shape of the human spatial CSF changes from bandpass to low pass as retinal illuminance falls to low levels.⁴ When only deterministic input images were considered, that effect could not be duplicated by any CV operator, so it is interesting to find that under more-realistic assumptions it proves to be an inevitable consequence of photon statistics.

Altogether, the present analysis shows that a simple one-parameter CV model can give at least a qualitative account of the major changes that occur in human spatial vision as the retinal illuminance varies between the absolute threshold [$\sim 10^{-4}$ troland (Td)] and 1000 Td. In the model considered here, photon-noisy retinal images (the quantum catches of the photoreceptors over a 250-msec time interval) are transformed by a fixed-parameter CV operator (the Gaussian operator G defined below in Section 3, with its scale parameter σ held constant across all illuminance levels), and it is assumed that any test image becomes discriminable from a uniform field when the peak value of d' across its output image reaches some fixed threshold level. This model correctly predicts that as the retinal illuminance rises from 10^{-4} to 10^3 Td, (1) Ricco's area shrinks⁵; (2) visual acuity rises (for gratings, by an overall factor on the order of 100, which matches the increase in human acuity⁶); (3) peak spatial-contrast sensitivity rises: the sensitivity at the best spatial frequency grows as the square root of the mean retinal illuminance up to 0.1 Td (Ref. 4) and reaches an asymptote of the order of 100 (threshold contrast $\approx 1\%$) at 10 Td (Ref. 4); (4) the shape of the spatial CSF changes from low pass to bandpass as the illuminance rises above 1 Td (Ref. 4); (5) increment thresholds obey the deVries-Rose law at low background illuminances and Weber's law at high ones; (6) the background illuminance at which Weber's law begins to hold is higher for small test spots than for large ones.² The model also correctly predicts that two sinusoidal gratings whose frequencies are both above the resolution limit at a given mean luminance level (and thus are invisible when viewed individually) can give rise to visible beats when superimposed.⁷

It seems rather remarkable that such a broad range of phenomena can all be created by a single self-adjusting mechanism driven only by photon noise. One gets a sense that, despite their lack of resemblance to familiar modeling devices, CV operators must somehow capture a fundamental property of retinal image processing. That would not be

entirely surprising, because it can be shown that an operator of the general form of Eq. (1.1) is, in a sense, axiomatically dictated by two fundamental problems faced by any visual system: the need to maximize spatial resolution in the face of photon noise and the need to map retinal illuminances spanning an enormous dynamic range into nerve signals with a relatively small one. (This argument is spelled out in Sections 2 and 3.)

However, a theory that only predicts the global qualitative properties of spatial vision is clearly inadequate: the real question is whether a model based on CV operators can make uniformly accurate quantitative predictions. The present analysis shows that this cannot be achieved by the bare-bones CV model considered here, a model in which the photoreceptor quantum catch is simply accumulated for a fixed period of time and then filtered by a CV operator. That model has two major flaws. First, at low retinal illuminances (<0.1 Td, the range where CV operators become effectively linear), the illuminance-related changes produced by the CV operation are too sluggish to match human performance. For example, at low light levels human visual acuity grows rapidly with retinal illuminance, nearly proportionally to its square root,⁶ whereas the Gaussian CV operator predicts a much slower rate of growth. This is a by-product of its effective linearity in this range, which causes it (and all CV operators) to predict incorrectly that at low light levels the spatial CSF should translate rigidly upward (in a log-log plot) as the mean retinal illuminance increases. [Experimentally, we find instead that between absolute threshold and 0.1 Td, the CSF shifts both upward and sideways (i.e., parallel to the log frequency axis) as illuminance rises.⁴]

The other major flaw appears at moderate-to-high light levels, where the nonlinear properties of the CV operation fully express themselves. Here, the problem is that the illuminance-related changes in contrast sensitivity produced by CV operators are too drastic to match human performance. For example, the Gaussian CV operator implies that above 10 Td the entire spatial CSF should shift rigidly along the (log) frequency axis as the mean retinal illuminance varies, so that both the peak of the CSF and its high-frequency cutoff (i.e., visual acuity) should grow proportionally with the square root of the retinal illuminance. (It seems virtually certain that this is true for any CV operator, regardless of its spread function, although mathematical scruples preclude a blanket assertion.) Neither of these effects occurs in human vision: above 10 Td the peak and the cutoff of the CSF for humans both change hardly at all.⁴ In addition, the CV model predicts that as retinal illuminance increases, contrast sensitivity at low spatial frequencies should decrease. That effect is never observed in human vision.

Of course the model analyzed here represents a rather primitive implementation of the basic CV idea: its linear analog would be a model with a single working part, a single linear filter, such as a difference of Gaussians, whose parameters are required to remain constant across all retinal illuminance levels. In comparison with that model, the achievements of the CV model analyzed here are quite impressive and might encourage us to construct more-elaborate theories in which the basic CV operation is supplemented by other mechanisms. For scotopic and mesopic vision,

this could be a useful exercise, because in that range the basic change that needs to be made seems fairly obvious, and it would be interesting to see whether it actually works. For photopic vision, on the other hand, I do not see any obvious way to design a viable model based on CV operators; the operation itself is not well adapted to perform in a high light environment, where photon noise is no longer a significant factor. (These points are discussed in Section 11.)

From the standpoint of psychophysical theory, then, the usefulness of CV operators seems at best problematic. Certainly it would be premature to say that such a basic mathematical tool will find no place in psychophysical theory. That would be like saying that linear operators are useless because they cannot model all the properties of spatial vision. The results presented here do not rule out the possibility that CV operators could be used to construct precise models of spatial vision for a restricted range of retinal illuminances. However, the results do show that there are fundamental obstacles that must be overcome for such a program to be successful. Even if such models could be devised, physiology suggests that the results would probably be more a formal exercise than a description of retinal reality. While it is surprisingly hard to find physiological evidence that clearly rules out the possibility of CV-like operations in the retina, what evidence does exist is negative, and recent studies suggest that the vertebrate retina achieves the same general goal (illuminance-dependent spatial filtering) by entirely different means. (These points are discussed in Section 11.)

For visual science, then, CV operators may well remain only theoretical curiosities: devices for modeling a retina that might have been but not the one that we have. For image engineering, however, their future seems more promising. The results presented here show that they provide a simple and effective algorithm for transforming photon-noisy input images whose mean intensities (and, consequently, signal-to-noise ratios) span a large range into output images that have a constant dynamic range, a noise level uniformly smaller than any desired upper bound, and a spatial resolution that automatically adjusts to match the prevailing light level. Although a discussion of specific applications is beyond the scope of this paper, it seems likely that such an algorithm could be useful in the design of surveillance systems and visual robots.

Organization of the Paper

In Sections 2–4 the stage is set for the analysis, and Sections 5–9 give the results. In Section 2 we review the constraints imposed on contrast detection by photon noise and show how they motivate an intensity-dependent spatial-summation operation located at the level of the photoreceptors, an operation implemented by intensity-dependent point-spread functions. In Section 3 we show how CV operators can be arrived at deductively, starting from a general family of variable-point-spread operators, by requiring two properties: (1) an area-intensity trade-off adapted to Poisson statistics and (2) dc suppression. As in the preceding paper, we focus here on a special case: the Gaussian CV operator. In Section 3 we define that operator, explain its unique mathematical convenience, and show that its output image values are uniformly bounded. Section 3 also includes an explanation of the relationship between physically realizable (i.e.,

discrete) CV operators and their continuous approximations, which are analytically indispensable but require tactful handling.

In Section 4 we review the mathematical properties of CV operators applied to deterministic input images, properties that were derived in the preceding paper.¹ These are still useful in the noisy-input case, because when the illuminance is not too low (10 or more mean quanta/receptor), the expected output image produced by a Gaussian CV operator for any photon-noisy input image is essentially the same as its output to the corresponding expected input image, i.e., the deterministic image obtained by replacing the input random variables with their expected values. In Section 5 we prove that convenient fact and also show that the output variance of a Gaussian CV operator has an upper bound (for all possible input images) that varies inversely with the square of its scale parameter. When the scale parameter is adjusted to fit psychophysical data, this upper bound proves to be small compared with the output values themselves: the maximum possible standard deviation is of the order of 2% of the mean output value, and for most images the actual standard deviation proves to be about 0.2% of the mean. In other words, the output images produced by this operator can be made virtually noise free.

In Sections 6–9 we deal with specific types of input images, the photon-noisy versions of the inputs considered in Ref. 1. In Section 6 we derive the mean and the variance of the response to uniform fields. Surprisingly, the mean response to a noisy uniform field with intensity (mean quanta/receptor) I is not a constant, as it is for deterministic images, but it grows as $1 - \exp(-I)$. This is the first indication of what proves to be a general rule: CV operators produce qualitatively different effects, depending on whether the illuminance is more or less than 10 quanta/receptor.

In Section 7 we deal with the response to edges. A key result of the previous paper was that in the deterministic case, edges in the input image create Mach bands in the output image, and the peak amplitudes of these Mach bands obey Weber's law. Here, it is shown that the same is true of the expected response to Poisson noisy edges once the input illuminance reaches 10 mean quanta/receptor. Moreover, the peak amplitude value is independent of the scale parameter of the CV operator, which means that the edge-response signal-to-noise ratio can be made arbitrarily large by adjusting that parameter (though at a cost in spatial resolution). Below 10 quanta/receptor the expected responses to Poisson noisy edges lose their Mach bands, which is a rather surprising result, since the deterministic case reveals no such effect: its edge response contains Mach bands whose amplitudes are constant (for a fixed contrast edge) across all input levels. The loss of Mach bands in the noisy-input case indicates that CV operators automatically transform themselves from bandpass to low-pass filters as the mean quantum catch falls to low levels.

In fact, below 0.1 mean quantum/receptor, CV operators become effectively equivalent to linear low-pass filters. In Section 8 this metamorphosis is shown directly in the response to sinusoidal gratings. For input contrasts of the order of 10% or less, CV operators act as linear operators at all light levels, returning sinusoidal outputs for sinusoidal inputs. (In the noisy-input case, this means that when the expected input image is a sinusoid, the expected output

image is a sinusoid.) Consequently, a modulation transfer function (MTF) can be defined, and, combined with the variances derived in Section 6, these MTF's can be used to predict spatial CSF's. Section 8 shows that the Gaussian CV operator's MTF undergoes two changes as the mean illuminance decreases: it shifts down along the frequency axis (indicating a loss in spatial resolution) and, below 10 quanta/receptor, its shape changes from bandpass to low pass. The same changes also appear in the CSF, and in addition (because of changes in the output noise level) it shows an overall loss of contrast sensitivity at mean illuminances of <1 quantum/receptor, a loss quite similar to the one observed in human CSF's. Section 8 concludes with a demonstration that CV operators also predict the fact that gratings that are individually invisible, because their frequencies are higher than the resolution limit at a given retinal illuminance level, give rise to visible beats when superimposed.

In Section 9 responses to spots (targets of the sort used in psychophysical increment threshold measurements) are discussed. The Gaussian CV operator's threshold-versus-background-intensity (TVI) curves, which obey the DeVries–Rose law (quantum-limited detection) at low background intensities and Weber's law at high ones, are discussed. If a small amount of dark light is assumed, to limit detection on zero backgrounds (i.e., absolute threshold), the CV operator's TVI curve provides a fairly good (but intrinsically never perfect) fit to human data. In Section 9 we also show how CV operators give rise to Ricco's law and cause the size of Ricco's area to shrink as background illuminance rises. Section 9 concludes with a calculation of the quantum efficiency of CV operators (i.e., of an observer whose input is a photon-noisy image filtered by a Gaussian CV operator).

In Section 10 some unsolved mathematical problems posed by CV operators are listed. In contrast to the deterministic case dealt with in the preceding paper,¹ the noisy-input case has not proved to be mathematically docile, and most of its properties are known analytically only for special cases (in particular, the Gaussian case) or, even worse, are known only empirically, from the results of simulations. Thus, while the main outlines of the consequences of image processing by CV operators now seem clear from a mixture of analysis and computation, many open problems still remain. Those described in Section 10 are only the ones that seem most immediate.

Finally, in Section 11 we summarize the successes and failures of the Gaussian CV operator as a model of image processing in the retina and discuss the problems involved in remedying its defects.

2. PHOTON NOISE AND SPATIAL SUMMATION

As Rose⁸ and deVries⁹ pointed out more than 40 years ago, contrast detection by any visual system is ultimately limited by the ability of the system to determine whether the mean of one Poisson random variable differs from that of another. For example, suppose that a patch of retina contains R photoreceptors and that its illuminance, the mean quantum catch per receptor, is I in one time interval and $I + CI$ ($C > 0$) in the next. Any sensible mechanism for detecting this increase will at least require the actual total catch in the

second interval to be greater than that in the first. Using that criterion alone (and the normal approximation to the Poisson distribution), the probability of detection is $N\{C[RI/(2+C)]^{1/2}\}$, where N is the normal distribution function:

$$N(z) = \int_{-\infty}^z [1/(2\pi)^{1/2} \exp(-x^2/2)] dx. \quad (2.1)$$

For this probability to be at least 0.99, it is necessary that $C[RI/(2+C)]^{1/2} \geq 2.3$, or, as an order-of-magnitude requirement,

$$RI > 10/C^2. \quad (2.2)$$

It is instructive to compare relation (2.2) with the parameters of human vision. The working range of the visual system is roughly -4 to $+6 \log \text{Td}$, and the mean quantum catch of the photoreceptors is about 4 per sec per receptor per troland.¹⁰ If 0.25 sec is taken as a conservative upper bound for the temporal integration period of photoreceptors, relation (2.2) implies that the minimum contrast that can reliably be detected from the quantum catch of a single receptor ranges from $C = 316$ at 10^{-4}Td through $C = 3.16$ at 1Td to $C = 0.003$ at 10^6Td . Human contrast thresholds are generally much lower than this (of course, for targets larger than a receptor): observers can detect contrasts of the order of $C = 1$ when retinal illuminance is 10^{-4}Td and of the order of $C = 0.01$ above 1Td .^{4,11} Clearly this would be impossible without some mechanism that, in effect, sums the quantum catches of many receptors, creating a signal that can satisfy the kind of constraint represented by relation (2.2).

On the other hand, spatial summation necessarily limits spatial resolution: a signal based on the total quantum catch of many receptors cannot carry information about the catch of any single one. Thus there is always a conflict between the ability to detect small contrasts and the ability to detect any contrast at all in small areas. Requirement (2.2) suggests resolving this conflict by causing the summation area R to vary inversely with the illuminance level I , permitting a fixed minimum contrast to be detected with a fixed reliability across all illuminance levels and simultaneously maximizing spatial resolution at any given level.

Assuming that intensity-dependent spatial summation is desirable, how should it be accomplished? The first design question is, Where should the size of the summation area be set: at the level of the summation units that collect signals from photoreceptors or at the level of the receptors themselves? Any attempt to implement the first solution must contend with an awkward "catch-22" dilemma: in order to adjust the size of its summation area according to the light level, a summation unit must estimate that level from signals it gets from receptors in its current summation area, but it cannot decide whether any given receptor should be within that area until it knows the light level. A more attractive option is to let each photoreceptor vote on the appropriate size of the summation area, guided by its own quantum catch, i.e., spread its signal over an area that varies according to that catch. In that case there is no catch-22 problem: no receptor needs to know anything more than its own input, and locally optimal spatial-summation areas can be created by parallel computations occurring simultaneously throughout the retina.

3. CONSTANT-VOLUME OPERATORS

In Section 2 we showed that photon noise motivates an intensity-dependent spatial-summation operation, and design considerations suggest implementing that operation at the level of the photoreceptor point-spread functions. At that early level it seems pointless to build in any orientation bias, so the operation should be rotation invariant, and parsimony dictates that it should also be translation invariant. Combining all four requirements, we are led to a class of image-processing operators of the general form

$$\begin{aligned} \mathbf{O}[I](x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F[I(u, v)] S[G[I(u, v)]][(x-u)^2 \\ + (y-v)^2] du dv, \end{aligned} \quad (3.1)$$

where $I(u, v)$ is the value of the input image I at point (u, v) [$I(u, v) \geq 0$, since it represents the light intensity]; $\mathbf{O}[I](x, y)$ is the value of the output image $\mathbf{O}[I]$ at point (x, y) ; S is an arbitrary real-valued spread function, which is assumed to be nonnegative to capture the idea of summation; and F and G are real functions that remain to be determined. This class includes both linear operators (the case in which F is the identity function and G is a constant) and the constant-volume operators defined in Section 1 (the case in which both F and G are the identity function).

We shall now specify F and G . Operator (3.1) describes a two-stage process in which each input point (u, v) first gives rise to a point-spread function whose value at the output point (x, y) is

$$F[I(u, v)] S[G[I(u, v)]][(x-u)^2 + (y-v)^2], \quad (3.2)$$

and the entire set of point-spread functions is then summed to produce the output image. The center height of the point-spread function around any input point (u, v) is $F[I(u, v)] S(0)$, and its volume is the integral of expression (3.2) over the (x, y) plane, which after the change of variable $x' = (x-u)$ $y' = (y-v)$ $\{G[I(u, v)]\}^{1/2}$, equals

$$\{F[I(u, v)]/G[I(u, v)]\} V_s, \quad (3.3)$$

where V_s is the constant given by

$$V_s = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(x^2 + y^2) dx dy. \quad (3.4)$$

If it is assumed that $S(0) > 0$, the equivalent area covered by the point-spread function around input point (u, v) is its volume divided by its center height, i.e.,

$$\text{area} = V_s/S(0)G[I(u, v)]. \quad (3.5)$$

The signal-detection argument in Section 2 implies that this area should vary inversely with the input intensity $I(u, v)$, and so we are led to the specialization $G(I) = I$.

This leaves F to be identified. Nothing about photon noise *per se* seems to force that choice, so its rationale must be sought elsewhere. One consideration here is to ensure that the output signals created by $\mathbf{O}$ should always fall within a fixed dynamic range, regardless of the actual retinal illuminance level (since that level spans a range of 10 log units, while the optic nerve has an effective working range of about 2). Also, at any fixed mean illuminance level, it seems sensible that the responses to input contrasts around that level should be able to exploit the entire response range.

This suggests that our operator should be designed to create dc suppression, i.e., to cause all uniform fields, whatever their intensity, to produce the same baseline response. If the input image I in Eq. (3.1) is a uniform field, i.e., $I(u, v)$ equals any positive constant L , then, after the change of variables $u' = (u - x)[G(L)]^{1/2}$, $v' = (v - y)[G(L)]^{1/2}$, the output image is also a uniform field whose value at every point is

$$[F(L)/G(L)]V_s. \quad (3.6)$$

Consequently, to guarantee dc suppression, we require $F(L)/G(L)$ to be a constant, which we can arbitrarily take to be 1. However, the uniform field response, expression (3.6), is simply the volume under the point-spread function when the input intensity is L [i.e., expression (3.3), with $I(u, v) = L$]. Thus we have shown that the only operators of the form of Eq. (3.1) that ensure dc suppression are those in which the volume under the point spread is a constant, independent of the input intensity. When that condition is coupled with the requirement that the point-spread area vary inversely with the input intensity, we are left with the class of operators

$$\begin{aligned} O[I](x, y) = & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(u, v) S\{I(u, v)[(x - u)^2 \\ & + (y - v)^2]\} du dv, \end{aligned} \quad (3.7)$$

i.e., the CV operators described in Section 1.

[Of course, all operators of the form of Eq. (3.1) with $F = G$ cause the point-spread volume to be constant. However, Eq. (3.7) can be considered the generic form of a CV operator: any other case of Eq. (3.1) with $F = G$ can be regarded as a two-stage operation in which the input image is subjected first to a transformation F and then to an operator of the form of Eq. (3.7).]

As was noted in Section 1, what we call CV operators here were called IDS operators in the preceding paper.¹ The reason for abandoning that name is that any operator of the form of Eq. (3.1) with G nonconstant can legitimately be regarded as an IDS operator, since the point-spread area [Eq. (3.5)] will vary with the input intensity. The antithesis of intensity-dependent spatial summation is the case in which G is constant, notably, the linear operators. In the same sense that Eq. (3.7) represents the generic form of a CV operator, linear operators represent the generic forms of constant-area operators, i.e., operators in which the spatial extent of the influence on the output image of any point in the input image is independent of the input value at that point.

Gaussian Constant-Volume Operators

In this paper, as in the preceding one, we focus mainly on a special case: the Gaussian CV operator, denoted by G and defined as follows:

$$\begin{aligned} G[I](x, y) = & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(u, v) (1/2\pi\sigma^2) \exp\{(-1/2\sigma^2)I(u, v) \\ & \times [(x - u)^2 + (y - v)^2]\} du dv. \end{aligned} \quad (3.8)$$

Here, the point-spread function around an input point (u, v) with an intensity $I(u, v)$ is a two-dimensional Gaussian

probability-density function whose equivalent area is $2\pi\sigma^2/I(u, v)$. The distance unit in both the input and output image planes is assumed to be the diameter of one photoreceptor, and so $2\pi\sigma^2/I(u, v)$ describes the size of the point spread in receptor areas. The scale parameter σ entirely determines the numerical properties of G ; sometimes we use the notation G_σ to make that dependence explicit. From a mathematical standpoint, the Gaussian is a uniquely convenient form of the point-spread function, because it is the only separable case, i.e., $S(x^2 + y^2) = S(x^2)S(y^2)$ for all x, y if and only if $S(x^2) = \exp(ax^2)$. (Strictly, this is true if S is continuous at least one point.¹²) The constants in Eq. (3.8) are chosen so that the volume under the point-spread function is 1.0. That choice is arbitrary: it simply sets the value of the response to uniform field input images, i.e., the value of $G[I](x, y)$ when the input image $I(u, v)$ is any positive constant.

Discrete Constant-Volume Operators

Continuous CV operators such as that described by Eq. (3.8) are analytically useful because they admit the power of the calculus, but from a physical standpoint they can be regarded only as approximations to discrete CV operators, operators that could actually be constructed. To appreciate the limits of these continuous approximations, we need a precise model of the discrete CV operators that they are supposed to emulate. Such operators can be described as follows. Imagine that an image falls upon a checkerboardlike array of densely packed square photoreceptors, each measuring 1×1 (so that the distance unit in the input image plane is one receptor diameter). The center of one receptor is taken to be the origin of the input plane, and each receptor is identified by its centerpoint coordinates (u, v) , where $u = 0, \pm 1, \pm 2, \dots$, and $v = 0, \pm 1, \pm 2, \dots$, relative to that origin. An input image I corresponds to some function mapping the set of receptor coordinates $\{(u, v)\}$ into nonnegative integers $\{I(u, v)\}$: the quantum catches of the receptors. Next, imagine a second plane, the output image plane, containing another densely packed checkerboard array of square summation units, each again 1×1 (so that the distance unit in the output image plane is one photoreceptor diameter). Each summation unit is identified by the coordinates of its centerpoint, say, (x, y) , and the summation unit (x, y) is thought of as lying directly below receptor (u, v) . When receptor (u, v) catches $I(u, v)$ quanta, it creates a point-spread function over the output image plane whose value at any point (p, g) is

$$I(u, v) S\{I(u, v)[(p - u)^2 + (q - v)^2]\}. \quad (3.9)$$

The summation unit centered at (x, y) integrates the point-spread function (3.9) over its 1×1 surface area, so that the point-spread contribution from receptor (u, v) to summation unit (x, y) is

$$\int_{y-0.5}^{y+0.5} \int_{x-0.5}^{x+0.5} I(u, v) S\{I(u, v)[(p - u)^2 + (q - v)^2]\} dp dq. \quad (3.10)$$

The total point-spread contribution from all the receptors to the summation unit (x, y) is the sum of expression (3.10) over all u, v , and that number is the output image value at coordinates (x, y) ; i.e., the discrete output image is exactly

$$O[I](x, y) = \sum_u \sum_v \int_{y-0.5}^{y+0.5} \int_{x-0.5}^{x+0.5} I(u, v) S\{I(u, v)[(p-u)^2 + (q-v)^2]\} dp dq. \quad (3.11)$$

Now we make approximations. If the point-spread function (3.9) is essentially constant over the 1×1 area of summation unit (x, y) , the integral expression (3.10) can be replaced by $I(u, v)S\{I(u, v)[(x-u)^2 + (y-v)^2]\}$. The output image then becomes

$$O[I](x, y) = \sum_u \sum_v I(u, v) S\{I(u, v)[(x-u)^2 + (y-v)^2]\}. \quad (3.12)$$

If it is assumed that $I(u, v)S\{I(u, v)[(x-u)^2 + (y-v)^2]\}$ is approximately constant over areas the size of a single receptor, the sum in Eq. (3.12) can be replaced by the integral Eq. (3.7), i.e., by a continuous CV operator.

Saturation

The error created by approximating Eq. (3.11) with Eq. (3.7) depends on the size of the input values $I(u, v)$ relative to the spatial extent of the basic spread function S . In general, as I becomes large, the area covered by a point-spread function $IS(Ir^2)$ shrinks, and eventually that function can no longer be treated as constant across a single summation unit. The most important practical consequence of this is that the continuous approximation [Eq. (3.7)] fails to reveal a saturation effect that limits the performance of any real CV operator at high light levels. Saturation occurs when the quantum catch of a receptor creates a point-spread function whose entire volume is confined to an area smaller than the receptor itself, smaller, that is, than the 1×1 area of the summation unit below that receptor. In that case, the true point-spread contribution from the receptor to the summation unit directly below it, i.e., the integral [Eq. (3.10)] for $x = u, y = v$, is the volume constant V_s , but the approximation treats it as $I(u, v)S(0)$, which is potentially unbounded. When the quantum catch at every receptor exceeds the saturation limit of a discrete CV operator, its true output image is a uniform field with the value V_s everywhere, so all contrast is lost. However, that loss will not necessarily be apparent from its continuous approximation, since the approximation envisages infinitely small receptors and summation units, which would make saturation impossible. For example, when the input is a grating of the form $L(1 + m \cos 2\pi fu)$, with $m^2 < 1$, and L grows without bound, any discrete CV operator will eventually saturate and produce a uniform output field for every frequency f . However its continuous approximation will disguise that fact and imply instead that the operator's high-frequency cutoff simply increases without bound.

To avoid such pitfalls, we must know the input level at which a discrete CV operator will begin to show significant saturation effects. Here, we are concerned chiefly with the Gaussian operator G approximated by Eq. (3.8), and in all numerical examples its parameter σ is taken to be 100. (That choice is based on psychophysical modeling considerations discussed below.) Analysis and computational experience show that this operator begins to saturate when the input level reaches about 10^4 quanta/receptor. (At that

level the equivalent area of the point-spread function is roughly 6 receptors, and the point spread from a receptor to its own summation unit is about 15% of the total spread.) Consequently, its continuous approximation will be used only for images of considerably $<10^4$ quanta/receptor.

Integer-Valued Inputs and Bounded Output Images

Whether we deal with discrete CV operators or their continuous approximations, it is important to bear in mind that the input image values $I(u, v)$ are receptor quantum catches and consequently must be integers, never fractions. The importance of this constraint can be seen if we attempt to determine what the maximum output value can be for a given CV operator, e.g., the Gaussian case [Eq. (3.8)]. It was noted above that the response of this operator to all uniform field inputs is a constant uniform field with a value of 1.0, and in Ref. 1 it was found that the maximum response of the operator to a wide variety of input images was always less than 2.0. On that basis it was asserted, without proof, that the Gaussian CV operator compresses all possible input images into a common finite range of output values. But when we try to prove that claim, an apparent problem arises. Suppose that we seek to construct the input image I that will create the maximum possible value of $G[I](x, y)$ in Eq. (3.8). Without a loss of generality we can consider only the output point $(x, y) = (0, 0)$, and because of the circular symmetry of the point-spread functions we need consider only circular symmetric input images: $I(u, v) = I(r)$, where $r^2 = u^2 + v^2$. When it is converted to polar coordinates, Eq. (3.8) becomes

$$G[I](0, 0) = \int_0^\infty [I(r)r/\sigma^2] \exp\{(-1/2)I(r)r^2/\sigma^2\} dr, \quad (3.13)$$

and we seek the $I(r)$ that maximizes Eq. (3.13). Differentiation shows that the integrand $[I(r)r/\sigma^2] \exp\{(-1/2)I(r)r^2/\sigma^2\}$ is maximized at every r by the input image $I(r) = 2\sigma^2/r^2$, and for that input Eq. (3.13) becomes

$$G[I](0, 0) = \int_0^\infty (2/r) \exp(-1) dr, \quad (3.14)$$

which is infinite, a disturbing result. Of course, the solution of Eq. (3.14) is unrealistic because it fails to represent the fact that the unbounded input image $2\sigma^2/r^2$ would saturate the receptor at the origin, making its output 1 rather than ∞ . To correct for this, we can break Eq. (3.14) into two parts, one giving a realistic account of the point-spread contribution of the receptor at the origin and the other giving the total contribution of all other receptors:

$$G[I](0, 0) = 1 + \int_{1/2}^\infty (2/r) \exp(-1) dr. \quad (3.15)$$

However, this concession to realism does not solve the problem; the maximum output is still infinite. In fact, any input image of the form $I(r) = K$, for $r \leq R$, and $I(r) = (2\sigma^2/r^2)$, for $r > R$, will make Eq. (3.13) infinite, since $G[I](0, 0)$ will be

$$1 - \exp\{(-1/2)KR^2/\sigma^2\} + \int_R^\infty (2/r) \exp(-1) dr. \quad (3.16)$$

Thus it is not the large values of the maximal input image $(2\sigma^2/r^2)$ near the origin that cause the output to be infinite; rather, it is the small values of $2\sigma^2/r^2$ far from the origin. And therein lies the fallacy of the solution. Since the values

of any real input image must be integers, the maximizing input function $I(r) = (2\sigma^2/r^2)$ describes a possible input only when $r \leq \sigma\sqrt{2}$. For larger values of r , the integer-valued $I(r)$ that maximizes the integrand in Eq. (3.13) is $I(r) = 1$. Consequently, the maximum output value for any physically possible input image is in fact bounded:

$$\begin{aligned} G[I_{\max}](0, 0) &\leq 1 + \int_{1/2}^{\sigma 2^{1/2}} (2/r) \exp(-1) dr \\ &\quad + \int_{\sigma 2^{1/2}}^{\infty} (r/\sigma^2) \exp[(-1/2\sigma^2)r^2] dr \\ &= 1 + 2 \exp(-1) [\ln(\sigma 2^{1/2}) - \ln(1/2)] \\ &\quad + \exp(-1). \end{aligned} \quad (3.17)$$

For $\sigma = 100$, this upper bound is 5.5. [The upper bound of 2.0 that is conjectured, more or less on empirical grounds, in Ref. 1 proves to be not far wrong, because the analysis in that paper focused on the case $\sigma = 1$. For that value of σ , relation (3.17) yields an upper bound of 2.1.] Analysis and computational experience with the case $\sigma = 100$ show that most input images create output values in the range 0–2.

4. RESPONSES TO DETERMINISTIC INPUT IMAGES

To understand the effects of CV operators on photon-noisy images, integer-valued stochastic processes, it is analytically convenient to begin with their responses to deterministic input images whose values need not be integers, i.e., arbitrary nonnegative real functions $I(u, v)$. These can be thought of as the expectations of actual noisy images. Cornsweet and Yellott¹ derived the main results of applying CV operators to such images. Their analysis focused on the Gaussian operator G [i.e., Eq. (3.8)] but emphasized that that case is unique only in its mathematical convenience: all CV operators share the same general properties. To demonstrate this, Cornsweet and Yellott proved a number of theorems that apply to all operators of the form of Eq. (1.1), independent of the exact form of the spread function S . In this section we review some of these results for later use. (The first two are obvious consequences of the construction of CV operators. All proofs can be found in Ref. 1.)

Result 1: Invariance under Translation and Rotation

For all CV operators, translating or rotating the input image by any amount leaves the output image unchanged except for translation or rotation by the same amount.

Result 2: Uniform Field Response

For all CV operators, the output image produced by any nonzero uniform field input $I(u, v) \equiv L$ is a uniform field whose value at every point is the volume constant V_s given by expression (3.4). (For the Gaussian operator G , $V_s = 1$.)

Result 3: Scaling Theorem

For any positive constant k and any input image $I(u, v)$,

$$O[kI(u, v)](x, y) = O[I(u/k^{1/2}, v/k^{1/2})](xk^{1/2}, yk^{1/2}). \quad (4.1)$$

This result is the technical key to the general properties of CV operators. It means that the effect of multiplying all the input intensities by a constant k is the same as rescaling the

Fig. 2. Edge-response profiles for the Gaussian CV operator G_{100} applied to deterministic input images [Eq. (4.2)].

original image by the factor $k^{1/2}$ on both dimensions, applying O to that image, and then rescaling the output image by the factor $1/k^{1/2}$, i.e., back to the original size. Weber's-law behavior and many other properties follow from this.

Result 4: Edge Responses, Mach Bands, and Weber's Law

When the input image is an edge of the form $I(u, v) = L$, for $u \leq 0$, and $I(u, v) = L + D$, for $u > 0$ (i.e., a step), the output of the Gaussian CV operator is

$$G[I](x, y) = N[(x/\sigma)(L + D)^{1/2}] + N[-(x/\sigma)L^{1/2}], \quad (4.2)$$

where N is the normal integral [Eq. (2.1)]. Figure 2 illustrates the response of G (with the parameter $\sigma = 100$) to three different input edges. Here L ranges from 1 to 100 quanta/receptor, but the contrast D/L is fixed at 0.5. It can be seen that all the responses exhibit Mach bands, and the peak and trough amplitudes of these Mach bands are the same in all cases. Analysis of Eq. (4.2) shows that the maximum response occurs at $x_{\max} = \sigma[(1/D)\ln(1 + D/L)]^{1/2}$, and the output value there is independent of the scale parameter σ and the baseline illuminance L :

$$\begin{aligned} G(x_{\max}) &= N\{[(1 + L/D)\ln(1 + D/L)]^{1/2}\} \\ &\quad + N\{-(L/D)\ln(1 + D/L)\}^{1/2}, \end{aligned} \quad (4.3)$$

which depends only on the Weber fraction D/L . The minimum response occurs at $-x_{\max}$ and falls as far below the baseline response value (i.e., 1.0) as the maximum is above it. In other words, the peak and trough values of the Mach bands in any edge response always obey Weber's law. This is a general property of all CV operators. Another general property is that the distances of the peak and the trough of the Mach bands from the edge itself vary as $1/L^{1/2}$. This suggests that spatial resolution will vary as $L^{1/2}$, which proves to be true.

(In view of result 1, all the statements in result 4 apply to any edge, whatever its location and orientation.)

Result 5: Responses to Sinusoids

Since CV operators are nonlinear, they must exhibit some harmonic distortion, and Fig. 3 shows the form that distortion takes in the Gaussian case when the input is a high-

Fig. 3. Response profiles for G_{100} applied to deterministic sinusoidal gratings of high (top) and low (bottom) contrast. The distance unit is photoreceptor diameters.

Fig. 4. MTF's [expression (4.5)] of G_{100} for deterministic input images at different retinal illuminance levels.

contrast sinusoidal grating. However, as Fig. 3 also shows, when the input contrast is 10%, the output appears quite sinusoidal. Analysis confirms that impression: when G is applied to an input of the form $I(u, v) = L(1 + m \cos 2\pi fu)$ and m is 0.1 or less, the output image is approximately

$$G[I](x, y) = 1 + (2\pi^2\sigma^2 f^2/L) \exp(-2\pi^2\sigma^2 f^2/L) (m \cos 2\pi fx), \quad (4.4)$$

with an error on the order of m^2 . Consequently, for low-contrast inputs it is sensible to speak of the MTF of the Gaussian CV operator: the ratio of output contrast to input contrast as a function of the input frequency f . Equation (4.4) shows that this MTF is

$$(2\pi^2\sigma^2 f^2/L) \exp(-\pi^2\sigma^2 f^2/L). \quad (4.5)$$

By a remarkable coincidence, at any fixed mean illuminance L this MTF is the same as that of Marr and Hildreth's¹³ $\text{del}^2 - G$ operator, i.e., the linear operator whose impulse response is the negative Laplacian of a Gaussian. However, in contrast to the MTF of a linear operator, expression (4.5) changes with the mean illuminance: plotted in log-log coordinates, as shown in Fig. 4, it shifts bodily along the frequency axis as L changes, so that the best frequency is $(1/\sigma\pi^{1/2})L^{1/2}$, and visual acuity (defined as the highest frequency at

Fig. 5. Response profiles (along the x axis) for G_{100} applied to a deterministic square spot of intensity $L + D$ surrounded by a background of intensity L [Eq. (4.6)]. The spot width W is 100 receptor diameters. D was adjusted to produce a constant peak response amplitude of 1.15.

which the MTF exceeds any fixed threshold) varies directly as $L^{1/2}$. The MTF's of all CV operators vary with illuminance in this same way, and so all imply that acuity grows as $L^{1/2}$.

Psychophysics shows that human visual acuity for gratings does grow roughly as the square root of retinal illuminance⁶ (although only up to about 10 Td). Thus CV operators give a fairly accurate account of that aspect of visual performance. However, the human spatial CSF does not maintain a fixed shape across all light levels and simply shift left or right along the (log) frequency axis as the mean illuminance changes. Instead, as the illuminance increases from zero, the CSF initially shifts but also rises (sensitivity increases at all frequencies) and then, at around 1 Td, changes shape from low pass to bandpass.⁴ Beyond 10 Td, the peak frequency of the CSF remains quite constant at roughly 5 cycles/deg. Applied to deterministic images, no CV operator can create a MTF that rises or changes shape with the mean illuminance, and so when only such images were considered it seemed impossible for CV operators to model those two properties of human spatial vision. As we shall see, both prove to be natural consequences of CV operators applied to photon-noisy images.

Result 6: Spot Responses and Increment Thresholds

When the input image I is a square spot of width W and intensity $L + D$, centered at the origin, and surrounded by a uniform background field of intensity L , the output image of the Gaussian CV operator is

$$\begin{aligned} G[I](x, y) = & 1 + \{[N[A(W/2 - x)] - N[-A(W/2 + x)]] \\ & \times [N[A(W/2 - y)] - N[-A(W/2 + y)]]\} \\ & - \{[N[B(W/2 - x)] - N[-B(W/2 + x)]] \\ & \times [N[B(W/2 - y)] - N[-B(W/2 + y)]]\}, \quad (4.6) \end{aligned}$$

where $A = (1/\sigma)(L + D)^{1/2}$ and $B = (1/\sigma)L^{1/2}$. Figure 5 shows profiles (across the x axis) of the response to spots of a fixed size for three different background intensities. The increment D was adjusted to keep the peak response value constant, simulating an increment threshold measurement. It can be seen that as the background intensity rises, the Mach bands at the edges of the spot become narrower and eventually cease to overlap. Once that level is reached, the peak response value always occurs at the peaks of the Mach bands, and the size of that peak obeys Weber's law. Consequently, a plot of the threshold value of D against the background intensity L (a TVI curve) will indicate that increment thresholds obey Weber's law above some critical background intensity level, a level that is higher for smaller spot sizes, as illustrated in Fig. 6. Human TVI curves show a qualitatively similar behavior.² Equation (4.6) can also be used to show that the increment threshold obeys Ricco's law for spot sizes smaller than a critical area, an area that shrinks as the background intensity increases. (Figure 10 in Ref. 1 illustrates that effect.) Of course, that is one of the properties of human vision⁵ that originally motivated our interest in CV operators.

Altogether, then, when CV operators are applied to deterministic images, they create a wide range of effects resembling well-known properties of human spatial vision: Mach bands, Weber's-law behavior, Ricco's-law behavior, and J -

shaped TVI curves. However, because these images are noise free, there is nothing in the model at this stage that creates a natural threshold; i.e., there are no intrinsic limits to contrast detection. Consequently the choice of a threshold response value is arbitrary, and meaningful comparisons with psychophysical data are impossible. To make such comparisons, we need to introduce photon noise and then determine the input parameters required to produce threshold-level signal-to-noise ratios. In the next section we begin that process.

5. NOISY IMAGES: NOTATION AND PRELIMINARY RESULTS

We now apply CV operators to photon-noisy input images, i.e., images in which the input values $I(u, v)$ are Poisson random variables. To distinguish between these random inputs and deterministic ones, $Q(u, v)$ is used to denote the random variable corresponding to the quantum catch at the input point (u, v) . The expected value $E\{Q(u, v)\}$ is denoted by $q(u, v)$. Thus a noisy input image Q is a stochastic process $\{Q(u, v): -\infty < u < \infty, -\infty < v < \infty\}$, with

$$P[Q(u, v) = n] = \exp[-q(u, v)] [q(u, v)]^n / n!, \quad n = 0, 1, 2, \dots, \quad (5.1)$$

so that $E\{Q(u, v)\} = \text{Var}\{Q(u, v)\} = q(u, v)$. The random variables $Q(u, v)$ are assumed to be mutually independent. The output image corresponding to input Q is the stochastic process $\{O[Q](x, y): -\infty < x < \infty, -\infty < y < \infty\}$ created from Q by the operation

$$\begin{aligned} O[Q](x, y) = & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Q(u, v) S\{Q(u, v)\} [(x - u)^2 \\ & + (y - v)^2] du dv. \quad (5.2) \end{aligned}$$

In particular, the output of the Gaussian CV operator G_σ is

Fig. 6. Deterministic TVI curves for G_{100} for test spots of different widths (W denotes the spot width in receptor diameters). Each curve shows the increment D required to produce a constant peak response value of 1.15 when G_{100} is applied to a deterministic square spot of intensity $L + D$ surrounded by a background of intensity L .

$$G_\sigma[Q](x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [Q(u, v)/2\pi\sigma^2] \exp\{(-Q(u, v)/2\sigma^2) \\ \times [(x - u)^2 + (y - v)^2]\} dudv \quad (5.3)$$

Interest centers on the mean and the variance of the output values, i.e., $E\{O[Q](x, y)\}$ and $\text{Var}\{O[Q](x, y)\}$. For the Gaussian operator G , the following theorem shows that when the mean input values $q(x, y)$ are not too small, the expected output image $E\{G[Q](x, y)\}$ is closely approximated by $G[q](x, y)$, i.e., the output image that would be created by applying the Gaussian operator G to the expected input image $q(u, v)$. This permits us to apply results already obtained for deterministic input images to a significant number of noisy input images—a great convenience, because the former are typically simple closed-form expressions whose implications are relatively easy to grasp, whereas the exact expressions for expected output images usually involve infinite series whose implications are not apparent from inspection.

Theorem 1

If $q(u, v) \geq 10$ for all (u, v) ,

$$E\{G[Q](x, y)\} \approx G[q](x, y), \quad (5.4)$$

with an error of at most 0.045.

To put this value in context, recall that the uniform-field response of G is 1.0, and for $\sigma = 100$, all response values fall between 0 and 5.5, with most between 0 and 2. Computational experience shows that the error created by approximation (5.4) is generally much less than 0.045: that maximum value corresponds to the worst possible case, as will appear in the proof. As q increases above 10, the maximum error decreases quickly.

Proof

We proceed by constructing the input image that maximizes the difference between the two sides of relation (5.4), and we show that the difference is at most 0.045. For this purpose it is sufficient to consider an arbitrary output point (x, y) , and it is convenient to choose $(x, y) = (0, 0)$. If we write q for $q(u, v)$, Q for $Q(u, v)$, and r^2 for $u^2 + v^2$ to reduce the notational burden, the right-hand side of relation (5.4) becomes

$$G[q](0, 0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (q/2\pi\sigma^2) \exp(-qr^2/2\sigma^2) dudv, \quad (5.5)$$

and the left-hand side becomes

$$\begin{aligned} E\{G[Q](0, 0)\} &= E\left\{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (Q/2\pi\sigma^2) \exp(-Qr^2/2\sigma^2) dudv\right\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E\{(Q/2\pi\sigma^2) \exp(-Qr^2/2\sigma^2)\} dudv \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (1/2\pi\sigma^2) \left\{ \sum_{n=0}^{\infty} [n \exp(-nr^2/2\sigma^2)] \right. \\ &\quad \times \exp(-q)(q^n/n!) \left. \right\} dudv \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (q/2\pi\sigma^2) \exp(-r^2/2\sigma^2) \exp[(-q) \\ &\quad \times (1 - \exp(-r^2/2\sigma^2))] dudv. \end{aligned} \quad (5.6)$$

The difference $G[q](0, 0) - E\{G[Q](0, 0)\}$ [i.e., Eq. (5.5) minus Eq. (5.6)] is then

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (q/2\pi\sigma^2) (\exp(-qr^2/2\sigma^2) - q \exp(-r^2/2\sigma^2) \exp\{(-q)[1 - \exp(-r^2/2\sigma^2)]\}) dudv, \quad (5.7)$$

and we seek the expected input image, $q(u, v)$, that maximizes the absolute value of expression (5.7). We can break expression (5.7) into two integrals, one (called A) for (u, v) values such that $(r^2/2\sigma^2) \leq 0.1$ and the other (B) for $(r^2/2\sigma^2) > 0.1$. In integral A the quantity $\exp(-r^2/2\sigma^2)$ can be replaced by its close approximation $1 - (r^2/2\sigma^2)$, and with that substitution we have

$$A = \int \int_{(r^2/2\sigma^2) \leq 0.1} (r^2/2\sigma^2) (q/2\pi\sigma^2) \exp(-qr^2/2\sigma^2) dudv.$$

Clearly, the input image $q(u, v)$ that maximizes A will depend only on r , and differentiation shows that the integrand in A is maximized at every r by the input image $q(u, v) = 2\sigma^2/r^2$. Evaluation of the integral for that q shows that $A_{\max} = 0.037$. The minimum possible value for A is obviously zero, produced by $q = 0$.

Finally, we need to maximize integral B . A computer search shows that for $r^2/2\sigma^2 > 0.1$ and $q \geq 10$, the integrand in expression (5.7) is nonnegative for $r^2/2\sigma^2 \leq 0.16$; in that range its maximum value is produced by $q = 10$, and the minimum value (0) is produced when q is any constant of the order of 30 or more. For $r^2/2\sigma^2 > 0.16$ the integrand is never positive, and its maximum value is 0, obtained when q is any constant on the order of 30 or more. Numerical integration shows that when the minimum possible integrand is used for all r in this range, the total value of the integral is not less than -0.034 . Consequently, this negative contribution cannot cancel the positive contribution from A . It follows that the total error $|A + B|$ is maximized by setting $q = 10$ for $0.1 < r^2/2\sigma^2 \leq 0.16$ and $q = 30$ for $0.16 < r^2/2\sigma^2$. For that image, $B_{\max} = 0.008$, and when this is added to $A_{\max}$, the maximum total error is 0.045. ■

When $q(u, v)$ falls below 10, approximation (5.4) rapidly becomes a poor one: the expected response of G can no longer be predicted from $G[q]$.

From a practical standpoint, the expected output image $E\{G[Q]\}$ is interesting only to the extent that it is a good predictor of the actual output images $G[Q]$ that will be generated by any illuminance function q . Theorem 2 shows that the scale parameter σ can be adjusted to make the variance of $G[Q]$ arbitrarily small across all input images, and so the accuracy of that prediction can be made arbitrarily good. In other words, the output image noise level can be preset below any required bound by a one-team adjustment of the parameter σ .

Theorem 2

Let G_σ denote the Gaussian CV operator with the scale parameter σ . For any expected input image $q(u, v)$, the variance of the output value $G_\sigma[Q](x, y)$ satisfies the inequality

$$\text{Var}\{G_\sigma[Q](x, y)\} \leq 4.54/\sigma^2. \quad (5.8)$$

It will be shown below that to match certain human psychophysical data, σ must be around 100. For that σ , the upper

bound on the right-hand side of relation (5.8) implies a maximum standard deviation of 0.02, and so most output values will be close to their expectations: the output noise will be on the order of 2% of the mean. [In fact, both analysis and computer simulation suggest that the upper bound given by relation (5.8) is too large by a factor of 10: for most input images the output standard deviations are of the order of 0.002. A sharper theorem seems in order here, although it is not easy to see how to prove one.] It is interesting to compare this 2% value with the noise of the input image. The operator G_{100} has a working range (i.e., is below saturation) from 0 to about 10^4 mean quanta/receptor, and throughout this range its output values have a standard deviation less than 0.02. The standard deviation of the receptor quantum catch is $q^{1/2}$ when the mean is q , so its noise-to-signal ratio (i.e., $1/q^{1/2}$) reaches 0.02 only when $q = 2500$ quanta/receptor, that is, when the retinal illuminance is of the order of 1000 Td. To reach the noise-to-signal ratio of 0.002, which characterizes G_{100} for most input images, requires 2.5×10^5 quanta/receptor, a retinal illuminance of more than 10^5 Td. In comparison with its input images, it seems fair to describe the output images created by G_{100} as virtually noise free. (Figures 9 and 10 below provide a direct comparison between input noise and output noise for this CV operator.)

Proof

It is sufficient to show that relation (5.8) holds for the particular output point $(x, y) = (0, 0)$ regardless of the mean input image $q(u, v)$. As a first step, we prove the following: for any input Q with mean values $q(u, v)$,

$$\text{Var}\{G_\sigma[Q(u, v)](0, 0)\} = (1/\sigma^2)\text{Var}\{G_1[Q(\sigma u, \sigma v)](0, 0)\}. \quad (5.9)$$

To show this, we begin by evaluating the left-hand side:

$$\begin{aligned} \text{Var}\{G_\sigma[Q(u, v)](0, 0)\} &= \text{Var}\left\{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [Q(u, v)/2\pi\sigma^2] \right. \\ &\quad \times \exp[-Q(u, v)/2\sigma^2](u^2 + v^2) du dv \Big\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{Var}\{[Q(u, v)/2\pi\sigma^2] \\ &\quad \times \exp[-Q(u, v)/2\sigma^2](u^2 + v^2)\} du dv \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (1/2\pi\sigma^2)^2 \text{Var}\{Q(u, v) \\ &\quad \times \exp[-Q(u, v)/2\sigma^2](u^2 + v^2)\} du dv, \end{aligned} \quad (5.10)$$

since the random variables $Q(u, v)\exp[-Q(u, v)/2\sigma^2](u^2 + v^2)$ are mutually independent.

To save notation, we now write Q for $Q(u, v)$, q for $q(u, v)$, and r^2 for $u^2 + v^2$ and omit the limits of integration, which are always the entire (u, v) plane. Using the fact that $\text{Var}\{X\} = E\{X^2\} - E^2\{X\}$, we can rewrite Eq. (5.10) in the form

$$\begin{aligned} \text{Var}\{G_\sigma[Q](0, 0)\} &= \int \int (1/2\pi\sigma^2)^2 E\{[Q \exp(-Qr^2/2\sigma^2)]^2\} du dv \\ &\quad - \int \int (1/2\pi\sigma^2)^2 E^2\{Q \exp(-Qr^2/2\sigma^2)\} du dv. \end{aligned} \quad (5.11)$$

This quantity will depend on the expected input image $q(u, v)$ and the scale parameter σ . Let $A(\sigma, q)$ denote the first integral in Eq. (5.11), and let $B(\sigma, q)$ denote the second (so that the variance is $A(\sigma, q) - B(\sigma, q)$). Then

$$\begin{aligned} A(\sigma, q) &= \int \int (1/2\pi\sigma^2)^2 \left[\sum_{n=1}^{\infty} n^2 \exp(-nr^2/\sigma^2) q^n \right. \\ &\quad \times \exp(-q)/n! \Big] du dv \\ &= \int \int (1/2\pi\sigma^2)^2 q \exp[-q - (r^2/\sigma^2) + q \exp(-r^2/\sigma^2)] \\ &\quad \times \left\{ \sum_{n=0}^{\infty} [(n+1)/n!] p^n \exp(-p) \right\} du dv \\ &= \int \int (1/2\pi\sigma^2)^2 q \exp[-q - (r^2/\sigma^2) + q \exp(-r^2/\sigma^2)] \\ &\quad \times [1 + q \exp(-r^2/\sigma^2)] du dv, \end{aligned} \quad (5.12)$$

where $p = q \exp(-r^2/\sigma^2)$. After the change of variables $s = u/\sigma$ and $t = v/\sigma$, Eq. (5.12) becomes

$$\begin{aligned} A(\sigma, q) &= (1/\sigma^2) \int \int (1/2\pi)^2 q' \exp[-q' - v^2 + q' \exp(-v^2)] \\ &\quad \times [1 + q' \exp(-v^2)] ds dt \\ &= (1/\sigma^2) A(1, q'), \end{aligned} \quad (5.13)$$

where $q'(s, t) = q(\sigma s, \sigma t)$ and $v^2 = s^2 + t^2$.

Next we apply the same process to $B(\sigma, q)$:

$$\begin{aligned} B(\sigma, q) &= \int \int (1/2\pi\sigma^2)^2 \left[\sum_{n=1}^{\infty} n \exp(-nr^2/2\sigma^2) q^n \right. \\ &\quad \times \exp(-q)/n! \Big]^2 du dv \\ &= \int \int (1/2\pi\sigma^2)^2 \{q \exp[-q - (r^2/2\sigma^2) \\ &\quad + \exp(-r^2/2\sigma^2)]\}^2 du dv \\ &= (1/\sigma^2) \int \int (1/2\pi)^2 \{q' \exp[-q' - (v^2/2) \\ &\quad + q' \exp(-v^2/2)]\}^2 ds dt, \\ &= (1/\sigma^2) B(1, q'), \end{aligned} \quad (5.14)$$

where q' and v are defined as for Eq. (5.13). Combining Eqs. (5.13) and (5.14), we have

$$\begin{aligned} \text{Var}\{G_\sigma[Q(u, v)](0, 0)\} &= A(\sigma, q) - B(\sigma, q) \\ &= (1/\sigma^2)[A(1, q') - B(1, q')] \\ &= (1/\sigma^2)\text{Var}\{G_1[Q(\sigma u, \sigma v)](0, 0)\}; \end{aligned}$$

thus Eq. (5.9) is proved. It remains only to establish an upper bound on $\text{Var}\{G_1[Q](\sigma u, \sigma v)\}$. This is the same as finding an upper bound for $\text{Var}\{G_1[Q](u, v)\}$, since the set of all expected input images $\{q(\sigma u, \sigma v)\}$ is the same as the set of all input images $\{q(u, v)\}$. For any random variable X , $\text{Var}\{X\} = E\{X^2\} - E^2\{X\} \leq E\{X^2\}$, and so we know at least that

$$\text{Var}\{G_1[Q(\sigma u, \sigma v)](0, 0) \leq \max E\{[G_1[Q(u, v)](0, 0)]^2\}. \quad (5.15)$$

We can now appeal to the upper bound on the possible values of $G[Q]$ that was established in Section 3, i.e., relation (3.17). Evaluated for $\sigma = 1$, relation (3.17) shows that $G_1[Q](0, 0)$ cannot exceed 2.13, and so

$$\max E\{[G_1[Q(u, v)](0, 0)]^2\} \leq (2.13)^2 = 4.54.$$

By making this substitution in relation (5.15) and applying Eq. (5.9), we obtain

$$\text{Var}\{G_\sigma[Q](0, 0)\} \leq 4.54/\sigma^2.$$

Thus theorem 2 is proved. ■

(For the purpose of establishing the smallest possible upper bound on the variance, the defect of this proof is its use of the weak inequality $\text{Var}\{G\} \leq E\{G^2\}$. A sharper result could probably be obtained by somehow exploiting the fact that $E^2\{G\}$ is generally nonzero.)

6. UNIFORM-FIELD RESPONSE

Recall from Section 4 that for any CV operator, all uniform-field input images produce the same output image: a uniform field whose intensity is the volume constant V_s given by Eq. (3.4). That property was built into CV operators during their construction in Section 3. Consequently, it is a bit surprising to find that, for noisy uniform fields, the expected output image is not constant but increases (to an asymptote of V_s) as retinal illuminance rises.

Theorem 3

If the expected input image $q(u, v)$ equals a constant L for all (u, v) , the expected output image of any CV operator is

$$E\{O[Q](x, y)\} = [1 - \exp(-L)]V_s. \quad (6.1)$$

Proof

Since the expected input image is uniform, the expected output image values will be the same for all (x, y) , and so for convenience we pick $(x, y) = (0, 0)$. The left-hand side of Eq. (6.1) is then

$$\begin{aligned} E\left\{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Q(u, v) S[Q(u, v)(u^2 + v^2)] du dv\right\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E\{Q(u, v) S[Q(u, v)(u^2 + v^2)]\} du dv \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left\{ \sum_{n=0}^{\infty} n S[n(u^2 + v^2)] L^n \exp(-L) (1/n!) \right\} du dv \\ &= \sum_{n=1}^{\infty} [\exp(-L) (L^n/n!)] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} n S[n(u^2 + v^2)] du dv \\ &= \sum_{n=1}^{\infty} [\exp(-L) (L^n/n!)] V_s \\ &= [1 - \exp(-L)] V_s \end{aligned}$$

which is the right-hand side of Eq. (6.1). ■

Since $1 - \exp(-L) = 1 - P(Q = 0)$, an alternative statement of theorem 3 is that the expected response value for a uniform field with $q(u, v) \equiv L$ is the volume constant V_s times $1 - P(Q = 0)$, i.e., times the probability that the quantum catch is not zero. It is interesting to note (from the form of the proof) that when theorem 3 is expressed that way, its validity does not depend on the assumption that Q is a Poisson random variable.

Why is the expected response to noisy uniform fields not constant, like the response to deterministic uniform fields? An intuitive answer runs as follows. If the input function $I(u, v) \equiv L$ is permitted to assume values between 0 and 1, the area of the point-spread function $LS(Lr^2)$ can become infinitely large as L becomes small. As L decreases, and with it the size of the point-spread contribution from any given receptor to the output value at a given point, that decrease is exactly compensated for by the increased number of receptors whose point spreads can reach that point. However, for photon-noisy images the quantum catch can only be 0, 1, 2, ..., and so the maximum area point spread is produced by $Q = 1$. Thus, as $E\{Q\}$ decreases to low levels, the number of receptors whose point spreads can reach any given output point does not increase indefinitely but instead reaches an upper limited determined by the area of $S(r^2)$. At the same time, the expected number of receptors that actually contribute any point spread at all decreases (since $P(Q = 0)$ is rising), and so the expected size of the total point spread arriving at any point decreases.

Theorem 3 holds for any spread function S . I have not been able to obtain a comparably general result for the variance of the uniform-field response. However, for Gaussian CV operators, the variance turns out to have a simple expression.

Theorem 4

For the Gaussian CV operator G_σ , the expected output value for a noisy, uniform-field input image Q with $q(u, v) \equiv L$ is

$$E\{G_\sigma[Q](x, y)\} = 1 - \exp(-L), \quad (6.2)$$

Fig. 7. Mean [top graph; Eq. (6.2)] and standard deviation [bottom graph; square root of Eq. (6.3)] of the response of G_{100} to Poisson noisy uniform fields.

and the variance is

$$\text{Var}\{G_\sigma[Q](x, y)\} = (1/8\pi\sigma^2)[1 - \exp(-2L)]. \quad (6.3)$$

In Fig. 7 the mean and the standard deviation are plotted as functions of L for the case $\sigma = 100$. The standard deviation reaches its asymptotic value of 0.002 at around $L = 5$.

Proof

Equation (6.2) follows immediately from Eq. (6.1), since V_s for G_σ is 1. To demonstrate Eq. (6.3) takes a bit more work. Clearly, the choice of the output point (x, y) is arbitrary, and so for convenience we pick $(x, y) = (0, 0)$. We then appeal to the fact, established earlier in the proof of theorem 2, that for any input image $Q(u, v)$ with expected value $q(u, v)$,

$$\text{Var}\{G_\sigma[Q](0, 0)\} = (1/\sigma^2)[A(1, q') - B(1, q')], \quad (6.4)$$

where $A(1, q')$ is given by Eq. (5.13) and $B(1, q')$ is given by Eq. (5.14). In those expressions, q' was an arbitrary input image derived from q : $q'(u, v) = q(\sigma u, \sigma v)$. Here, $q(u, v)$ is the constant L , and so $q'(u, v) \equiv L$; we need to evaluate the integrals $A(1, L)$ and $B(1, L)$. The first takes the form

$$\begin{aligned} A(1, L) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (1/2\pi)^2 L \exp[-L - r^2 + L \exp(-r^2)] du dv \\ &+ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (1/2\pi)^2 L^2 \exp[-L - 2r^2 \\ &+ L \exp(-r^2)] du dv, \end{aligned} \quad (6.5)$$

where $r^2 = u^2 + v^2$. Let $A1$ denote the first integral in Eq. (6.5), and let $A2$ denote the second. When changed to polar coordinates, $A1$ becomes

$$A1 = (1/2\pi) \int_0^{\infty} L \exp(-L) \exp(-r^2) \exp[L \exp(-r^2)] r dr,$$

and, after the change of variable $w = \exp(-r^2)$, this becomes

$$\begin{aligned} A1 &= (1/4\pi) \int_0^1 L \exp(-L) \exp(Lw) dw \\ &= (1/4\pi)[1 - \exp(-L)]. \end{aligned} \quad (6.6)$$

For $A2$, we have, after changing to polar coordinates,

$$A2 = (1/2\pi) \int_0^{\infty} L^2 \exp(-L) \exp(-2r^2) \exp[L \exp(-r^2)] r dr.$$

After the change of variable $w = r^2$, this becomes

$$A2 = (1/4\pi) \int_0^{\infty} L^2 \exp(-L) \exp(-2w) \exp[L \exp(-w)] dw,$$

and, after a second change of variable $z = L \exp(-w)$, we have

$$\begin{aligned} A2 &= (1/4\pi) \exp(-L) \int_0^L z \exp(-z) dz \\ &= (1/4\pi) \exp(-L)[1 + (L - 1)\exp(L)]. \end{aligned} \quad (6.7)$$

By adding Eqs. (6.6) and (6.7), we obtain

$$A(1, L) = L/4\pi. \quad (6.8)$$

For $B(1, L)$, we begin with

$$\begin{aligned} B(1, L) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (1/2\pi)^2 L^2 \exp(-2L) \exp(-r^2) \\ &\times \exp[2L \exp(-r^2/2)] du dv. \end{aligned} \quad (6.9)$$

By changing this to polar coordinates, we obtain

$$\begin{aligned} B(1, L) &= (1/2\pi) \int_0^{\infty} L^2 \exp(-2L) \exp(-r^2) \\ &\times \exp[2L \exp(-r^2/2)] r dr, \end{aligned}$$

and, after the change of variable $w = r^2/2$, this becomes

$$\begin{aligned} B(1, L) &= (1/2\pi) \int_0^{\infty} L^2 \exp(-2L) \exp(-2w) \\ &\times \exp[2L \exp(-w)] dw. \end{aligned}$$

After making a second change of variable $z = 2L \exp(-w)$, we obtain

$$\begin{aligned} B(1, L) &= (1/8\pi) \exp(-2L) \int_0^{2L} z \exp(z) dz \\ &= (1/8\pi)[2L - 1 + \exp(-2L)]. \end{aligned} \quad (6.10)$$

Finally, by subtracting Eq. (6.10) from Eq. (6.8) and applying Eq. (6.4), we arrive at

$$\text{Var}\{G_\sigma[Q](0, 0)\} = (1/\sigma^2)(1/8\pi)[1 - \exp(-2L)];$$

So Eq. (6.3) is proved. ■

7. EDGE RESPONSES

Theorem 5

Suppose that $Q(u, v)$ is a noisy input image whose mean $q(u, v)$ is the step function $q(u, v) = L$ for $u \leq 0$ and $q(u, v) = L + D$ for $u > 0$, with $D \geq 0$ (in other words, an infinitely extended vertical edge). The expected output image for the Gaussian CV operator G_σ is

$$\begin{aligned} E\{G_\sigma[Q](x, y)\} &= 1 - e^{-L} + \sum_{n=1}^{\infty} (1/n!) N[(x/\sigma)n^{1/2}] \\ &\times [(L + D)^n \exp(-L - D) - L^n e^{-L}], \end{aligned} \quad (7.1)$$

where N is the normal integral Eq. (2.1).

In Fig. 8 plots of Eq. (7.1) are compared with the deterministic edge response [Eq. (4.2)]. Because CV operators are translation and rotation invariant, Eq. (7.1), appropriately rotated and shifted, describes the expected response to edges of all orientations at all locations.

Proof

For this input image,

$$\begin{aligned}
 E\{G_\sigma[Q](x, y)\} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{n=1}^{\infty} \{(n/2\pi\sigma^2) \exp\{(-n/2\sigma^2)[(x-u)^2 + (y-v)^2]\} L^n (1/n!) e^{-L}\} du dv + \int_{-\infty}^{\infty} \int_0^{\infty} \sum_{n=1}^{\infty} (n/2\pi\sigma^2) \\
 &\quad \times \exp\{(-n/2\sigma^2)[(x-u)^2 + (y-v)^2]\} [(L+D)^n/n!] \exp(-L-D) du dv \\
 &= \int_{-\infty}^0 \sum_{n=1}^{\infty} \{[1/\sigma(2\pi)^{1/2}] n^{1/2} \exp\{(-n/2\sigma^2)(x-u)^2\} L^n (1/n!) e^{-L}\} du + \int_0^{\infty} \sum_{n=1}^{\infty} \{[1/\sigma(2\pi)^{1/2}] n^{1/2} \\
 &\quad \times \exp\{(-n/2\sigma^2)(x-u)^2\} (L+D)^n/n! \exp(-L-D)\} du \\
 &= \sum_{n=1}^{\infty} \left\{ L^n (1/n!) e^{-L} \int_{-\infty}^0 [1/\sigma(2\pi)^{1/2}] n^{1/2} \exp\{(-n/2\sigma^2)(x-u)^2\} du \right\} + \sum_{n=1}^{\infty} \left\{ (L+D)^n (1/n!) \right. \\
 &\quad \times \exp(-L-D) \int_0^{\infty} [1/\sigma(2\pi)^{1/2}] n^{1/2} \exp\{(-n/2\sigma^2)(x-u)^2\} du \left. \right\} \\
 &= \sum_{n=1}^{\infty} (L^n/n!) e^{-L} \{1 - N[(x/\sigma)n^{1/2}]\} + \sum_{n=1}^{\infty} [(L+D)^n/n!] \exp(-L-D) N[(x/\sigma)n^{1/2}] \\
 &= 1 - e^{-L} + \sum_{n=1}^{\infty} (1/n!) N[-(x/\sigma)n^{1/2}] [(L+D)^n \exp(-L-D) - L^n e^{-L}],
 \end{aligned}$$

which is Eq. (7.1). ■

As Fig. 8 shows, the expected edge response [Eq. (7.1)] becomes indistinguishable from the deterministic edge response [Eq. (4.2)] for $L \geq 10$. In that range it contains symmetrical Mach bands, and the peak and trough values of those Mach bands are independent of both the scale parameter σ and the mean illuminance L : they are determined entirely by the Weber fraction (i.e., the edge contrast) D/L . Since we know from theorem 4 [Eq. (6.3)] that the standard deviation of the uniform-field response of G_σ varies inversely with σ , this means that, for any given contrast D/L , the signal-to-noise ratio of the peak edge response can be made arbitrarily large by increasing σ . In other words, any edge (and consequently any target bounded by sharp edges) can be made as detectable as we like. However, there is a cost for this increase in edge detectability, because, although the size of the peak response is independent of σ , its location is not: the peak occurs at $x = x_{\max} = \sigma[(1/D)\ln(1 + D/L)]^{1/2}$, and the width of the entire Mach band grows as σ . This is indicative of the fact (discussed in the next section) that the spatial resolution of G_σ varies inversely with σ , with the result that, in the range $L \geq 10$, the product of the visual acuity times the edge-response signal-to-noise ratio is a constant.

Figure 8 shows that when L falls below 10 quanta/receptor, the expected edge response departs from the deterministic response: first it becomes less than the deterministic response, and its Mach bands become attenuated; then at L values < 1 the Mach bands entirely disappear. This behavior can be understood analytically from Eq. (7.1): if both L and $L + D$ are small enough that their squares can be treated as zero, Eq. (7.1) becomes approximately

$$E\{G_\sigma[Q](x, y)\} = L + DN[x/\sigma], \quad (7.2)$$

which increases monotonically with x . If G were a linear operator, the Mach bands that it creates above 10 quanta/receptor would imply that it is a bandpass (or high-pass) filter, and the loss of those Mach bands below 10 quanta/receptor would imply that it is a low-pass filter.

I have not been able to derive an exact expression for the variance of the edge response of G_σ . Theorem 2, of course, shows that for any fixed L and D , the edge response varies inversely with σ^2 , but the upper bound given by relation (5.8) in that theorem is too large to be practically useful. Consequently, to gain an idea of the actual variance, we resort to simulation. Figures 9 and 10 show the results of applying the discrete version of G_σ , with $\sigma = 100$, to simulated photon-noisy images of a vertical edge with 50% contrast. The input images here were 600×600 pixel arrays in which the input value at each pixel was a Poisson random variable with a mean of L for points to the left of the edge (which ran lengthwise down the center of the image) and a mean of $L + D$ for points to the right of it, with $D = 0.5L$ in every case. L ranged from 0.1 to 10,000. The left-hand panels in Figs. 9 and 10 show the quantum catch per pixel across two horizontal scan lines through each input image. (In the figures, only the central 256 pixels are shown for each scan line.) The discrete version of G_{100} was applied to each image, and the central 256×256 pixels of the resulting output images were treated as an approximation to the output images that would have been created by infinitely extended input images. (For input images of finite size, the output pixels near the border do not receive the full point spread that they would get from an input image of infinite size, which is the case envisaged by the analysis. But since the maximum size of the point-spread function (i.e., its area for an input of one photon) can be calculated, a set of output pixels sufficiently far from the border can be identified as having the same output image for a finite-sized input image as it would have for one of infinite size. Hence the 600×600 pixel input image yields a usable 256×256 pixel output image.) The right-hand panels in Figs. 9 and 10 show the output image values across the same two horizontal scan lines displayed for the input images.

From this simulation (which was carried out by Reuman) the variance of the edge response can be estimated for each point along the response profile parallel to the x axis. This variance proves to be the same across all points: its value is the uniform-field variance corresponding to the background

Fig. 8. Expected edge responses of G_{100} applied to Poisson noisy inputs [dotted lines; Eq. (7.1)] compared with its responses to the corresponding deterministic edges [solid lines, Eq. (4.2)]. The edge is a step from L to $L + D$ (mean quanta per receptor) with $D/L = 5$. For $L \geq 10$ the expected response to the noisy-input image becomes indistinguishable from the response to the deterministic input image.

illuminance L , i.e., the variance given by Eq. (6.3). For $L \geq 10$ and $\sigma = 100$, Eq. (6.3) implies a standard deviation of 0.002, and the peak height of the Mach band in the expected edge response for 50% contrast is 1.048, and so in that range the peak value of the response signal-to-noise ratio (i.e., the peak expected edge response minus the expected uniform-field response, all divided by the uniform-field standard deviation) is 24. When L is small (≤ 0.1), so that Eq. (7.2) describes the expected edge response and there are no Mach

bands, the response simply increases monotonically from a low-side asymptote of L to a high-side asymptote of $L + D$. In the latter region the signal-to-noise ratio at each output point is $2\sigma D(\pi/L)^{1/2}$ or, if the contrast D/L is denoted by C , $2\sigma C(\pi L)^{1/2}$. Thus, in the low-light range, the operator's output signal-to-noise ratio is not independent of the illuminance level, as it is for $L \geq 10$, but instead varies as $L^{1/2}$, as does its input. However, the output signal-to-noise ratio is much larger than the input ratio: at $L = 0.1$, where the input signal-to-noise ratio for a 50% contrast edge is 0.16, the signal-to-noise ratio for G_{100} is 56. The effect of this improvement can be appreciated by comparing the input and output image profiles for the case $L = 0.1$ in Fig. 9.

The bottom right-hand panel in Fig. 10 shows how saturation begins to occur when the illuminance level reaches 10^4 quanta/receptor. For $L = 10$ through $L = 1000$ the peak value of the edge response (i.e., the peak of the Mach band) remains essentially constant at 1.05, which is the value implied by the deterministic equation [Eq. (4.3)] for $\sigma = 100$ and $D/L = 0.5$. However, when L reaches 10,000, the peak value falls to 1.03. This still represents a signal-to-noise ratio of 15, and so the edge remains easily detectable, as the figure shows. However, by the time L reaches 10^5 , saturation is complete: the output image becomes simply a uniform field, with no modulation to indicate that the input image contained any edge at all. (Figure 10 does not show this case.) In other words, the upper end of the effective working range of G_{100} is around 10^4 mean quanta/receptor, or 10^4 Td, assuming a temporal-summation interval of 250 msec.

8. RESPONSE TO SINUSOIDAL GRATINGS

Theorem 7

Suppose that the expected input image $q(u, v)$ is a sinusoidal grating of the form $q(u, v) = q(u) = L(1 + m \cos 2\pi fu)$, with $m^2 \leq 1$. The expected output image for the Gaussian CV operator G_c is then

$$E\{G_c[Q](x, y)\} = \sum_{n=1}^{\infty} \left\{ (L^n/n!) e^{-L} \int_{-\infty}^{\infty} [n^{1/2}/\sigma(2\pi)^{1/2}] [1 + c(u)]^n \right. \\ \left. \times \exp[-(n/2\sigma^2)(x - u)^2] \exp[-c(u)L] du \right\}, \quad (8.1)$$

where $c(u) = m \cos 2\pi fu$.

Proof

$$E\{G_c[Q](x, y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{n=1}^{\infty} \left\{ (n/2\pi\sigma^2) \exp[-(n/2\sigma^2)((x - u)^2 \right. \\ \left. + (y - v)^2)] [q(u)]^n \exp[-q(u)] (1/n!) \right\} du dv \\ = \sum_{n=1}^{\infty} \left\{ (1/n!) \int_{-\infty}^{\infty} n^{1/2} [1/\sigma(2\pi)^{1/2}] [q(u)]^n \right. \\ \left. \times \exp[-q(u)] \exp[-(n/2\sigma^2)(x - u)^2] du \right\},$$

which is Eq. (8.1) when $q(u) = L(1 + m \cos 2\pi fu)$.

Fig. 9. Response profiles for G_{100} applied to Poisson noisy edges at different illuminance levels. The input images were noisy vertical edges in which the mean quanta per pixel were L to the left of 0 and $L + D$ to the right. $D/L = 0.5$ in all images. The left-hand panels show the actual quantum catch per pixel across two horizontal scan lines. The solid line in each left-hand panel marks the mean input intensity level (i.e., $L + D/2$). In the top two panels the ordinate scale is the actual quantum catch: the scale for the upper scan line is shown on the left; the scale for the lower scan line is shown on the right. In the bottom left-hand panel the ordinate scale is 20% of the mean value; i.e., tick marks indicate the mean value $L + D/2$ plus (or minus) 20%, 40%, etc. The right-hand panels show the output image values across the same two horizontal scan lines displayed for the input images. Solid horizontal lines indicate the expected response value at the edge itself [i.e., the value of Eq. (7.1) at $x = 0$]. The scale for the upper scan line is marked on the left-hand ordinate; the scale for the bottom scan line is marked on the right-hand ordinate. In the bottom panels, tick marks on the ordinate indicate increments of 0.025 above or below 1.0. Note that these curves are actual response profiles, not expected values.

Fig. 10. Continuation of Fig. 9 at higher illuminance levels.

Unless the mean illuminance L is quite small (≤ 0.1), it seems impossible to simplify Eq. (8.1) any further; to see what it implies, all that we can do is to plot it. For $L \geq 10$, of course, the expectation [Eq. (8.1)] is closely approximated by the response of G_σ to a deterministic sinusoid:

$$G_\sigma[q(u)](x) = \int_{-\infty}^{\infty} [1/\sigma(2\pi)^{1/2}] [L(1 + m \cos 2\pi fu)]^{1/2} \times \exp[-(L/2\sigma^2)(1 + m \cos 2\pi fu)(x - u)^2] du. \quad (8.2)$$

However, Eq. (8.2) is hardly more enlightening than Eq. (8.1). Computation shows that when the input contrast is 0.3 or more, Eq. (8.2) is distinctly nonsinusoidal: it seems to contain a spurious frequency component at $2f$. (Figure 3 shows this for $m = 0.9$.) However, it was shown analytically in Ref. 1 (and illustrated in Fig. 3 above) that for input contrasts of the order of $m \leq 0.1$, Eq. (8.2) is well approximated by a sinusoidal grating of the form of Eq. (4.4). Thus for $L \geq 10$ and $m \leq 0.1$ we have, to a close approximation,

$$E\{G_\sigma[Q](x, y)\} = 1 + \{(2\pi^2 f^2 \sigma^2 / L) \exp[-(2\pi^2 \sigma^2 f^2 / L)]\} \\ \times (m \cos 2\pi f x). \quad (8.3)$$

In this range, then, it is possible to define the MTF of G_σ as the ratio of the output contrast in Eq. (8.3) to the input contrast m . By letting $g(\sigma, f, L)$ denote this MTF, we obtain, for $L \geq 10$,

$$g(\sigma, f, L) = (2\pi^2 \sigma^2 f^2 / L) \exp[-(2\pi^2 \sigma^2 f^2 / L)]. \quad (8.4)$$

If spatial resolution is defined operationally by the highest spatial frequency f at which the MTF exceeds some fixed threshold value, then Eq. (8.4) shows that, in the range $L \geq 10$, resolution will be directly proportional to $L^{1/2}$ and inversely proportional to σ . Combined with the fact that the edge-response signal-to-noise ratio varies directly with σ , this last relationship justifies the remark (in Section 7 above) that G creates an exact trade-off between spatial resolution and contrast detectability.

The MTF $g(\sigma, f, L)$ can be extended to values below $L = 10$ [although not in the analytic form of Eq. (8.4)] because direct evaluation of Eq. (8.1) by numerical integration shows that for $m \leq 0.1$ the expected response to sinusoidal inputs is always sinusoidal to a close approximation. Consequently, the MTF can be computed for all values of L . Figure 11 shows the MTF's of G_σ , with $\sigma = 100$, for L values ranging from 0.01 to 1000 mean quanta/receptor. The frequency scale here is in cycles per receptor, and for $L = 100$ the peak frequency of the MTF is $f = 0.023$ cycle/receptor. The diameter of human photoreceptors is about 0.5 min of visual angle, and so this peak frequency would be 2.8 cycles/deg. Human spatial CSF's for mean illuminances of the order of 100 Td (i.e., a mean quantum catch around 100 per receptor per 0.25 sec) peak at about 3 cycles/deg,⁴ and so the round-number value $\sigma = 100$ provides a reasonable fit to that aspect of visual performance, which is the reason that this σ was selected for special attention. The use of the MTF peak frequency to estimate σ is motivated by the fact (shown below) that this is also the peak of the CSF at high light levels, and that value does not depend on the d' that is assumed to correspond to threshold. The choice $\sigma = 100$ implies a peak frequency of 6 cycles/deg at $L = 500$, which provides a good match to the peak frequency measured by Campbell and Green¹⁴ at a mean illuminance of 500 Td. Campbell and Green found that the peak frequency remains at around 6 cycles/deg when the optics of the eye are bypassed by interferometry, and so that value presumably reflects the purely neural mechanisms that we seek to model.

As Fig. 11 shows, the MTF $g(\sigma, f, L)$ undergoes a metamorphosis as L falls below 10 quanta/receptor: its shape changes from the bandpass form given by Eq. (8.4) to a low-pass form. Analysis of Eq. (8.1) for low values of L ($L \leq 0.1$) shows that, as L approaches zero, the MTF $g(\sigma, f, L)$ ap-

proaches the form $\exp(-2\pi^2 \sigma^2 f^2)$, which is the MTF of a linear operator with the Gaussian impulse response $(1/2\pi\sigma^2) \exp(-r^2/2\sigma^2)$. In fact, for $L \leq 0.1$ we can obtain a simple approximation to the exact form of the expected response profile [Eq. (8.1)], i.e., one that holds for all values of m . For that purpose we note that when $L \leq 0.1$, the Poisson probability factor $L^n(1/n!)e^{-L}$ in Eq. (8.1) is negligibly small for $n > 1$, and the factor $\exp[-mL(\cos 2\pi f x)]$ can be approximated by $[1 - mL(\cos 2\pi f x)]$. Consequently, for $L \leq 0.1$, we have approximately

$$E\{G_\sigma[Q](x, y)\} = Le^{-L} \int_{-\infty}^{\infty} [1 + m(1 - L)(\cos 2\pi f u) \\ - m^2 L (\cos 2\pi f u)^2] [1/\sigma(2\pi)^{1/2}] \\ \times \exp[-(x - u)^2/2\sigma^2] du \\ = Le^{-L} \{1 - m^2 L/2 + m(1 - L) \exp[-2(\pi\sigma f)^2] \\ \times \cos 2\pi f x - m^2 (L/2) \exp[-2(\pi\sigma f)^2] \\ \times \cos 2\pi(2f)x\}.$$

For $L \leq 0.1$, $Le^{-L} \approx L - L^2$. All terms in the above equation that contain L^2 are $\leq L^2$ and therefore ≤ 0.01 , and when they are all set equal to zero, the equation becomes simply

$$E\{G_\sigma[Q](x, y)\} = L[1 + m \exp[-2(\pi\sigma f)^2] \cos 2\pi f x], \quad (8.5)$$

which holds for all m . This shows that for $L \leq 0.1$, the CV operator G_σ becomes equivalent to the low-pass linear filter whose impulse response is the Gaussian function $(1/2\pi\sigma^2) \exp(-r^2/2\sigma^2)$.

Intuitively, this change occurs because an expected quantum catch of 0.1 or less means that almost all receptors catch at most one quantum. Consequently the only point-spread function that has a chance to express itself is the one corresponding to $Q = 1$; all the rest occur with negligible probability. The effect is thus equivalent to applying the linear operator whose impulse response is the CV point-spread function for $Q = 1$. This is obviously a general property of CV operators—at light levels for which the mean quantum catch per receptor is uniformly ≤ 0.1 , the CV operator with the spread function S becomes equivalent to the linear operator whose impulse response is $S(x^2 + y^2)$.

Fig. 11. MTF's of G_{100} for Poisson noisy-input images at different mean illuminance levels.

The MTF's in Fig. 11 indicate how much the expected contrast of any input grating $L(1 + m \cos 2\pi fu)$ will be attenuated by G_σ (for $m \leq 0.1$), and theorem 4 gives the mean and the standard deviation of the uniform-field response at any point. Consequently, for any point in the output image, we can calculate d' as the difference between the expected value of the output at that point for the grating versus a uniform field of the same illuminance, divided by the standard deviation of the uniform-field response. In particular, at any peak (say, at $x = 0$), d' will be

$$d'_{\max} = \{2\sigma m(2\pi)^{1/2}[1 - \exp(-L)]g(\sigma, f, L)\}/[1 - \exp(-2L)]^{1/2}. \quad (8.6)$$

If it is assumed that the grating is discriminable from a uniform field when $d'_{\max}$ reaches some threshold value, Eq. (8.4) can be used to determine the input contrast m needed to achieve that value for any σ, f , and L . A curve plotting the inverse of the contrast threshold against f , for fixed σ and L , is a CSF. Figure 12 shows the CSF's of G_σ , for $\sigma = 100$, based on the assumption that the threshold is achieved when $d'_{\max} = 2^{1/2}$. (Ignoring probability summation, i.e., assuming that the observer uses only the response at a single output point, that d' corresponds to a hit probability of 0.75 for a false-alarm probability of 0.25 in a yes-no signal-detection experiment or to a correct-response probability of 0.85 in an unbiased two-alternative forced-choice experiment).

Figure 12 shows that for mean illuminance $L \geq 10$ quanta/receptor, the contrast sensitivity at the best frequency is constant, but the best frequency itself (and the location of the entire CSF) varies proportionally as $L^{1/2}$. As L falls below 10, the shape of the CSF changes from bandpass to low pass, and its level decreases; the overall contrast sensitivity decreases. Below $L = 0.1$, that decrease is proportional to $L^{1/2}$, as would be expected, since in that range G acts as a linear filter with a fixed impulse response, and consequently its contrast sensitivity is directly governed by Poisson statistics.

The CSF's in Fig. 12 do not extend down to a sensitivity of 1 (threshold contrast 100%) because the nonlinearity of G can become appreciable for inputs whose contrasts are more than about 30%. Consequently the acuity G_{100} cannot be read directly from its CSF's (except for illuminances below 0.1 mean quantum/receptor, where G becomes effectively linear). To find the highest detectable frequency, it is necessary to evaluate Eq. (8.1) to find the peak response value and then determine the highest-frequency f that yields a threshold value of d' at $m = 1$. However, in practice it is much more convenient to base the acuity calculation on an input contrast of $m = 0.9$, because in this case the expected input image $q(u) = L(1 + m \cos 2\pi fu)$ will be uniformly ≥ 10 for all $L \geq 100$, and so in that range Eq. (8.1) can be replaced by the computationally simpler form of Eq. (8.2). Because Eq. (8.2) (as do all responses of G to deterministic images) obeys the scaling theorem [Eq. (4.1)], acuity for $L \geq 100$ needs to be computed exactly only for the case $L = 100$; all other acuities $A(L)$ in this range can then be calculated immediately from the scaling theorem relationship $A(L)/A(100) = L^{1/2}/100^{1/2}$ (which holds regardless of nonlinearities). In the range $L \leq 0.1$, $A(L)$ can be calculated from the linear approximation [Eq. (8.5)]. This leaves only the range $0.1 < L < 100$, where $A(L)$ must be calculated directly from Eq. (8.1).

Figure 13 shows visual acuity as a function of L for G_{100} , based on the assumption that the threshold corresponds to a d' of $2^{1/2}$ at the peak of the response to a 90% contrast sinusoidal grating. Acuity rises slowly at L values of < 0.1 quantum/receptor. In that range, where G is linear, the highest-frequency f at which a grating with a contrast m produces a given d' is determined by the relationship

$$\exp[2(\pi\sigma f)^2] = (m/d')(2\sigma\pi^{1/2}/L^{1/2}). \quad (8.7)$$

For $m = 0.9$, $d' = 2^{1/2}$, and $\sigma = 100$, Eq. (8.7) implies an acuity of 0.002 cycle/receptor, or 0.24 cycle/deg, at $L = 10^{-4}$ quantum/receptor. If it is assumed that 1 Td corresponds to a quantum catch of 4 quanta/sec per receptor and that the integration time of the visual system is 0.25 sec, then L can be identified with 10^{-4} Td. Human acuity at that retinal illuminance appears to be about 0.6 cycle/deg.⁶ At $L = 1000$, the acuity of G_{100} has risen to 26 cycles/deg (0.5 cycle/receptor), so the overall increase from -4 to $+3$ log Td is 2 log units. Human acuity at 1000 Td is 60 cycles/deg, and so it too increases by 2 log units from -4 to $+3$ log Td. (To reach 60 cycles/deg, G_{100} needs a mean illuminance of 5000 quanta/receptor.) Thus the overall growth of acuity with illum-

Fig. 12. CSF's of G_{100} for Poisson noisy sinusoidal gratings at different mean illuminance levels. Each curve shows the inverse of the input contrast that is needed to produce a constant peak value of d' (here, $d' = \sqrt{2}$) in the output image.

Fig. 13. G_{100} predictions of visual acuity versus retinal illuminance for Poisson noisy sinusoidal gratings. The curve shows the highest spatial frequency at which a 90% contrast grating can be discriminated from a uniform field with a d' value of $\sqrt{2}$. The left-hand ordinate shows acuity in cycles per receptor. The right-hand ordinate shows acuity in cycles per degree (cpd), assuming a receptor width of 0.5 min of visual angle.

inance is predicted fairly well by a CV operator. However, the form of the curve for $\log(\text{acuity})$ versus $\log(L)$ in Fig. 13 does not provide a good match to psychophysical results: it is concave upward (acuity grows faster at higher illuminances), whereas plots of comparable human data are always concave downward. In other words, human acuities are always significantly higher than those predicted by G_{100} throughout the range -4 to $+3.7 \log \text{Td}$. This discrepancy can be eliminated for any given value of L by adjusting the scale parameter σ , but the overall shape of the curve will remain concave upward. Evidently, no Gaussian CV operator will provide a good fit to psychophysical plots of acuity versus mean retinal illuminance.

This defect has two causes. One is the fact that at low light levels, CV operators become equivalent to low-pass linear filters, and a linear filter allows grating acuity to grow proportionally with the square root of mean retinal illuminance only if its MTF is proportional to $1/f$. [More precisely, if acuity grows as $L^{1/2}$ from f_1 to f_2 as L increases from L_1 to L_2 , then the MTF of the filter must be proportional to $1/f$ from f_1 to f_2 . The MTF of a Gaussian filter is $\exp(-2\pi^2\sigma^2 f^2)$, which decreases much faster than $1/f$.] At high light levels (about 10 quanta/receptor, or 10 Td, assuming a 250-msec time frame) the problem is the opposite; here, the Gaussian CV operator does cause acuity to grow proportionally with the square root of mean retinal illuminance, but this rapid growth occurs too late to match human performance; above 10 Td, human acuity grows slowly, if at all.

The failures of G_σ at high light levels are not confined to its acuity predictions; the latter are only symptomatic of a more general problem. This is the fact that, for illuminances above 10 quanta/receptor, the CSF's in Fig. 12 shift bodily along the log frequency axis as L increases, so that the best frequency varies as $L^{1/2}$, while the sensitivity at that frequency remains constant. Human CSF's do not shift bodily as retinal illuminance rises from 10 to 1000 Td; instead, they maintain a roughly constant best frequency over that range, and as illuminance rises, the sensitivity increases at that frequency and all higher ones.⁴ In addition, human CSF's show no absolute loss of low-frequency sensitivity as retinal illuminance rises, but because of the shift property, all Gaussian CV operators create such a loss above $L = 10$ quanta/receptor. Broadly speaking, one can say that in comparison with the visual system, Gaussian CV operators respond too drastically to changes in the mean light level above 10 quanta/receptor. The visual system evidently prefers to use extra quanta in that range to improve its high-frequency performance while maintaining a constant best frequency and a fixed passband, whereas G uses them to shift the entire passband toward higher frequencies, as though obsessed with improving visual acuity regardless of the cost to low-frequency performance.

Neither the MTF's for G_σ shown in Fig. 11 nor the CSF's shown in Fig. 12 take into account contrast reductions imposed on the retinal image by the optics of the eye. That factor can be incorporated into the model by multiplying the MTF of G_σ by that of the eye. For illuminances in the range 10^{-4} – 10^3 quanta/receptor, this makes essentially no difference. Figure 14 shows the MTF's of G_σ multiplied by an optical contrast reduction factor of the form $\exp(-2\pi^2 p^2 f^2)$, as though the eye has a Gaussian point-spread function of the form $(1/2\pi p^2)\exp(-0.5r^2/p^2)$. A p value of 1.2 provides a

Fig. 14. MTF's of G_{100} corrected for the optics of the eye. The thick curves are the uncorrected MTF's from Fig. 11. The thinner curves show the corrected MTF's. The correction is made with the assumption that the point-spread function of the eye is a Gaussian probability-density function with a standard deviation of 1.2 receptor diameters. Below 1000 mean quanta/receptor, the corrected and uncorrected MTF's are indistinguishable.

reasonably good fit to Gubisch's¹⁵ optical MTF for a 2.4-mm pupil; that value has been used to create the curves shown in Fig. 14. It can be seen that for mean illuminances of <1000 , incorporation of the optical MTF has a negligible effect on the overall MTF of G_σ and consequently cannot significantly affect its CSF predictions. In other words, those predictions still remain well off the mark after optical factors are taken into account.

It seems almost certain that these high-light defects are common to all CV operators, since they stem from the scaling theorem [Eq. (4.1)], and, in the deterministic case, that theorem holds for all spread functions. However, we have not shown that that theorem can be applied generally in the noisy-input case. To do so, it would be necessary to prove two results: that the variance of the uniform-field response is always asymptotically independent of the mean illuminance level [a general version of Eq. (6.3) in theorem 4] and that the expected output image for any photon-noisy input image always becomes asymptotically equal to the deterministic response to the corresponding expected input image (a general version of theorem 2). It seems to be virtually certain that both of these statements are true, but I have not found the key to proving either one. The brute-force methods used here to derive the properties of G do not carry over to the general case. This seems to be the outstanding open problem in the theory of CV operators: until it is solved, we cannot say definitely that the scaling theorem holds at high light levels regardless of the choice of the spread function or, consequently, that the major defects of the Gaussian model are common to all CV operators.

Visible Beats from Invisible Gratings

If two sinusoidal gratings with frequencies f_1 and f_2 are imaged on one's retina simultaneously, both with a contrast near 1.0 (using interferometry to bypass the optical transfer function), and f_1 and f_2 are above the resolution limit (so that each grating alone appears to be a uniform field), one sees a grating whose apparent frequency is the difference $f_1 - f_2$

and whose apparent contrast is the reverse of that of the input gratings (i.e., it has a trough instead of a peak at $x = 0$).⁷ Figure 15 shows that the operator G_σ can produce this effect. For $\sigma = 100$, the highest spatial frequencies passed by this operator are 0.07 cycle/receptor at $L = 100$ and 0.10 cycle at $L = 200$. The top part of Fig. 15 shows the expected response profiles for cosine-phase input gratings with $L = 100$ and a contrast of 0.9, with $f = 0.10$ in one case and $f = 0.11$ in the other. Neither grating by itself produces a spatially modulated response. The bottom part shows the expected response to an input image that is the sum of the 0.10 and 0.11 gratings. (The input image profile is shown at the

bottom of the figure). The response is essentially a reverse-contrast cosine whose frequency is the difference 0.01 and whose contrast is 7%. Since the output standard deviation here is 0.002, this contrast represents a signal-to-noise ratio of 35, and so the beat frequency should be quite visible.

9. RESPONSE TO SPOTS

A spot here means a region of some constant retinal illuminance $L + D$ surrounded by a uniform field of illuminance L , the sort of target used in increment threshold measurements. The simplest case analytically is a square spot, dealt

contrast = 0.9 $f = 0.11$ $L = 100$

contrast = 0.9 $f = 0.10$ $L = 100$

$L = 100$ $f_1 = 0.10$ $f_2 = 0.11$ contrast = 0.9

Fig. 15. G_{100} creates visible beats from invisible gratings. The upper two panels illustrate the expected input and output image profiles for 90% contrast sinusoidal gratings of mean illuminance 100 quanta/receptor and frequency 0.11 (left) and 0.10 (right) cycles/receptor. The bottom panel shows the expected input and output image profiles obtained when the input image is the sum of the two sinusoids. The output image is essentially a reverse-contrast cosine whose frequency is 0.01 cycle/receptor.

with below in theorem 8. That shape is sufficient to illustrate the essential properties of the response of G to all spots, because these response properties depend not on the shape of the test spot but on its size. For any spot there is some illuminance level beyond which the maximum response value occurs at the peak of the Mach bands generated by its edges. Since the amplitude of that peak depends only on the Weber fraction D/L , the detectability of any spot will obey Weber's law once L reaches the critical level. That level depends not on the exact shape of the spot but on the distance between its edges: the smaller that distance is, the higher L must be before the Mach bands from opposite sides shrink enough so that they no longer overlap, and Weber's law can begin to hold.

Theorem 8

Suppose that the expected input image $q(u, v)$ is a square spot described by $q(u, v) = L + D$, for $|u| \leq W/2$, and $|v| \leq W/2$, and $q(u, v) = L$ elsewhere (so that the spot width is W). The expected output image for the Gaussian operator G_σ is then

$$E\{G_\sigma[Q](x, y)\} = 1 - \exp(-L) + \sum_{n=1}^{\infty} (1/n!)(L + D)^n \times \exp(-L - D) - L^n \exp(-L) [(A - B)(C - D)], \quad (9.1)$$

where

$$\begin{aligned} A &= N\{(1/\sigma)n^{1/2}[(W/2) - x]\}, \\ B &= N\{(1/\sigma)n^{1/2}[-(W/2) - x]\}, \\ C &= N\{(1/\sigma)n^{1/2}[(W/2) - y]\}, \\ D &= N\{(1/\sigma)n^{1/2}[-(W/2) - y]\}, \end{aligned}$$

and N is the normal integral Eq. (2.1)

Proof

Let T denote the set of points $\{(u, v) | -W/2 \leq u, v \leq W/2\}$ (i.e., T is the region within the test spot), and let S denote its complement in the (u, v) plane. For this input image,

$$\begin{aligned} E\{G_\sigma[Q](x, y)\} &= \iint_T \left(\sum_{n=1}^{\infty} (n/2\pi\sigma^2) \exp\{(-n/2\sigma^2)[(x-u)^2 \right. \\ &\quad \left. + (y-v)^2]\} (L + D)^n (1/n!) \right. \\ &\quad \left. \times \exp(-L - D) \right) dudv + \iint_S \left[\sum_{n=1}^{\infty} (n/2\pi\sigma^2) \right. \\ &\quad \left. \times \exp\{(-n/2\sigma^2)[(x-u)^2 + (y-v)^2]\} \right. \\ &\quad \left. \times (L^n)(1/n!)e^{-L} \right] dudv. \quad (9.2) \end{aligned}$$

Let I_t denote the first integral in Eq. (9.2), and let I_s denote the second one. I_s would be the expected response to a uniform field with an intensity L if S were the entire (u, v) plane, and so it can be rewritten as that response [i.e., $1 - \exp(-L)$] minus the part of that response arising from integration over the complement of S , which is T . Consequently, Eq. (9.2) can be rewritten as

$$I_t + 1 - \exp(-L) - \iint_T (\text{the integrand of } I_s) dudv.$$

Interchanging of summation and integration and then evaluation of the resulting normal integrals leads immediately to Eq. (9.1). ■

Figure 16 shows profiles of Eq. (9.1) for G_{100} along the line $y = 0$ (i.e., horizontally through the center of the spot) for spots 1000 receptors in width, with $D/L = 5$ and $L = 10, 1$, and 0.1 quanta/receptor. The figure also shows profiles of the response of G_{100} to the corresponding deterministic input images $q(u, v)$, i.e., Eq. (4.6) with $y = 0$. At $L = 10$ (and all higher values) the expected response $E\{G[Q]\}$ is the same as $G[q]$, as we would expect from theorem 2. For lower L values, $E\{G[Q]\}$ falls below $G[q]$, and the Mach bands disap-

Fig. 16. Expected response profiles of G_{100} for Poisson noisy spot images compared with its response profiles for deterministic spots. The dotted lines show the expected output image values along the x axis [i.e., Eq. (9.1) with $y = 0$] for a Poisson noisy-input image whose mean intensity is $L + D$ quanta/receptor inside a square spot of width 1000 receptor diameters and L outside the spot. The solid lines show the output image values for a deterministic spot with the same parameters [i.e., Eq. (4.6) with $y = 0$]. For $L \geq 10$, the two curves become indistinguishable.

pear, again as expected from earlier results. Profiles across a circular spot would show the same effects.

Threshold versus Illuminance

From Eq. (9.1) [or, for $L \geq 10$, Eq. (4.6)] we can calculate the expected response for each point in the output image for any L , D , and W , and that value, combined with the mean and the standard deviation of the uniform-field response from Section 6, determines a d' value for each point. Assuming that the spot is detectable when the largest d' ($d'_{\max}$) reaches some critical value, we can compute the increment D required for detection of a spot of width W on any background L , i.e., the TVI curve. Figure 17 shows two TVI curves for G_{100} , one for a square spot whose width is 1000 receptors (filled points) and the other for a spot width of 100 receptors (open points). The figure also shows Aguilar and Stiles's TVI curve for a (circular) spot 9 deg in diameter detected only by rods.¹⁶ Assuming a receptor diameter of 0.5 min, that test spot had a diameter of 1080 receptors, and so its TVI curve should be compared with the filled points for G_{100} . (Aguilar and Stiles's data are replotted by using the assumption that 1 scotopic Td corresponds to a mean quantum catch of 1 photon/receptor. The best current estimates indicate that 1 Td produces an average of 4.23 isomerizations per sec per rod,¹⁰ and Aguilar and Stiles's test flash lasted 0.2 sec; thus the actual effective total quantum catch per receptor over the course of the flash would have been 0.85/Td.) To estimate the value of $d'_{\max}$ needed to match the TVI curve for G_{100} to Aguilar and Stiles's data, we use the fact that in the region of background intensities where their TVI curve obeys Weber's law, their Weber fraction (D/L) is 0.1. In the same region the maximum response of G_{100} occurs at the peak of the Mach band produced by the spot's edges, and that value [given by Eq. (4.3)] depends only on D/L . For $D/L = 0.1$, the peak value is 1.012, and the response to a uniform field at these L values has a mean of 1.0 and a standard deviation of 0.002. Thus a Weber fraction of 0.1 corresponds to a $d'_{\max}$ of 6, a conservative threshold value but not an unreasonable one for cautious observers using the method of adjustment. This d' was used to calculate the increment threshold for all values of L for both $W = 1000$ and $W = 100$.

In addition to increment thresholds, Fig. 17 also shows predicted absolute thresholds for G_{100} , i.e., the spot intensity D that is needed to produce a $d'_{\max}$ of 6 when $L = 0$. Nothing in the model itself limits detection when the background quantum catch is zero, but we can calculate how much dark light would be required to produce the absolute threshold that Aguilar and Stiles measured for their test spot. That amount proves to be $10^{-4.4}$ quantum/receptor, and the predicted absolute threshold points in Fig. 17 are based on that value. Of course, the prediction for $W = 1000$ is constrained to match Aguilar and Stiles's absolute threshold, and so there is no significance to the goodness of fit of this point. However, it is significant that the amount of dark light needed to match the psychophysical absolute threshold is small enough to have no effect on increment thresholds for background intensities greater than $L = 10^{-4}$. (At $L = 10^{-4}$ the dark light raises the threshold by 0.06 log unit.) All the predicted points in Fig. 17 were obtained with the assumption that the nominal L value is augmented by a dark light equivalent to $10^{-4.4}$ mean quantum/receptor. Except at $L =$

10^{-4} , these thresholds are graphically indistinguishable from those predicted without dark light.

If the absolute threshold points are ignored, Fig. 17 shows that the TVI curves for G_{100} consist essentially of three branches: a low-light branch ($L \leq 0.1$) in which the slope is 0.5 (detection obeys the deVries-Rose law), a high-light branch in which the slope is 1.0 (detection obeys Weber's law), and a transitional branch connecting the two extremes. All three branches are to be expected from any Gaussian CV operator with any scale parameter. In the low-light region, when L is less than 0.1, all CV operators become equivalent to linear low-pass filters, and the deVries-Rose law follows directly from linearity combined with Poisson statistics. At high light levels, Weber's law follows from the fact that the peak amplitudes of the Mach bands depend only on the Weber fraction and that the mean and the variance of the uniform-field response are constant; thus $d'_{\max}$ depends only on the Weber fraction. Figure 17 also shows that the background illuminance at which Weber's law begins to hold depends on the size of the test spot: for $W = 1000$ the critical illuminance is 10 quanta/receptor; for $W = 100$ it is 100 quanta/receptor. However, once its critical background level is reached, the TVI curve for the smaller spot superim-

Fig. 17. G_{100} curves for increment threshold versus background intensity for test spots of different widths. Open points (spot width, 1000 receptors) show the increment D required for a square test spot of intensity $L + D$ mean quanta/receptor, surrounded by a background of intensity L , to be detected with a peak d' value of 6. Solid points show the same for a spot width of 100 receptors. The smooth curve is Aguilar and Stiles's scotopic TVI curve for a 200-msec 9-deg circular test spot, plotted on the assumption that 1 scotopic Td equals a mean intensity of 1 quantum/receptor.

poses upon that of the larger one. This will be true for any size spot once the background illuminance reaches its Weber's-law range. (Of course, a small spot may never actually reach that range, because saturation can occur before it is achieved. For G_{100} the TVI curve for $W = 10$ becomes asymptotic only at $L = 10,000$ quanta, which is the level at which saturation effects begin to become significant, and so smaller spots will never obey Weber's law.)

In the intermediate range, between the deVries-Rose region and its Weberian asymptote, the behavior of the G TVI curve reflects a complicated interaction between two effects: the development of Mach bands (owing to the growing influence of the operator's intrinsic nonlinearity) and the narrowing of those Mach bands (owing to the shifting of the operator's illuminance-dependent MTF). One curious consequence of this interaction is that at $L = 100$, the smaller spot has a lower increment threshold than the larger one.

Comparing Aguilar and Stiles's TVI curve with the G_{100} curve for $W = 1000$, we see that while the overall fit is fairly good, there are two serious discrepancies. The human curve begins to follow Weber's law at background illuminances of around $L = 10^{-2}$, whereas G_{100} requires $L = 10$ and obeys the deVries-Rose law quite closely up to $L = 1$. Also, at the high end of the background range, Aguilar and Stiles's curve begins to exhibit saturation at $L = 100$, and saturation is complete by $L = 1000$. G_{100} also saturates (although the figure does not show it), but not until L exceeds 10,000. Neither of these discrepancies can be remedied by altering the scale parameter σ : no σ will cause Weber's law to hold below $L = 0.1$, because G_σ will always be effectively a linear operator for illuminances in that range; to produce saturation at $L = 1000$ requires a σ value on the order of 7, which would cause $d'_{\max}$ in the Weber's law range to be only 0.4 (i.e., for $D/L = 0.1$). In addition, it should be noted that the dark-light value of $10^{-4.4}$ mean quantum/receptor required to fit Aguilar and Stiles's absolute threshold is roughly 30 times smaller than the direct estimate made by Baylor *et al.* of the dark light in primate rods.¹⁰ Altogether, then, whereas the Gaussian CV operator can duplicate in a general way the main features of human scotopic TVI curves, it does not provide exact quantitative predictions. More generally, because all CV operators become effectively linear operators for $L \leq 0.1$ and consequently obey the deVries-Rose law, we can say that, regardless of its spread function S , no CV operator will cause Weber's law to hold for background illuminances below 0.1 mean quantum/receptor.

Ricco's Law

For a square test spot of width W and illuminance $L + D$ surrounded by a background of illuminance L , Ricco's law holds if the detectability of the spot remains constant when the product DW^2 is held constant. We can take that to mean that the peak value of d' across the spot response is constant for a constant value of the product DW^2 . Analysis of Eq. (9.1) shows that G_σ never implies Ricco's law exactly, but for spots smaller than a critical size, Ricco's law holds as a close approximation, so that it would appear to be valid. The size of this critical area, Ricco's area, shrinks as L rises, but not until L reaches a value of 0.1.

To determine the size of Ricco's area, consider first the case of illuminance levels in the range where G_σ is effectively linear, i.e., $L + D \leq 0.1$. Evaluating Eq. (9.1) with L^2 and $(L$

$+ D)^2$ set equal to zero yields a mean response at the spot center of

$$E\{G_\sigma[Q](0, 0)\} = L + D\{2N[W/(2\sigma)] - 1\}^2, \quad (9.3)$$

and this will be the largest mean response in the output image. Consequently, the maximum value of d' across the output image will be Eq. (9.3) minus L , divided by the square root of Eq. (6.3) (the standard deviation of the uniform-field response), and Ricco's law holds exactly if Eq. (9.3) remains constant for all D and W values such that the product DW^2 is constant. Thus the CV model implies that Ricco's law never holds exactly, since the cumulative normal distribution function N is never exactly linear. However, for z values in the range -0.5 to $+0.5$, $N(z) = 0.4z + 0.5$ to an accuracy of two decimal places. When the argument of N in Eq. (9.3) falls in that range, we have approximately

$$E\{G_\sigma[Q](0, 0)\} = L + DW^2(0.4/\sigma)^2 \quad (9.4)$$

and

$$d' = DW^2(0.4/\sigma)^2 / (2L/8\pi\sigma^2)^{1/2} = DW^2(0.57)/\sigma L^{1/2}. \quad (9.5)$$

Consequently, Ricco's law holds to an accuracy of two decimal places for $W \leq \sigma$. For G_{100} , then, the apparent width of Ricco's area at low illuminance levels will be 100 receptors. If a receptor width of 0.5 min of visual angle is assumed, the width of Ricco's area is 0.83 deg. Note that this width will remain constant from $L = 0$ (i.e., for a background consisting only of dark light) up to around $L = 0.1$; the size of Ricco's area remains constant throughout the range in which G is effectively linear. The prediction 0.83 deg is not bad: Barlow's⁵ psychophysical estimate of the width of Ricco's area (at an eccentricity of 6.5 deg) for $L = 0$ is 0.71 deg, and his data for backgrounds at around 0.01 Td suggest that the size of Ricco's area remains unchanged up to at least that level.

At high background levels ($L \geq 10$), where the expected response [Eq. (9.1)] becomes the same as the deterministic response [Eq. (4.6)] and the mean and the variance of the uniform field response are constants [the mean is 1 and the variance is given by Eq. (6.3)], d' at the center of the test spot is

$$\{2N[(W/2\sigma)(L + D)^{1/2}] - 1\}^2 - \{2N[(W/2\sigma)L^{1/2}] - 1\}^2 / (1/8\pi\sigma^2)^{1/2}, \quad (9.6)$$

and again, Ricco's law will hold to an accuracy of two decimal places when the arguments of the normal integral N fall in its linear range, i.e., when $W \leq \sigma/(L + D)^{1/2}$. In that range, Eq. (9.6) simplifies to

$$d' = 0.8DW^2/\sigma. \quad (9.7)$$

If it is assumed that $D/L = 0.1$, then the width of Ricco's area at high background levels equals $0.95\sigma/L^{1/2}$ receptor diameters. At $L = 1000$, that width would be 3 receptors, or 0.025 deg, which is about 7 times smaller than the smallest diameter of Ricco's area obtained by Barlow at his highest background levels.² In other words, the CV operator overpredicts the extent to which Ricco's area shrinks as the background illuminance rises from 0 to 1000 Td: it implies that the diameter will decrease by a factor of about 30, whereas the observed decrease is around a factor of 6. On the other hand, from 0 to 100 Td the predicted linear shrinkage factor is only 10.5, which is not a bad prediction.

Quantum Efficiency

For the task of detecting a square spot of width W and illuminance $L + D$ mean quanta/receptor surrounded by a background of illuminance L , an ideal observer using the entire quantum catch within the spot has a d' of $DW/L^{1/2}$. Consequently, to achieve a fixed value of d' for any background level and spot size, an ideal quantum-limited detector requires a mean total increment $DW^2 = d'WL^{1/2}$. The quantum efficiency (QE) of any other detector is the ratio between that minimum and the mean total increment required by that detector to achieve the same d' . Here, we consider the QE of a spot detector based on a Gaussian CV operator: first the photon-noisy input image is filtered by G_σ , and then the detection decision is based on the value of the output image at its center point, i.e., $(x, y) = (0, 0)$. (For small spots that point always has the highest d' in the output image.) This is the same detection mechanism that was assumed above to derive the TVI curves in Fig. 17. It is undoubtedly not the best possible CV observer: the ideal observer would make use of more than one point in the output image. However, it seems practically impossible to calculate the joint statistical properties of the entire ensemble of random variables that compose the full output image, and so the ideal CV observer cannot readily be determined. The less-than-ideal CV observer considered here has the advantage of being mathematically tractable and permits us to calculate at least a sensible lower bound for the QE of a detection mechanism whose input is a photon-noisy image filtered by a CV operator.

We consider here only spots smaller than Ricco's area, as determined in the preceding section. For these spots the peak d' value is given by Eq. (9.5) for L values in the range 0–0.1 and by Eq. (9.7) for $L \geq 10$. For any fixed background illuminance $L \leq 0.1$ and any spot width $W \leq \sigma$ (to ensure that Ricco's law holds), d' for the CV detector equals d' for the ideal observer when their respective illuminance increments D_{cv} and D_{io} satisfy the relationship

$$D_{io}/D_{cv} = 0.57W/\sigma = \text{QE}. \quad (9.8)$$

In this range, the QE increases with the size of the test spot and reaches a peak value of 0.57 for $W = \sigma$, i.e., when the spot is exactly the size of Ricco's area. It should be noted that this QE value is based on effectively absorbed photons. Psychophysical estimates of the QE of human observers are usually based on quanta arriving at the cornea and, with that as a baseline, average around 0.05 when the task is spot detection at low background levels.¹⁷ The fraction of quanta at the cornea that actually result in photopigment isomerizations is estimated to be around 0.2,¹⁷ and so the QE of the neural visual system for effectively absorbed photons at low light levels is around 25%. In the low-light range, then, the QE of the Gaussian CV operator for optimal size targets is about twice that exhibited by the human visual system.

For high background illuminances, with $L \geq 10$ [and $W \leq \sigma/(L + D)^{1/2}$ to guarantee Ricco's law], d' for the CV observer matches d' for the ideal observer when

$$D_{io}/D_{cv} = 0.8WL^{1/2}/\sigma = \text{QE}. \quad (9.9)$$

Because $W \leq \sigma/(L + D)^{1/2} < \sigma/L^{1/2}$, the right-hand side of Eq. (9.9) cannot exceed 0.8. In this range, where the CV operator fully expresses its nonlinearity, it causes the QE for detecting a small fixed-size spot to grow as $L^{1/2}$ so long as $L <$

$(\sigma/W)^2$. For G_{100} , for example, the QE for a spot 10 receptors wide (5 min of visual angle) would be 0.25 at $L = 10$ and 0.8 at $L = 100$. The 80%-maximal QE value is much higher than QE values typically reported for human observers detecting spots on intense backgrounds, which are of the order of 1% for corneal quanta or 5% for isomerizing quanta.¹¹ Thus, at high light levels, the QE of a CV observer can be as much as 16 times higher than that of human observers. Of course, this maximum efficiency is only achieved for spots of the optimal size, those that exactly fill Ricco's area at any given background level.

10. OPEN PROBLEMS

For deterministic input images, Cornsweet and Yellott¹ showed that a fairly complete theory of CV operators could be developed by means of the scaling theorem [Eq. (4.1)], by which it can be shown that all CV operators share the same general properties, regardless of the exact form of the spread function. I have not found a similar key to obtaining general results in the case of photon-noisy images. The present analysis focused almost entirely on the Gaussian CV operator and relied on brute force methods that do not generalize to arbitrary spread functions. Even for the Gaussian case, there still remains one important question for which we have only simulation results [problem (3) below]. Altogether, then, the mathematical theory of CV operators for noisy images is left here in rough form: the Gaussian case permits us to guess how things must work in general, but it would be reassuring to substitute theorems for intuitions.

The following three problems seem to be especially important:

(1) Can theorem 1 be generalized to arbitrary spread functions? Theorem 1 shows that for the Gaussian operator G defined by Eq. (3.8), the expected output image $E\{G\{Q\}(x, y)\} \approx G\{E\{Q\}(x, y)\}$ when the expected input image $E\{Q(u, v)\}$ is uniformly ≥ 10 quanta/receptor. The problem is to determine whether this is true for all CV operators, i.e., all operators of the general form of Eq. (1.1). If it is, the entire deterministic theory could be applied generally in the photon-noisy case at high light levels, just as it has been here for the Gaussian operator.

(2) Can the variance result [Eq. (6.3)] of theorem 4 be generalized to arbitrary spread functions? Equation (6.3) showed that for the Gaussian operator, the variance of the uniform-field response is proportional to $1 - \exp(-2L)$, where L is the mean quantum catch per receptor. In other words, the variance becomes independent of the field intensity at high light levels. Is this true in general? [Equation (6.1) of theorem 3 shows that the mean of the uniform-field response is always proportional to $1 - \exp(-L)$, regardless of the spread function, and it seems likely that a similar result holds for the variance.] The practical significance of solving problems (1) and (2) was noted in Section 8: a yes answer to both questions would prove the conjecture that for all CV operators the CSF shifts bodily along the log frequency axis as retinal illuminance varies above 10 quanta/receptor, so that the peak spatial frequency and the visual acuity are both proportional to the square root of the mean illuminance level.

(3) Can it be shown that at high light levels, say, where

$E\{Q(u, v)\}$ is uniformly ≥ 10 , the variance is constant across all points in the output image, regardless of the exact form of the expected input image? (The natural value of this constant would be the asymptotic uniform-field variance.) The simulation results reported in Section 7 suggest that this is true for Gaussian operators, but we have no analytic proof for any CV operator, including the Gaussian one.

11. DISCUSSION: CONSTANT-VOLUME OPERATORS AS MODELS OF RETINAL IMAGE PROCESSING

The aim of this paper was to derive the basic consequences of applying CV operators to photon-noisy images and to use those results to determine how well a simple model of retinal image processing based on CV operators can predict the psychophysical properties of human spatial vision. In the model studied here it was assumed that the quantum catch of the photoreceptors is accumulated over a fixed time period (here, 250 msec, but the exact interval is not critical) and then filtered by a CV operator with a fixed-parameter Gaussian spread function [the operator G is given by Eq. (3.8), with $\sigma = 100$]. Test images were assumed to become discriminable from uniform fields when d' at any point in the output image exceeds a fixed threshold. This model proved to do a surprisingly good job of duplicating the global features of spatial vision over the range of retinal illuminances from the absolute threshold to 1000 Td: as illuminance increases over that range, it causes Ricco's area to shrink; visual acuity to rise; peak contrast sensitivity to grow to an asymptote on the order of 100 (threshold contrast $\approx 1\%$); the shape of the CSF to change from low pass to bandpass; and increment thresholds to shift from deVries-Rose-law behavior to Weber's law behavior, a change that occurs sooner for large targets than for small ones. The model also correctly predicts that two gratings whose frequencies are both higher than the resolution limit at a given mean luminance level (so that both resemble uniform fields when presented separately) can give rise to visible contrast at their difference frequency when viewed simultaneously.

However, this model generally fails to duplicate the exact quantitative parameters of spatial vision. In particular, at low light levels it underpredicts the growth of visual acuity with retinal illuminance, and at high levels it overpredicts it. In addition, it incorrectly implies that at illuminance levels above 10 Td the peak frequency of the spatial CSF should vary proportionally with the square root of the mean retinal illuminance and that low-frequency contrast sensitivity should decrease as retinal illuminance increases in that range.

Thus the simplest CV model succeeds in duplicating the overall qualitative properties of spatial vision up to at least 1000 Td but fails to make accurate quantitative predictions. Could these defects be cured by tinkering? The prospects for this seem rather different at high and low light levels. At low levels the basic problem is that the intrinsic nonlinearity of the CV operation has no opportunity to express itself. When the probability of a photoreceptor's catching more than one photon per time frame becomes negligible (< 0.1 Td) the only point spread that has a chance to occur is the one corresponding to a 1-quantum catch, and so a CV operator with a spread function S becomes effectively equivalent

to the linear operator whose impulse response is $S(r^2)$. Consequently, in this range CV operators share the principle defect of linear operators: no matter what the spread function is, increasing the mean illuminance level can only cause the CSF to translate rigidly upward (in a log-log plot) so that it can never match the lateral shifts seen in human CSF's.

The most obvious remedy for this low-light problem is to alter the assumption that the CV operation occurs at the level of individual photoreceptors. If we assumed instead that the quantum catches of some number of receptors are first pooled and then subjected to a CV operation, the operator's nonlinear effects could be made to reveal themselves at low light levels; acuity could be made to grow as the square root of retinal illuminance, and Weber's law could be made to hold. For example, if the CV operation were applied at the level of a second-stage unit that sums the quantum catch of 100 receptors, the effective nonlinear range of a CV model could be extended down to 0.01 Td. Moreover, the same change would cause the CV operator to begin to saturate at a level 100 times lower than before; e.g., G_{100} would begin to saturate at around 100 Td instead of at 10^4 Td, and so it could duplicate the saturation properties of the rod system.¹⁶ Altogether, it seems at least conceivable that a viable model of scotopic spatial vision could be constructed along these lines.

At high light levels the defects of CV models arise from the intrinsic properties of the operation itself: the fact that it is designed to cause spatial resolution to grow at the maximum possible rate (i.e., as the square root of the mean quantum catch). From a modeling standpoint, this singled-minded design philosophy has two unfortunate consequences: it forces both the peak frequency and the high-frequency cut-off of the CSF to grow faster than they should to match human performance, and it has the perverse effect of causing the low-frequency contrast sensitivity to decrease as the retinal illuminance increases. As the mean quantum catch increases to very high levels (above 10^3) this last problem becomes especially critical, because the MTF of the CV operator begins to eliminate all the spatial frequencies that can actually get through the optics of the eye (Fig. 14), creating a kind of self-imposed blindness.

It is not clear how these high-light defects could be remedied. One natural starting point (suggested in Ref. 1) would be to assume a compressive transformation of the receptor signal, e.g., to take the input to the CV operator to be some power function of the quantum catch rather than the catch itself. In this case the Gaussian CV operator would become

$$G[Q](x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [Q(u, v)]^p (1/2\pi\sigma^2) \exp\{(-1/2\sigma^2) \times [Q(u, v)]^p [(x - u)^2 + (y - v)^2]\} dudv. \quad (10.1)$$

In the deterministic case it is easy to show that a generalized CV operator of this sort satisfies the scaling theorem and consequently still obeys Weber's law.¹ However, acuity no longer grows as the square root of the mean illuminance L but grows instead $L^{p/2}$, and this is true of the entire MTF: plotted against log frequency, it shifts along the axis so that its peak frequency is proportional to $L^{p/2}$. If this remains true in the noisy-input case (which seems likely), the relationship between the mean illuminance and the CSF could

be made less volatile by an appropriate choice of the exponent p .

However, it would still be the case that as illuminance rises, the CSF would shift bodily toward higher frequencies, and so sensitivity would still fall at low frequencies. This last problem is frustrating. It arises from an excessive quantum catch (the visual analog of being too rich) and might seem to be solved easily by simply throwing away part of that catch, for example, by performing the CV operation on a quantum catch accumulated over a shorter time interval. However, the original (i.e., larger) catch must still be available to support high-frequency performance, and so this approach would require multiple memories: one for the full quantum catch at each receptor over the last 200 msec (for example), another for the catch over the last 100 msec, etc. This approach seems much too cumbersome to be worth pursuing, and I have not found any attractive alternative.

In the face of these difficulties, is it worthwhile to invest further effort in developing CV-based models for biological visual systems? The answer obviously depends on whether they correspond to any physiological reality, but that point is not so easily decided as one might suppose. Current physiologically informed descriptions of the retina¹⁸ certainly contain nothing resembling CV mechanisms, and it might seem that the issue is long since settled, since CV operators involve no lateral inhibition: their point-spread functions are never negative. But lateral inhibition itself is only a hypothesis designed to explain physiological results, e.g., the bandpass nature of the CSF's of retinal ganglion cells,¹⁹ and for the most part, the same results that are usually taken to demonstrate lateral inhibition would also be created by CV operators. Also, we have seen that CV operators provide a natural account of the apparent failure of lateral inhibition at low light levels, the apparent loss of the inhibitory surround portion of the ganglion-cell receptive field.²⁰ Thus the fact that CV operators do not appear in current physiological models does not prove that they do not exist in the retina. It only shows that physiologists have not been aware of them as theoretical tools.

What sort of experiment could definitely rule out the possibility of CV operations in the retina? The most direct approach would be to measure the CSF's of individual retinal ganglion cells at retinal illuminance levels spanning a broad range. Such measurements have been made for cat ganglion cells by Derrington and Lennie,²¹ who found that the CSF (for X cells) could always be fitted by adjusting the center and surround sensitivity parameters of a linear difference-of-Gaussians receptive-field model, with no need to vary the spatial scale parameter of either the center or the surround mechanism. Enroth-Cugell and Robson¹⁹ also reported essentially the same result in their classic paper on the X-Y distinction. (In their case the spatial scale parameters had to be altered to fit the CSF at low light levels, but only by a negligible amount.) This seems to show that the cat retina does not operate in a CV fashion: otherwise there should be much larger changes in the apparent size of receptive fields.

However, this does not necessarily mean that things work the same way in the primate retina: cat visual acuity increases hardly at all as the mean luminance rises from 10^{-5} to 100 cd/m^2 , while human acuity rises by a factor of the order of 50.²² Apparently there are rather substantial dif-

ferences in the retinal hardware of cats and monkeys.²³ I am not aware of any experiment along the lines of that done by Derrington and Lennie with monkeys, but data of that sort could provide a clear-cut answer to the question of whether anything resembling a CV operation is carried out the primate retina.

I suspect that the answer is no, first, because of the intrinsic defects of that operation as a psychophysical model, and second, because it seems so improbable that cat and primate retinas are designed according to fundamentally different principles. In addition, current work indicates that the vertebrate retina has found other ways of altering its spatial filtering properties as a function of the prevailing light level: in the fish retina, recent evidence suggests that the effectiveness of lateral inhibition mediated by horizontal cells is modulated by feedback signals carried from the inner to the outer plexiform layer by interplexiform cells.²⁴ In cat retina it is thought that a similar result is achieved by an illuminance-dependent modulation of the effectiveness of rod-cone gap junction.¹⁸

Altogether, then, it seems likely that the remarkable similarities between the effects of CV operators and the global properties of spatial vision stem not from an identity of mechanism but from a common preoccupation with fundamental physical problems faced by all visual systems: problems created by photon noise and by the extremely large dynamic range of retinal image intensities in natural environments. The Poisson statistics of photon noise dictate an intensity-dependent spatial-summation mechanism to maximize resolution across different light levels, and the mismatch between the dynamic ranges of retinal inputs and outputs dictates a dc suppression mechanism. In Sections 2 and 3 it was shown that if we try to achieve these two goals by a single operation located at the photoreceptors, we are led inevitably to the class of CV operators, which in turn automatically give rise to bandpass filtering and Weber's-law behavior. From that perspective, then, all the essential features of spatial vision (that is, those that we think of as being due to retinal processes) can be seen as arising from a single algorithm designed to solve two problems that must be solved by any visual system. If our retina solves these same problems in a non-CV fashion and creates the same essential properties, this suggests that the formal linkage between them is preserved even if we drop the assumption that both problems must be solved simultaneously at the receptor level. The interesting theoretical question is whether one can find an axiomatic framework that makes such a linkage apparent.

ACKNOWLEDGMENTS

I thank S. Reuman for programming assistance, A. Ahumada and D. I. A. MacLeod for advice, T. Cornsweet for proposing the idea of constant-volume operators, and D. Bosman for suggesting their name. This research was supported by the National Aeronautics and Space Administration under joint research interchange NCA2-5.

REFERENCES

1. T. N. Cornsweet and J. I. Yellott, Jr., "Intensity-dependent spatial summation," *J. Opt. Soc. Am. A* **2**, 1769-1786 (1985); Errata *J. Opt. Soc. Am. A* **3**, 165 (1986).

2. H. B. Barlow, "Increment thresholds at low intensities considered as signal/noise discriminations," *J. Physiol. (London)* **136**, 469-488 (1957).
3. There is no formal barrier to modeling lateral inhibition in a CV fashion, i.e., allowing the spread function to assume negative values. For example, the function S in Eq. (1.1) could be a difference of Gaussians. However this leads to grotesque output images: contrast reversals occur at low spatial frequencies, and the Mach bands at edges bulge in the wrong direction. In the nonlinear world of CV operators, lateral inhibition is both unnecessary and insufficient for modeling visual processes. This has been shown by J. I. Yellott, Jr., T. N. Cornsweet, and S. R. Reuman, "Intensity-dependent spatial summation and lateral inhibition," *Invest. Ophthalmol. Vis. Sci.* **26**, 138 (1985).
4. F. I. Van Nes and M. A. Bouman, "Spatial modulation transfer in the human eye," *J. Opt. Soc. Am.* **12**, 401-406 (1967). Van Nes and Bouman plotted their contrast-sensitivity functions upside-down relative to current conventions. Their data are replotted by G. Westheimer, "Visual acuity and spatial modulation thresholds," in *Handbook of Sensory Physiology. Visual Psychophysics*, D. Jameson and L. M. Hurvich, eds. (Springer-Verlag, Berlin, 1972).
5. H. B. Barlow, "Temporal and spatial summation in human vision at different background intensities," *J. Physiol. (London)* **141**, 337-350 (1958).
6. S. Shlaer, "The relation between visual acuity and illumination," *J. Gen. Physiol.* **21**, 165-188 (1937). Shlaer's data are shown by J. P. Thomas, "Spatial resolution and spatial interaction," in *Handbook of Perception*, E. C. Carterette and M. P. Friedman, eds. (Academic, New York, 1979), Vol. 5, Chap. 7, Fig. 3. His measurements of grating acuity do not extend below a retina illuminance of $-3 \log Td$, but extrapolation suggests that at $-4 \log Td$, acuity would be 0.6 cycle/deg.
7. G. J. Burton, "Evidence for nonlinear response processes in the human visual system from measurements of the thresholds of spatial beat frequencies," *Vision Res.* **13**, 1211-1225 (1973); D. I. A. MacLeod, D. R. Williams, and W. Makous, "Difference frequency gratings above the resolution limit," *Invest. Ophthalmol. Vis. Sci.* **26**, 11 (1985).
8. A. Rose, "The relative sensitivities of television pickup tubes, photographic film and the human eye," *Proc. IRE* **30**, 293-300 (1942); "The sensitivity performance of the human eye on an absolute scale," *J. Opt. Soc. Am.* **38**, 196-208 (1948).
9. H. deVries, "The quantum character of light and its bearing upon threshold of vision, the differential sensitivity and visual acuity of the eye," *Physica* **11**, 179-189 (1943).
10. D. A. Baylor, B. J. Nunn, and J. L. Schnapf, "The photocurrent, noise, and spectral sensitivity of rods of the monkey *Macaca fascicularis*," *J. Physiol. (London)* **357**, 575-607 (1984).
11. A. Rose, *Vision: Human and Electronic* (Plenum, New York, 1974).
12. J. Aczel, *Lectures on Functional Equations and Their Applications* (Academic, New York, 1966).
13. D. Marr and E. Hildreth, "Theory of edge detection," *Proc. R. Soc. London Ser. B* **207**, 187-217 (1980).
14. F. W. Campbell and D. G. Green, "Optical and retinal factors affecting visual resolution," *J. Physiol. (London)* **181**, 576-593 (1965).
15. R. W. Gubisch, "Optical performance of the human eye," *J. Opt. Soc. Am.* **57**, 407-415 (1967).
16. M. Aguilar and W. S. Stiles, "Saturation of the rod mechanism of the retina at high levels of stimulation," *Opt. Acta* **1**, 59-64 (1954).
17. H. B. Barlow, "Retinal and central factors in human vision limited by noise," in *Vertebrate Photoreception*, H. B. Barlow and P. Fatt, eds. (Academic, London, 1977).
18. P. Sterling, M. Freed, and R. G. Smith, "Microcircuitry and functional architecture of the cat retina," *Trends Neurosci.* **9**, 186-193 (1986).
19. C. Enroth-Cugell and J. G. Robson, "The contrast sensitivity of the ganglion cells of the cat," *J. Physiol. (London)* **187**, 517-552 (1966).
20. H. B. Barlow, R. Fitzhugh, and S. W. Kuffler, "Change of organization in the receptive fields of the cat's retina during dark adaptation," *J. Physiol. (London)* **137**, 338-354 (1957).
21. A. M. Derrington and P. Lennie, "The influence of temporal frequency and adaptation level on receptive field organization of retinal ganglion cells in cat," *J. Physiol. (London)* **333**, 343-366 (1982).
22. T. Pasternak and W. H. Merigan, "The luminance dependence of spatial vision in the cat," *Vision Res.* **21**, 1333-1339 (1981).
23. R. Shapley and V. H. Perry, "Cat and monkey retinal ganglion cells and their visual functional roles," *Trends Neurosci.* **9**, 229-235 (1986).
24. J. E. Dowling, "Dopamine: a retinal neuromodulator," *Trends Neurosci.* **9**, 236-240 (1986).