IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, VOL. SMC-7, NO. 3, MARCH 1977

A Nonlinear Model for the Spatial Characteristics of the Human Visual System

CHARLES F. HALL, MEMBER, IEEE, AND ERNEST L. HALL, MEMBER, IEEE

***Abstract*—Several recent papers have presented data from experimental investigations of the human ivsual system (HVS) which support the general hypothesis that the HVS is composed of spatial frequency channels. It has been pointed out, however, that a linear systems analysis of the entire system is not valid. Furthermore, a nonlinear model consisting of a log-bandpass filter produced some experimental results with deviations at high spatial frequencies. A new structure for a nonlinear mathematical model which is easily quantifiable, produces results which compare with experimental data, and has a physiological correlate is presented. The significance of this model is that the bandwidth of the visual system decreases as contrast increases. Thus the system appears to maximize the signal to noise ratio while attempting to maintain a constant "perceptual" spatial-frequency fidelity.**

I. Introduction

IT HAS BEEN suggested by Campbell and Robson [1] that the human visual system (HVS) may contain channels tuned to spatial frequencies. An excellent review of psychophysical experimentation in this area can be found in Campbell [2]. In addition to psychophysical support for this hypothesis, there exists neurophysiological evidence that neurons in the visual cortex of the cat and the monkey are sharply tuned to spatial frequency as well as orientation [3]–[8]. Pollen and Taylor [9] have suggested, based on their experimental work, that the complex cell structure of the striate cortex of the HVS may be performing two-dimensional spatial decompositions of subdomains of the visual space. They have specified electrophysiological properties of a neural substratum that could support such a spatial transformation.

In 1965 Kabrisky [10] suggested the type of neutral interconnections which could support a cross-correlation-type of computation in the visual cortex. The Kabrisky model for the HVS has been extended to include low spatial-frequency filtered transformations [11]–[13]. This model has been applied to various pattern recognition tasks [14]. Ginsburg [15], [16] has shown the Kabrisky model has psychological correlates which demonstrate the Gestalt principles of closure, similarity, and proximity. In addition, he has demonstrated that spatial filters based on the visual system modulation transfer function (MTF) are superior to

Manuscript received April 5, 1976; revised October 21, 1976. This work was supported by the Advanced Research Projects Agency under Contract No. F08606-72-0008 ARPA Order No. 1706 and by the Air Force Avionics Laboratory, AFSC, WPAFB, under Contract No. F33615-77-C-1016.

C. F. Hall is with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90007.

E. L. Hall is with the Department of Electrical Engineering, University of Tennessee, Knoxville, TN 37916.

simple low-pass spatial filters [17]. Andrews has found that the number of features required for minimum mean-square error is fewer in the discrete Fourier transform domain (spatial frequency domain) than in the Walsh/Hadamard or Haar transform domains [18], [19]. Thus the spatial frequency domain appears to be an effective and efficient representation space for visual imagery.

The wide diversity and large amount of evidence indicating that a spatial frequency decomposition occurs in the visual system is difficult to ignore. The fact that a simple pattern recognition system based on such a premise has been used successfully makes a model of the HVS based on spatial frequency transformations practical as well as satisfying. A convenient concomitance of such a model is that linear systems techniques can be used in analyzing it. The main objective becomes that of finding a MTF for the HVS. Once the MTF is specified, the response of the visual system to any input pattern can be computed. This capability has obvious applications in image processing, analysis, and recognition.

Unfortunately, the human visual system contains nonlinearities. This fact is easily established by analyzing the phenomena of "brightness constancy" and "simultaneous contrast" [20, pp. 270–284, 352–353]. One is left with the choice of ignoring the nonlinearities, developing a separate describing function for each region of interest, or performing the complex nonlinear systems analysis. The first choice is suitable for a first approximation. This approach will be considered in the next section. The linear model will then be extended to include a nonlinearity. This simple nonlinear model will be shown to be inadequate as a concise model of the HVS. In Section IV a modified nonlinear model will be presented. This model is consistent with the spatial frequency decomposition requirement, contains a nonlinearity in an appropriate location, and, therefore, reflects the physiological structure of the eye. A discussion of the relevance and application of our model is contained in the last section.

II. First Approximation Models

A first approximation model for the HVS can be made by considering the system to be linear, isotropic, and time- and space-invariant. Other assumptions which are often implied, although not explicitly stated, are that the system is monocular, monochromatic, and photopic.

The use of the MTF to estimate the system response to an input image assumes that the system is linear and thus obeys the principle of superposition. In general, if the intensity of the light radiated from an object is increased, the magnitude of the response of the HVS should increase proportionally. We will initially assume the response increase to be directly proportional to the stimulus, i.e., a linear system.

An isotropic system is one which has characteristics which are invariant to direction. The visual system is not isotropic [21]. The response of the system to a rotated contrast grating is a function of frequency of the grating as well as angle of orientation. The sensitivity of the system decreases to a minimum at 45° and then rises again reaching the original level at 90° rotation. At the point of maximum deviation, 45°, a frequency of 30 cycles/degree would be −3 dB below the response at 0° rotation. A frequency of 10 cycles/degree is reduced by only 15% when rotated from the vertical or horizontal to a 45° position. This anisotropic behavior will not be included in our models. It can be added easily when the application warrants the increased complexity.

The assumption of spatial and temporal invariance or homogeneity is difficult to modify quantitatively. The HVS is not spatially homogeneous, neither in optics nor receptors. However, optical spatial invariance is a good assumption near the optic axis, and even though the densities of rods and cones vary strongly with retinal position, the foveal region is relatively homogeneous [22, pp. 47–52]. Davidson has pointed out that the anatomical evidence against homogeneity is weakened by the fact that certain structurally inhomogeneous systems are functionally self-homogenizing [23]. The situation in the case of temporal homogeneity is even more complicated. The apparent existence of sustained and transient channels in the visual system [24]–[26] produce complex and interrelated responses to psychophysical stimuli which are difficult to interpret [27]–[30]. The models discussed in this paper will not consider temporal responses, and thus temporal homogeneity is not a factor.

Since we are not considering depth perception in our model, the differences between monocular and binocular vision are primarily due to the effects on absolute threshold and resolution. Campbell and Green [31] have shown that binocular threshold is lower than monocular threshold by a factor of $\sqrt{2}$. Furthermore, they found binocular resolution to be approximately 7% better than monocular resolution. Because of these simple relationships, a monocular model can be easily extended to an equivalent binocular model without depth perception. Since the psychophysical experiments referenced in developing our model stimulated one eye, we will model the monocular case.

Monochromatic vision may be defined as the inability to distinguish differences in hue. Since the illumination sources used in the experimental protocols of interest are of a constant spectral content (usually white or blue-green light), there are no hue variations in the stimuli. Thus monochromatic vision is an acceptable restriction to the model.

The assumptions detailed previously are valid for the system which produces the MTF derived via contrast gratings as described in Cornsweet [20, pp. 330–342]. This MTF is shown in Fig. 1. The question then becomes, "How can the system which produced this response be modeled?" Of course, one can just represent the system as a bandpass filter (Fig. 2(a)). This is difficult to defend from a physiological standpoint since the response is a compound one due to several mechanisms within the HVS.

A slightly more detailed model would be a combination of a low-pass and a high-pass filter (Fig. 2(b)). The high-pass portion, which produces the well-known Mach band phenomenon, is a result of lateral inhibition [20, p. 358].

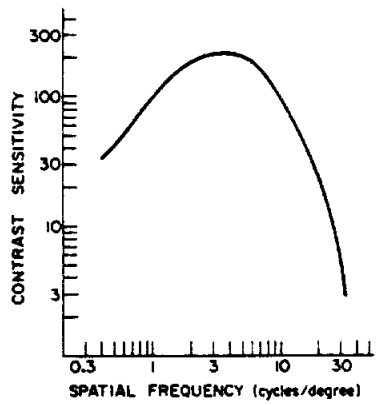


Fig. 1. Spatial frequency response function of human visual system.

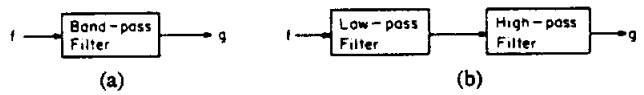


Fig. 2. (a) Simple bandpass filter model of the HVS. (b) Low-pass, high-pass filter model of the HVS.

The low-pass filter is determined by several mechanisms. The optical characteristics of the eye, including pupil size, obviously affect the high-frequency response of the system. The size and density of the photoreceptors and neural summation will also tend to limit resolution or high spatial frequency response. In addition, light scattering within the aqueous humor is probably an increasing function of frequency and would cause a lower sensitivity at these frequencies due to the increased losses. Campbell has concluded [32, p. 182] that the answer to the question, "Which is it that limits resolution, optics or the retina-brain complex?" is not a simple one. He further states that for large pupils (2.8 mm in diameter or greater) or poor focus, the optics does most of the attenuation. The retina-brain complex limits resolution when the pupil diameter is 2.4 mm and the system is focused. Riggs [33, pp. 333–334] has pointed out that for pupil diameters less than 2 mm, the limits of resolution are established by Rayleigh's diffraction limit.

Although the model depicted in Fig. 2(b) can be used to describe a system which produces an MTF similar to the HVS, it will not account for the phenomenon of "brightness constancy." In general, the brightness of an object tends to remain constant, despite variations in the illumination falling on it. This phenomenon can be accounted for by introducing a nonlinearity in the system [20, p. 373].

III. Simple Nonlinear Models

In the previous section it was pointed out that if a system is linear (or nonlinear operating in a linear range), Fourier techniques can be used for analysis. In particular, an MTF can be determined, and multiplication of the Fourier transform of any object by this MTF will give the transform of the image. In other words, the system response can be determined for any input stimulus. In the HVS these techniques are applicable up to the photoreceptors in the retina. This is the case since to this point we are dealing with electromagnetic radiations within the optics of the eye [34]. Beyond this point the assumptions necessary to justify a linear systems analysis are difficult to defend.

There are two types of photoreceptors in the retina, rods and cones. These receptors have different functions. The rods function primarily at low intensities and are not involved in color vision. The cones, on the other hand, function at higher light intensities and are responsible for color vision. In addition, the area of the retina which has maximum photoreceptor density and minimum neural summation, the fovea,[1] is rod free. Thus the cones contribute to maximal acuity (in terms of resolution). Kabrisky has likened this effect to viewing the world through a piece of frosted glass with a clear spot in the middle [10, p. 18]. The clear spot is analogous to the fovea, and we are not cognizant of the "blurring" since the clear spot is always where we are looking. This duality of the retinal receptors is sometimes referred to as the duplicity theory [35, p. 387].

It is not yet known exactly how a quantum of light stimulating a photoreceptor elicits an electrical response from an accompanying nerve cell. It is known that for a dark adapted eye one quantum is sufficient to stimulate a single rod, and approximately nine rods within a 50 μ diameter area must be stimulated within 1 ms of each other before the stimulus is perceived [20, pp. 23–26]. Thus an absolute minimum threshold nonlinearity exists in the system. This threshold varies as a function of retinal location and time in the dark [36, pp. 81–84]. If an input image were to remain above this threshold and no other nonlinearities existed in the HVS, then a linear systems analysis would be valid. Unfortunately, other nonlinearities exist.

In 1932 Hartline and Graham reported the recording of neurological signals from single receptors in the eye of the Limulus (the horseshoe crab) [37]. Subsequently, Fuortes measured the electrical properties of the nerve cells of the eye of the Limulus [38], [39]. His findings were consistent with the hypothesis that light stimulating the receptor causes a potential difference between the inside and outside of the receptor cell. This potential difference mediates the frequency of adjacent nerve impulse firings. Rushton felt that Fuortes' data indicate that the resistance of the cell membrane is proportional to the logarithm of the total light incident on the receptor [40]. Furthermore, he concluded that the relationship between the membrane potential and the frequency of impulses is a linear one and, therefore, the frequency is a logarithmic function of light intensity. Rushton compared this result with the stimulus–response relationship of mechanical receptors and observed that the logarithmic transformation seems to be a property of sense organs in general [40, p. 177], i.e., they follow the Fechner logarithmic law.

Stevens [41], [42, pp. 208–209] has replotted Hartline's data on a log-log scale and pointed out that the functional

[1] The fovea is a circular area of ~ 0.5 mm in diameter centered on the optical axis of the eye. The density of cones in this area is approximately 2×10^5 mm^{-2}.

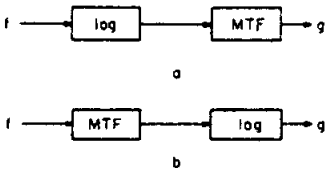


Fig. 3. Two simple nonlinear models of HVS.

relationship follows a power law with an exponent of 0.29. It is interesting to note that Stevens discussed his power law vis-a-vis Fechner's logarithmic law at the same symposium at which Rushton reported his analysis of Fuortes' data [43]. Stevens has also reported that the exponent for perceptual brightness of 5° targets in the dark is 0.33 [42, p. 15]. This number represents all of the nonlinearity in the system since the experimental paradigm now includes the perception of the stimulus and the response is not just the monitoring of a neurological signal. In other words, the 0.29 exponent is based on physiological data. The 0.33 exponent is from psychophysical data. Uttal, in discussing the coding of sensory magnitude, has concluded that the main site of response compression occurs at the receptors [44, p. 317], and the above data indicate this is true. In any case, the nonlinearity can be approximated by a logarithmic function, and we will do this for convenience.

Two simple ways this nonlinearity can be added to a model of the HVS are shown in Fig. 3. The question which immediately arises is, "How can an MTF, which requires a linear system, be included in such a model?" The argument that is usually made in support of such a structure is that for small variations in intensity of input patterns the system is operating in essentially a "linear" region of the nonlinearity [20, pp. 334–335]. Recently, Mansfield has reported that the system is, in fact, linear for intensities near absolute threshold and follows a power law with exponent 0.32 above in intensity of 10^{-2} trolands [45, p. 682]. Further, under increasing light adaptation conditions, the linear region becomes larger. Of the two structures shown in Fig. 3, (a) can be used to explain the phenomenon of brightness constancy [20, pp. 352–353].

Fig. 4(a) contains the intensity profiles of two hypothetical images. The difference between the two profiles is that the intensity of the second is five times greater everywhere. A human observer would not detect this difference, and in fact, the two images would be perceived as having the same "brightness." A linear model for the HVS would filter out the average intensity level, passing the higher frequency transition of the step change in intensity. Since the step itself has been increased by a factor of five, the change would be preserved by the linear system. If, however, the model is nonlinear and in particular, logarithmic, then the situation depicted in Fig. 4(b) occurs. As can be seen, the step in intensity is now equal in amplitude for both intensity levels. After high-pass filtering, this constant amplitude step is the only thing that passes through the system; thus brightness constancy results. We see then that the model of Fig. 3(a) not only is physiologically sound, but that it

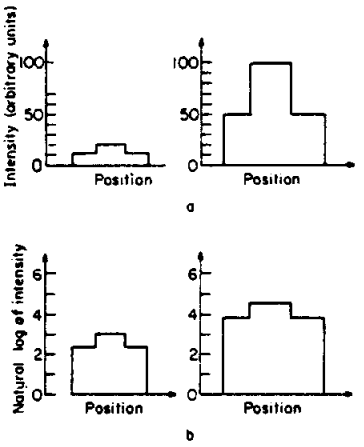


Fig. 4. Intensity profiles illustrating "brightness constancy" (after Cornsweet [20, p. 355]).

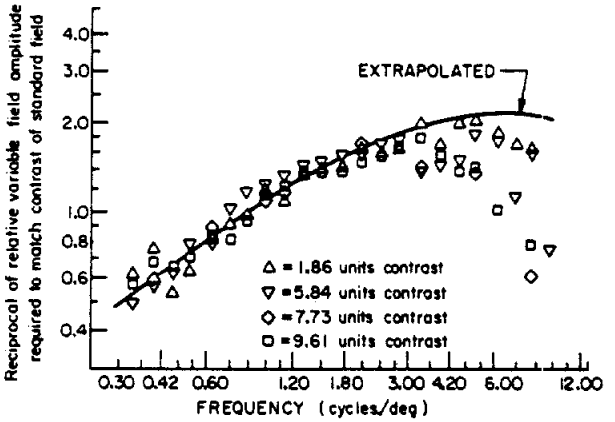


Fig. 5. HVS response to exponential sine wave gratings (from Davidson [23, p. 1305]).

predicts brightness constancy as well. Stockham has successfully applied a version of this model to image processing [46]. In addition, Mannos and Sakrison have found the same basic model to be appropriate for image coding [47]. They also found that images which were coded using a cube root function were judged subjectively as being "better" than those coded via a logarithmic function.

In 1968, Davidson [23] performed the contrast grating experiment with exponential sine-wave gratings. He reasoned that this stimulus should linearize the HVS. The results are shown in Fig. 5. Davidson noted the variation of the measured describing functions with contrast at high frequencies and observed that this result implies a highly nonlinear system at high frequencies. He then convincingly argued that the finding may have been an artifact.

Recently, Henning, Hertz, and Broadbent [48] have reported experimental results which indicate a nonlinear distortion of signals at high, but not low, spatial frequencies. Their results do not support a model consisting of a logarithmic nonlinearity followed by linear independent frequency channels. The model discussed in the next section is consistent with their findings and explains the high frequency variations in Davidson's data.

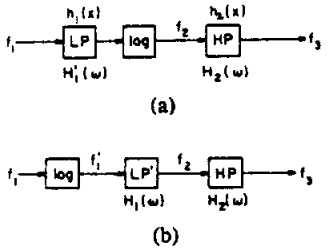


Fig. 6. (a) Proposed nonlinear model for spatial characteristics of HVS. (b) Nonlinear model of HVS assumed by Davidson.

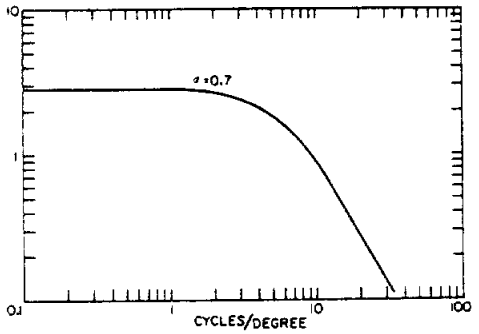


Fig. 7. Spatial frequency response function $H_1(\omega)$ as determined by Fourier transform of line spread function for pupil size of 3 mm.

IV. A New Model

Our model is shown in Fig. 6(a). The model has simply placed the low-pass filter in front of the nonlinearity. This modification can be justified by the physical location of the mechanisms which are responsible for this filter, as discussed in Section II. Westheimer and Campbell [49] measured the light distribution formed by the living human eye and found the line spread function (LSF) to be of the form

$$h_1(x) = \exp(-\alpha|x|). \tag{1}$$

For white light and a pupil diameter of 3 mm, α has a value of 0.7. The MTF of this system can be computed from the LSF. Assuming the one-dimensional case, the Fourier transform becomes

$$H_1(\omega) = \int_{-\infty}^{\infty} h_1(x) \exp(-j\omega x)\, dx. \tag{2}$$

Performing the integration gives

$$H_1(\omega) = \frac{2\alpha}{\alpha^2 + \omega^2}, \tag{3}$$

which determines the frequency response of the low-pass filter. A plot of this function with $\alpha = 0.7$ is shown in Fig. 7. The -3 dB point occurs at approximately 6.6 cycles/degree, well within the passband of the visual system MTF. Thus it can hardly be disputed that anatomically, this low-pass filter modifies all images before the receptors in the HVS are encountered. It will now be shown that this different structure has a significant effect on the spatial frequency response of the system. In particular, the response varies at high frequencies for different contrast levels, and the system requires both a high-pass spatial filter and exponentiation

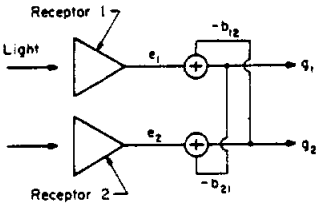


Fig. 8. Backward inhibition model for photoreceptors in HVS.

to linearize it. In contrast, the system of Fig. 6(b) requires only an exponentiation to linearize it.

The effect of the receptors will be modeled using the backward inhibitor model. This is one of four model types proposed for neural networks by Furman [50] and developed in detail by Cornsweet [20, pp. 290–299]. The model for two receptors is shown in Fig. 8. The receptors will be assumed to have a logarithmic response to the incoming light intensity. In Fig. 8 e_i is the signal produced when receptor i is illuminated, and g_i is the output of the ith network. The inhibitory coefficient b_{ij} represents the strength of the inhibition that g_j exerts on g_i. If we assume the threshold of the receptors to be zero and the $g_i \gg 0$, then

$$g_1 = e_1 - b_{12}g_2 \qquad g_2 = e_2 - b_{21}g_1 \tag{4}$$

define the system. Or in general,

$$g_i = e_i - \sum_{j=1}^{n} b_{ij}g_j, \qquad i = 1,2,\cdots,n. \tag{5}$$

This equation may be written in matrix form as

$$g = e - Bg \tag{6}$$

where

$$g = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_n \end{bmatrix}, \qquad e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix},$$

and

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & & & \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}.$$

Equation (6) can be rewritten as

$$(I + B)g = e \tag{7}$$

or

$$g = (I + B)^{-1}e. \tag{8}$$

Now we assume no self-inhibitory interaction and that the inhibitory interaction is an exponentially decreasing function of the distance between receptors, i.e.,

$$b_{ij} = \begin{cases} 0, & i = j \\ a_0 \exp(-a|i - j|), & i \neq j. \end{cases} \tag{9}$$

Under this assumption $(I + B)$ becomes

$$I + B = \begin{bmatrix} 1 & w_1 & w_2 & \cdots & w_n \\ w_1 & 1 & w_1 & \cdots & w_{n-1} \\ \vdots & & & & \\ w_n & w_{n-1} & w_{n-2} & \cdots & 1 \end{bmatrix} \quad (10)$$

where

$$w_k = a_0 \exp(-ka).$$

This matrix is Toeplitz and corresponds to a linear shift-invariant system. Thus the impulse response of the matrix becomes

$$g(x) = (1 - a_0)\delta(x) + a_0 \exp(-a|x|). \quad (11)$$

The Fourier transform of this impulse response is

$$G(\omega) = (1 - a_0) + a_0 \frac{2a}{a^2 + \omega^2}. \quad (12)$$

The matrix $(I + B)^{-1}$ will therefore have a frequency characteristic

$$H_2(\omega) = \frac{1}{G(\omega)} = \frac{a^2 + \omega^2}{2a_0 a + (1 - a_0)(a^2 + \omega^2)}. \quad (13)$$

This function represents a high-pass filter with

$$H_2(0) = \frac{a}{2a_0 + a(1 - a_0)}. \quad (14)$$

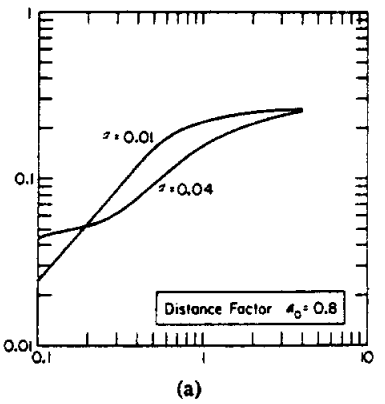
and

$$\lim_{\omega \to \infty} H_2(\omega) = \frac{1}{1 - a_0}. \quad (15)$$

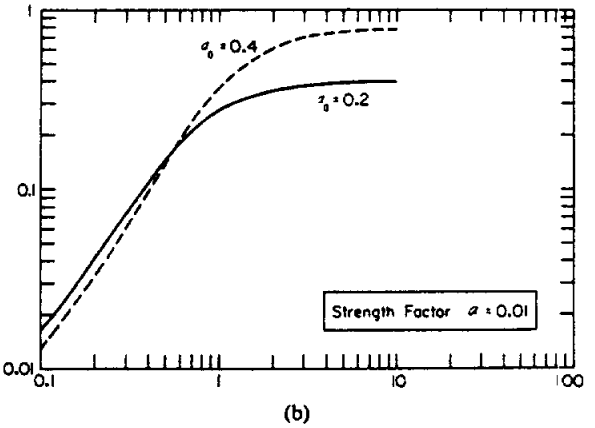
Plots of $H_2(\omega)$ as a function of a_0 and a are shown in Fig. 9. A comparison with the low frequency portion of Fig. 5 indicates that appropriate parameter values for normalized responses of the two systems are $a = 0.01$ and $a_0 = 0.2$. Using these parameters for the high-pass filter and those of the low-pass filter derived earlier, we will now analyze the two nonlinear models illustrated in Fig. 6.

Consider the input of the two nonlinear models to be exponentiated sine wave gratings. The resulting responses will permit an analytical comparison of the two systems, and they may also be compared to the experimental results obtained by Davidson [23]. The configuration shown in Fig. 6(b) is the model assumed by Davidson in his exponential sine wave cancellation experiment. This model performs an exponential-log cancellation at the input, which results in a simplified analysis. The model in Fig. 6(a) does not perform this simple cancellation, and therefore, the two systems are not equivalent.

A mathematical development which contrasts these two models is contained in the Appendix. The essence of this development is the set of frequency response functions shown in Fig. 11. Note that the functions vary at high frequencies as a function of contrast and thus predict Davidson's experimental data shown in Fig. 5. In particular, the high-frequency response increases in bandwidth as the input contrast is decreased. The implications of this result and applications of the new model will be discussed in the following section.



(a)



(b)

Fig. 9. High-pass spatial filter response $H_2(\omega)$ produced by neural inhibition for (a) two inhibition strength values a and a distance factor $a_0 = 0.8$, and (b) two distance factors a_0 and an inhibition strength factor $a = 0.01$.

V. Implications and Applications

In the previous section a new model for the spatial frequency characteristics of the HVS was presented. This model predicts a nonlinear spatial frequency response of the HVS which agrees with experimental data. The model is based on an arrangement of elements in the system in a manner which correlates with the physiological structure of the HVS. It is the placement of nonlinear elements which produces the predictable response. The results were obtained by a combination of mathematical analysis and computer calculations and agree with independent results obtained with psychovisual testing as reported in the literature [23], [48].

The major implication of this model is that the system is analogous to a variable bandwidth filter which is controlled by the contrast of the input image. As input contrast increases, the bandwidth of the system decreases. In an optimal system design one would limit the bandwidth in an attempt to maintain maximum signal to noise ratio. Since imagery noise contains predominently high spatial-frequencies, it would be reasonable to limit this end of the system transfer function. Unfortunately, in most real systems, the signal of interest contains frequencies in the same band as the noise. In our case, the "edge" information is of high spatial-frequency content. Thus a filtering out of noise can only be made at the expense of "blurring" edges.

An optimum system would contain an appropriate tradeoff between system performance (in terms of "edge fidelity") and signal to noise ratio. If we consider noise and edges in the context of gradient operators, high contrast increases the magnitude of the system response. Subjectively, if we view two prints of a low noise photograph which differ only in the contrast at which they were printed, the higher contrast print appears "sharper" and more "crisp." On the other hand, the contrast of a noisy image (such as a "snowy" TV screen) can be increased subjectively by squinting while viewing the screen (squinting increases the high-frequency roll-off). These two examples could illustrate the extreme performance ends of a system attempting to maximize the signal to noise ratio while attempting to maintain a constant perceptual spatial frequency fidelity. In any case, the new model has several applications as a preprocessor for image processing, image coding, pattern recognition, and scene analysis.

The nonlinear structure shown in Fig. 6(a) can be used to form a logical extension of the Kabrisky model of the HVS which has been used successfully in several pattern recognition applications [14]. A fundamental premise upon which his model is based is that the HVS uses low spatial-frequencies as features. In particular, a correlation between low spatial-frequency filtered images and stored prototypes is performed in area 18 or 19 of the striate cortex. If such is the case, then the model developed in this paper has significant relevance as a preprocessor in this type of recognition system. In particular, for complex scenes, which contain variable contrast information, such a preprocessor would have a significant affect on the spatial frequency content of the input signal. It may turn out that just as the human observer directs his attention to various subsections of a complex scene during cognition, an automated system based on this model should compute average contrast over localized areas and adjust the spatial frequency filter parameters accordingly. Such a model would be useful in scene analysis.

In the case of image coding, the structure in Fig. 6(a) is again important as a preprocessor. If our intent is to code an image for transmission over a noisy channel, then it makes sense to first preprocess the image in the same manner as the HVS. We are then left with the problem of coding only that data which is relevant and reduce any noise problems. The same argument holds for bandwidth reduction and efficient storage problems. Thus the functional form which has been proposed by Mannos and Sakrison [47] for studies of visual fidelity should perhaps be modified to

$$\tilde{A}(f_r) = [c_1 + (f_r/f_0)^{k_1}] \exp [-(f_r/m(c)f_0)^{k_2}], \quad (16)$$

where we have added $m(c)$, a function of contrast responsible for modifying the high-frequency roll-off.

Frei has proposed a model for color spatial vision [51] which can be extended in the same manner that we extended the monochromatic model of Fig. 3(a). Frei's model contains a transformation representing the change from spectral energy space $S(x,y,\lambda)$ to tristimulus space $t_1(x,y)$, $t_2(x,y)$, and $t_3(x,y)$. This transformation is placed in front of the configuration shown in Fig. 3(a). A more accurate model would have the low-pass filter portion of the MTF in front of the spectral transformation. In fact, there could be three filters, one for each band, which have different roll-off points caused by chromatic aberrations of the occular media. Investigations of this model in the context of image coding and a fidelity criterion are being pursued.

The low-pass, nonlinearity, high-pass structure is not limited to spatial response or even spectral-spatial response. Spekreijse has shown that this basic structure is valid for modeling the temporal response of the HVS as well [52]. Thus the first stages of the HVS, the optics, the photoreceptors, and retinal interconnectivity can be modeled with a structure which is consistent with spectral and spatio-temporal responses. We might hasten to add that there is nothing new in the physical structure we have discussed. What has been shown is that the assumptions made in the past which permitted simpler models are not valid. Moreover and most significantly, the effects of these assumptions are relevant to system performance, and they can be quantified. The resulting model should permit more accurate studies of the visual system and be a useful addition (as a preprocessor) to various image processing systems.

Appendix

We begin with the analysis of the model in Fig. 6(a) and first determine the contrast response to an exponentiated sinusoid of the low-pass-log system. If the input signal is defined as

$$f_1(x) = \exp (a + m \cos \omega_0 x) \quad (17)$$

where a is the average luminance and m is the peak amplitude of the sinusoidal variation, (17) can be rewritten as

$$f_1(x) = f_0 \exp \{m \cos (\omega_0 x)\} \quad (18)$$

where

$$f_0 = \exp (a). \quad (19)$$

Next, the contrast response of the log-low-pass system to the exponential sinusoid, which is simply its low-pass frequency response, is computed. Finally, computations of the contrast response of the new model are compared to the experimental data.

The Fourier transform of $f_1(x)$ becomes

$$F_1(\omega) = f_0 \mathcal{F}\{\exp [m \cos (\omega_0 x)]\}. \quad (20)$$

From Taylor's theorem this expression can be expanded to

$$F_1(\omega) = f_0 \sum_{n=0}^{\infty} \frac{m^n}{n!} \mathcal{F}\{2^{n-1}[\cos (\omega_0 x)]^n\}. \quad (21)$$

Performing a binomial expansion yields

$$F_1(\omega) = f_0 \sum_{n=0}^{\infty} \frac{m^n}{n!} \frac{1}{2^n} \sum_{i=0}^{n} \binom{n}{i} 2\pi\delta[\omega - (2i - n)\omega_0]. \quad (22)$$

From (3), with $\alpha = 0.7$, we have

$$\mathcal{F}\{h_1(x)\} = \frac{0.14}{0.49 + \omega^2}. \quad (23)$$

Thus the response of $h_1(x)$ to $f_1(x)$ becomes

$$\mathcal{F}\{f_1(x) \circledast h_1(x)\} = \mathcal{F}\{f_1(x)\} \cdot \mathcal{F}\{h_1(x)\} = \frac{0.14}{0.49 + \omega^2} f_0 \sum_{n=0}^{\infty} \frac{m^n}{n!} \frac{2\pi}{2^n} \sum_{i=0}^{n} \binom{n}{i} \cdot \delta[\omega - (2i - n)\omega_0] \tag{24}$$

where $\circledast$ denotes convolution. Therefore, from the inverse Fourier transform,

$$f_1(x) \circledast h_1(x) = f_0 \sum_{n=0}^{\infty} \frac{m^n}{n!} \frac{1}{2^n} \sum_{i=0}^{n} \binom{n}{i} \cdot \frac{0.14 \exp[j(2i - n)\omega_0 x]}{0.49 + [(2i - n)\omega_0]^2}. \tag{25}$$

This convolution is a maximum when

$$\exp[j(2i - n)\omega_0 x] = 1$$

or

$$x = \frac{2\pi m}{\omega_0}, \qquad m = 0,1,2,\cdots. \tag{26}$$

Therefore,

$$[f_1(x) \circledast h_1(x)]_{\max} = f_0 \sum_{n=0}^{\infty} \frac{m^n}{n!} \frac{1}{2^n} \sum_{i=0}^{n} \binom{n}{i} \cdot \frac{0.14}{0.49 + [(2i - n)\omega_0]^2}. \tag{27}$$

Similarly, the minimum response, which occurs for

$$\exp[j(2i - n)\omega_0 x] = 0$$

or

$$x = \frac{\pi(2m + 1)}{\omega_0}, \qquad m = 0,1,2,\cdots \tag{28}$$

becomes

$$[f_1(x) \circledast h_1(x)]_{\min} = f_0 \sum_{n=0}^{\infty} \frac{m^n}{n!} \frac{(-1)^n}{2^n} \sum_{i=0}^{n} \binom{n}{i} \frac{0.14}{0.49 + [(2i - n)\omega_0]^2}. \tag{29}$$

Taking the logarithm of (27) and (29) yields

$$\log[f_1(x) \circledast h_1(x)]_{\max} = \log\left(\frac{2f_0}{0.7}\right) + \log\left\{1 + \frac{0.7}{2} \sum_{n=1}^{\infty} \frac{m^n}{n!} \frac{1}{2^n} \sum_{i=0}^{n} \binom{n}{i} \frac{0.14}{0.49 + [(2i - n)\omega_0]^2}\right\} \tag{30}$$

and

$$\log[f_1(x) \circledast h_1(x)]_{\min} = \log\left(\frac{2f_0}{0.7}\right) + \log\left\{1 + \frac{0.7}{2} \sum_{n=1}^{\infty} \frac{m^n}{n!} \frac{(-1)^n}{2^n} \sum_{i=0}^{n} \binom{n}{i} \cdot \frac{0.14}{0.49 + [(2i - n)\omega_0]^2}\right\}. \tag{31}$$

By using the series expansion for $\log(1 + x)$ and letting

$$P = \frac{0.7}{2} \sum_{n=1}^{\infty} \frac{m^n}{n!} \frac{1}{2^n} \sum_{i=0}^{n} \binom{n}{i} \frac{0.14}{0.49 + [(2i - n)\omega_0]^2} \tag{32}$$

and

$$Q = \frac{0.7}{2} \sum_{n=1}^{\infty} \frac{m^n}{n!} \frac{(-1)^n}{2^n} \sum_{i=0}^{n} \binom{n}{i} \frac{0.14}{0.49 + [(2i - n)\omega_0]^2}, \tag{33}$$

one may obtain

$$\log[f_1(x) \circledast h_1(x)]_{\max} = \log\left(\frac{2}{0.7} f_0\right) + \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} P^j \tag{34}$$

and

$$\log[f_1(x) \circledast h_1(x)]_{\min} = \log\left(\frac{2}{0.7} f_0\right) + \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} Q^j. \tag{35}$$

If we define contrast as the difference between the maximum and minimum logarithmic responses, then

$$C = \log\left(\frac{2}{0.7} f_0\right) + \sum_{j=1}^{\infty} \frac{(-1)^{j+1} P^j}{j} - \log\left(\frac{2}{0.7} f_0\right) - \sum_{j=1}^{\infty} \frac{(-1)^{j+1} Q^j}{j} = \sum_{j=1}^{\infty} \frac{(-1)^{j+1}}{j} (P^j - Q^j). \tag{36}$$

Now consider the model shown in Fig. 6(b). In this case we first take the logarithm of the input giving

$$\log[f_1(x)] = \log\{f_0 \exp[m \cos(\omega_0 x)]\} = a + m \cos(\omega_0 x). \tag{37}$$

Since the log and exponent operations cancel, the system is linearized with respect to the sinusoidal input. The frequency response $H_1(\omega)$ could be determined by observing the modification of the input sinusoidal amplitude at different frequencies. The Fourier transform of (37) becomes

$$\mathcal{F}\{\log f_1(x)\} = 2\pi a\delta(\omega) + m\pi[\delta(\omega + \omega_0) + \delta(\omega - \omega_0)]. \tag{38}$$

For the purpose of analysis, it is interesting to determine the type of filter in the new model which would correspond to a given filter $H_1(\omega)$ in the log-low-pass model. This filter would, unlike $H_1(\omega)$, be input signal dependent. However, the filter shape and variation could then be compared to the empirical results obtained in Davidson's experiment. To determine an equivalent signal dependent filter let

$$H_1'(\omega) \triangleq \mathcal{F}\{h_1'(x)\}. \tag{39}$$

Then the transform of the output becomes

$$F_2'(\omega) = H_1'(\omega)\mathcal{F}\{\log f_1(x)\} = H_1'(\omega)\{2\pi a\delta(\omega) + m\pi[\delta(\omega + \omega_0) + \delta(\omega - \omega_0)]\}. \tag{40}$$

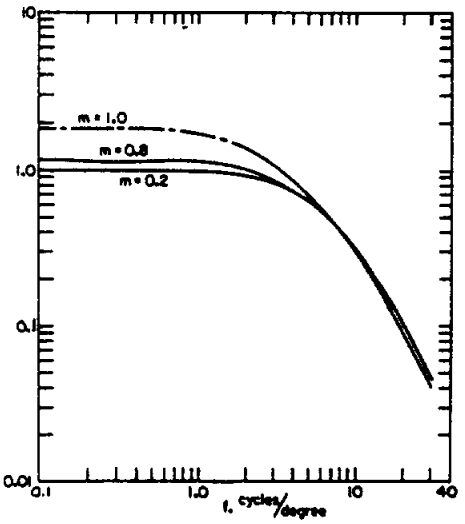


Fig. 10. Spatial frequency response of equivalent signal dependent filter $H_1'(\omega)$ for several values of sinusoidal modulation m.

The inverse transform gives

$$f_2'(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} F_2'(\omega)\exp(j\omega x)\,d\omega$$

$$= aH_1'(0) + \frac{m}{2}[H_1'(\omega_0)\exp(j\omega_0 x) + H_1'(-\omega_0)\exp(j\omega_0 x)]. \quad (41)$$

However, $H_1'(\omega_0) = H_1'(-\omega_0)$, therefore,

$$f_2'(x) = aH_1'(0) + mH_1'(\omega_0)\cos\omega_0 x. \quad (42)$$

Thus

$$[f_2'(x)]_{max} = aH_1'(0) + mH_1'(\omega_0) \quad (43)$$

and

$$[f_2'(x)]_{min} = aH_1'(0) - mH_1'(\omega_0). \quad (44)$$

The contrast for this model becomes

$$C' = aH_1'(0) + mH_1'(\omega_0) - H_1'(0) + mH_1'(\omega_0)$$

$$= 2mH_1'(\omega_0). \quad (45)$$

Equating the contrast of the two models gives

$$2mH_1'(\omega_0) = \sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j}(P^j - Q^j) \quad (46)$$

or

$$H_1'(\omega_0) = \frac{1}{2m}\sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j}(P^j - Q^j). \quad (47)$$

A plot of $H_1'(\omega_0)$ for several values of m is shown in Fig. 10. When this characteristic function is combined with that for the high-pass filter, the curves illustrated in Fig. 11 are obtained. These curves compare favorably with the experimental data of Davidson. In particular, the high-frequency response increases in bandwidth and sensitivity as the input contrast is decreased. Furthermore, this variation is a predictable inherent response of the new model but not of the previous one.

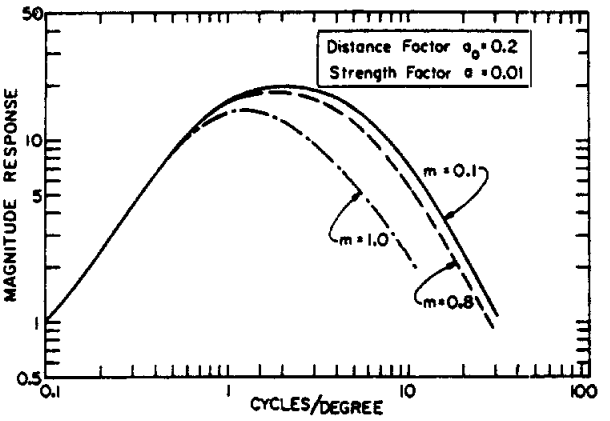


Fig. 11. Overall spatial frequency response $H_1'(\omega)H_2(\omega)$ for several values of sinusoidal modulation m.

Acknowledgment

The authors gratefully acknowledge the assistance of Hideo Murakami in the early stages of this work.

References

[1] F. W. Campbell and J. G. Robson, "Application of Fourier analysis to the visibility of gratings," *J. Physiology (London)*, vol. 197, pp. 551–566, 1968.
[2] F. W. Campbell, "The transmission of spatial information through the visual system," in *The Neurosciences Third Study Program*, F. Worden and F. O. Schmitt, Eds. Cambridge, MA: MIT Press, 1973, ch. 9, pp. 95–103.
[3] D. H. Hubel and T. N. Wiesel, "Receptive fields of single neurones in the cat's striate cortex," *J. Physiology (London)*, vol. 148, pp. 574–591, 1959.
[4] ——, "Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex," *J. Physiology (London)*, vol. 160, pp. 106–154, 1962.
[5] ——, "Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat," *J. Neurophysiology*, vol. 28, pp. 229–289, 1965.
[6] ——, "Receptive fields and functional architecture of monkey striate cortex," *J. Physiology (London)*, vol. 195, pp. 215–243, 1968.
[7] F. W. Campbell, G. F. Cooper, and C. Enroth-Cugell, "The spatial selectivity of the visual cells of the cat," *J. Physiology (London)*, vol. 203, pp. 223–235, 1969.
[8] F. W. Campbell, G. F. Cooper, J. G. Robson, and M. B. Sachs, "The spatial selectivity of visual cells of the cat and the squirrel monkey," *Proceedings of the Physiological Society, J. Physiology (London)*, vol. 204, pp. 120P–121P, July 1969.
[9] D. A. Pollen and J. H. Taylor, "The striate cortex and the spatial analysis of visual space," in *The Neurosciences Third Study Program*, F. Worden and F. O. Schmitt, Eds. Cambridge, MA: MIT Press, 1973, ch. 21, pp. 239–247.
[10] M. Kabrisky, *A Proposed Model for Visual Information Processing in the Human Brain*. Urbana, IL and London: University of Illinois Press, 1966.
[11] ——, "A proposed model for visual information processing in the brain," in *Models for the Perception of Speech and Visual Form*, Wathen-Dunn, Ed. Cambridge, MA: MIT Press, 1967, pp. 354–361.
[12] M. Kabrisky *et al.*, "A theory of pattern perception based on human physiology," *Ergonomics*, vol. 13, pp. 129–142, 1970.
[13] M. Kabrisky, "An introduction to a model of the human visual system," in *IEEE NAECON '73 Record*, Dayton, OH, May 1973.
[14] J. W. Carl and C. F. Hall, "The application of filtered transforms to the general classification problem," *IEEE Trans. Computers*, vol. C-21, no. 7, pp. 785–790, July 1972.
[15] A. P. Ginsburg, "Psychological correlates of a model of the human visual system," in *IEEE NAECON '71 Record*, Dayton, OH, May 1971.
[16] —— *et al.*, "Psychological aspects of a model for the classification of visual images," in *Advances in Cybernetics and Systems*, vol. III, J. Rose, Ed. London: Gordon and Breach, 1974.
[17] ——, "Pattern recognition techniques suggested from psychological correlates of a model of the human visual system," in *IEEE NAECON '73 Record*, Dayton, OH, May 1973.
[18] H. C. Andrews, "Multidimensional rotations in feature selection," *IEEE Trans. Computers*, vol. C-20, pp. 1045–1051, 1971.

[19] ——, *Introduction to Mathematical Techniques in Pattern Recognition*. New York: Wiley-Interscience, 1972.
[20] T. N. Cornsweet, *Visual Perception*. New York: Academic Press, 1970.
[21] F. W. Campbell, J. J. Kulikowski, and J. Levinson, "The effect of orientation on the visual resolution of gratings," *J. Physiology (London)*, vol. 187, pp. 427–436, 1966.
[22] J. L. Brown, "The structure of the visual system," in *Vision and Visual Perception*, C. H. Graham, Ed. New York: Wiley, 1965.
[23] M. Davidson, "Perturbation approach to spatial brightness interaction in human vision," *J. Opt. Soc. Amer.*, vol. 58, pp. 1300–1309, 1968.
[24] J. J. Kulikowski and D. J. Tolhurst, "Psychophysical evidence for sustained and transient detectors in human vision," *J. Physiology (London)*, vol. 232, pp. 149–163, 1973.
[25] D. J. Tolhurst, "Sustained and transient channels in human vision," *Vision Research*, vol. 15, pp. 1151–1155, 1975.
[26] B. G. Cleland, M. W. Dubin, and W. R. Levick, "Sustained and transient neurones in the cat's retina and lateral geniculate nucleus," *J. Physiology (London)*, vol. 217, pp. 473–496, 1971.
[27] P. Tynan and R. Sekuler, "Perceived spatial frequency varies with stimulus duration," *J. Opt. Soc. Amer.*, vol. 64, pp. 1251–1255, Sept. 1974.
[28] D. J. Tolhurst, "Reaction times in the detection of gratings by human observers: A probabilistic mechanism," *Vision Research*, vol. 15, pp. 1143–1149, 1975.
[29] R. Spitzberg and W. Richards, "Broad band spatial filters in the human visual system," *Vision Research*, vol. 15, pp. 837–841, 1975.
[30] J. J. McCann, R. L. Savoy, J. A. Hall, Jr., and J. J. Scarpetti, "Visibility of continuous luminance gradients," *Vision Research*, vol. 14, pp. 917–927, 1974.
[31] F. W. Campbell and D. G. Green, "Monocular versus binocular visual acuity," *Nature*, vol. 208, pp. 191–192, 1965.
[32] F. W. Campbell, "Visual acuity via linear analysis," in *Proceedings of the Symposium on Information Processing in Sight Sensory Systems*, Pasadena, CA. Nov. 1965.
[33] L. A. Riggs, "Visual acuity," in *Vision and Visual Perception*, C. H. Graham, Ed. New York: Wiley, 1965.
[34] G. Westheimer, "Application of Fourier methods to the human visual system," in *Proceedings of the Symposium on Information Processing in Sight Sensory Systems*, Pasadena, CA, Nov. 1965.
[35] D. J. Aidley, *The Physiology of Excitable Cells*. Cambridge: The University Press, 1971.
[36] L. Kaufman, *Sight and Mind, An Introduction to Visual Perception*. New York: Oxford University Press, 1974.
[37] H. K. Hartline and C. H. Graham, "Nerve impulses from single receptors in the eye," *J. Cell. Comp. Physiol.*, vol. 1, pp. 277–295, 1932.
[38] M. G. F. Fuotes, "Electric activity of cells in the eye of Limulus," *Amer. J. Opthal.*, vol. 46, pp. 210–223, 1958.
[39] ——, "Initiation of impulses in the visual cells of Limulus," *J. Physiology (London)*, vol. 148, pp. 14–28, 1959.
[40] W. A. H. Rushton, "Peripheral coding in the nervous system," in *Sensory Communication*, W. A. Rosenblith, Ed. Cambridge, MA: MIT Press, pp. 169–181, 1961.
[41] S. S. Stevens, "Neural events and the psychophysical law," *Science*, vol. 170, pp. 1043–1050, 1970.
[42] ——, *Psychophysics*. New York: Wiley, 1975.
[43] ——, "The psychophysics of sensory function," in *Sensory Communication*, W. A. Rosenblith, Ed. Cambridge, MA: MIT Press, pp. 1–33, 1961.
[44] W. R. Uttal, *The Psychobiology of Sensory Coding*. New York: Harper & Row, 1973.
[45] R. J. W. Mansfield, "Visual adaptation: Retinal transduction, brightness and sensitivity," *Vision Research*, vol. 16, pp. 679–690, 1976.
[46] T. G. Stockham, Jr., "Image processing in the context of a visual model," *Proc. IEEE*, vol. 60, pp. 828–842, 1972.
[47] J. L. Mannos and D. J. Sakrison, "The effects of a visual fidelity criterion on the encoding of images," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 525–536, 1974.
[48] G. B. Henning, B. G. Hertz, and D. E. Broadbent, "Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequency," *Vision Research*, vol. 15, pp. 887–897, 1975.
[49] G. Westheimer and F. W. Campbell, "Light distribution in the image formed by the living human eye," *J. Opt. Soc. Amer.*, vol. 52, pp. 1040–1045, 1962.
[50] G. G. Furman, "Comparison of models for subtractive and shunting lateral-inhibition in receptor-neuron fields," *Kybernetik*, vol. 2, pp. 257–274, 1965.
[51] W. Frei, "A new model of color vision and some practical implications," USC Semiannual Tech. Report, USCIPI 530, Sept. 1, 1973–Feb. 28, 1974, pp. 128–143.
[52] H. Spekreijse, "System analysis approach to the problems of vision," in *Interdisciplinary Investigation of the Brain*, J. P. Nicholson, Ed. London: Plenum Press, pp. 193–211, 1972.