

# Image processing by the human eye

Larry N. Thibos

Indiana University, Department of Visual Sciences, School of Optometry, Bloomington, IN 47405

## ABSTRACT

Image processing by the eye is treated as a classical example of concatenated linear filters followed by a sampling operation. The first filter is optical and is characterized by an optical point-spread function. The second filter is neural and is characterized by the neural point-spread function, which is shown to be related to the receptive fields of retinal neurons. Sampling renders the internal "neural image" a discrete signal subject to the effects of aliasing. Conditions responsible for aliasing are formulated in terms of the amount of overlap of retinal samplers. Evidence of aliasing in human vision is presented along with a simulation of an aliased neural image in the peripheral visual field.

## 1. INTRODUCTION

Initial processing of visual input by the eye can be conceived as a two-stage process. The first stage is low-pass spatial filtering which occurs when the eye's optical system forms a retinal image. The second stage involves the sampling of the continuous optical image by a discrete array of retinal neurons. Although the transduction of light into neural signals is performed by the tiny photoreceptor cells of the retina, signals from many receptors are pooled by subsequent second-order retinal neurons and again by the third-order neurons which form the optic nerve. These pooling operations cause further spatial filtering of the discrete, neural representation of the visual scene. Taken together, optical and neural spatial filtering causes each optic nerve fiber to be receptive to light over an appreciable area of the visual field. Thus, we may begin to describe the early stages of visual processing by modelling the retina as a locally homogeneous array of neurons in which each neuron samples the retinal image by summing over large, overlapping regions. The purpose of this paper is to develop such a model and then use it to assess the importance of optical and retinal processing on image coding in the human visual system.

Although the behavior of human optic-nerve fibers has never been observed experimentally, recent optical and perceptual experiments on humans and also physiological and anatomical experiments on animal models yield a clear picture of their likely characteristics. Thus, part of the motivation for this paper was to succinctly present the main features of physiological optics and retinal architecture for the non-specialist who is interested mainly in the constraints placed on the man-machine interface by early stages of visual processing. Accordingly, in preparing this brief review most references to the primary literature of visual science have been omitted in favor of recent, comprehensive reviews and other reference works.

## 2. CHARACTERIZATION OF THE NEURAL IMAGE

The purpose of the next section is to introduce the concept of a neural image and to describe its primary attributes following the early stages of visual processing. Linear filter theory is used to develop a quantitative model of the neural image as a way of accounting for the effects of optical and neural spatial filtering performed by the eye and to account also for the effects of neural sampling. In section 3 it will be shown that throughout the retina, with the exception of a central region called the fovea, the array of optic fiber neurons under-samples the retinal image. This occurs because their spatial density is well below the Nyquist limit required to faithfully represent the retinal image. Parameters of the model are estimated from recent experimental data and the results used to account perceptual aliasing in human vision, which is simulated in section 4.

### 2.1 Stages of signal processing in the eye

2.1.1 Optics Vision begins with the formation of a light image upon the retina by the optical system of the eye as illustrated in Fig. 1. Optical imperfections and diffraction inevitably reduce image contrast in a way that may be described as low-pass spatial-filtering. If pupil diameter is less than about 2.5 mm, optical quality of the human eye for foveal vision can be nearly diffraction

limited but for larger pupils, ocular aberrations limit any further improvement in retinal image quality<sup>1</sup>. Recent experiments have shown that peripheral optical quality out to 30 deg of eccentricity is about the same as in the fovea<sup>2</sup>. So, at least for central and mid-peripheral vision through a mid-sized pupil, it is not unreasonable to suppose that the optical system of the eye is a linear, shift-invariant system. Accordingly, we may calculate<sup>3</sup> the retinal image  $i(x)$  by convolution ( $\ast$ ) of the point spread function (p.s.f.) of the eye  $p(x)$  with the intensity distribution of the object  $o(x)$ . Thus the first stage of the visual system will be characterized by the equation

$$i(x) = o(x) \ast p(x) \quad (1)$$

where  $x$  is a unitless dimension in a Cartesian coordinate reference frame and is related to the visual direction  $\theta$  by the equation  $x = \sin(\theta)$ .

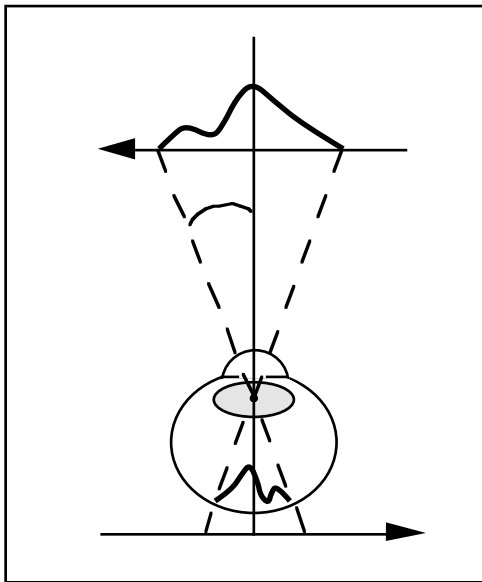


Fig. 1 A coordinate reference frame for vision.

**2.1.2 Transduction** Neural processing of the retinal image begins with the transduction of light energy into corresponding changes of membrane potential of individual photoreceptor cells. Photoreceptors are laid out in a thin sheet which varies in composition across the retina. At the very center of the foveal region, which corresponds to our central field of view, the photoreceptors are exclusively cones but in the parafoveal region rods appear and in the peripheral retina rods are far more numerous<sup>4</sup>. This paper will emphasize

daylight vision by cones. Each cone is thought to integrate the total amount of light energy entering the cell through its own tiny aperture just a few microns in diameter. Since this entrance aperture is wholly within the body of the cone, it will not physically overlap the aperture of neighboring cones. (This is not to say that a point source of light will only stimulate only one cone at a time. In fact, the optical system of the eye will spread the image of a point source over a retinal area which may include several cones.) Based on this arrangement of the cone mosaic, we may characterize the first neural stage of the visual system as a sampling process wherein a continuous retinal image is transduced by an array of non-overlapping samplers. The result is a discrete array of neural signals which will be called a "neural image".

**2.1.3 Optic nerve output** The optic nerve in humans contains roughly one million individual fibers, each of which is an outgrowth of a third order neuron of the retina called a ganglion cell. In general, ganglion cells are functionally connected to many rods and cones by means of intermediate, second order neurons. As a result, a given ganglion cell may respond to light falling over a relatively large region of the retina called its "receptive field", with the middle of the field typically weighted most heavily. The neural connectivity underlying the receptive fields of ganglion cells is known well enough for mammals to draw a schematic wiring diagram<sup>5</sup> which we presume to be a reasonable blueprint for humans as well. Neighboring ganglion cells may receive input from the same cone, which implies that receptive fields of third-order neurons can overlap. In the highly specialized foveal region, however, the receptive fields of ganglion cells are about the size expected for individual cones which gives rise to the notion of essentially one-to-one connectivity of cones to ganglion cells<sup>4</sup>.

Ganglion cells come in many varieties, but one particular class (denoted  $\alpha$ -cells) seems to dominate throughout the primate retina<sup>6</sup>. Physiological experiments on cat and monkey indicate that a  $\alpha$ -ganglion cell responds to a linear combination of light falling on its receptive field<sup>5,7</sup>. Accordingly, a mathematical model of a  $\alpha$ -ganglion cell should be designed to respond by amount  $r$  to the light falling within its receptive field ( $r_f$ ) according to the equation

$$r = \int_{rf} w(x) \cdot i(x) dx \quad (2)$$

where  $w(x)$  denotes the spatial weighting function of the receptive field. Such a model is common currency among neurophysiologists. Notice that by concentrating here upon the weighting function  $w(x)$  for the output neurons of the retina we subsume the effects of two previous stages of neural processing, namely, sampling by photoreceptors and manipulation of the cone neural image by second-order inter-neurons. Also note that for the present investigation of spatial vision it is not necessary to consider the final stage of retinal processing in which the time-continuous quantity  $r$  is encoded (perhaps nonlinearly) for asynchronous digital transmission along the optic nerve.

## 2.2 Spatial description of the neural image

The goal of this section is to build upon the framework erected above in order to give a mathematical description of what the neural image looks like as it leaves the eye *via* the optic nerve. A global analysis encompassing the whole of the visual field is too difficult to attempt here because of the complications introduced by retinal inhomogeneity between fovea and periphery. Instead, attention will be focussed on a local region where the neural architecture of the retina is relatively uniform. Accordingly, consider a homogeneous population of ganglion cells which are responsible for representing the retinal image in a small patch of retina as illustrated in Fig. 2. Although the visual field is two-dimensional, a simpler one-dimensional analysis will be sufficient for developing the main results which follow. By the assumption of homogeneity, the weighting function of each receptive field has the same form but is centered on different  $x$  values for different neurons. The cells need not be equally spaced for the following general results to hold.

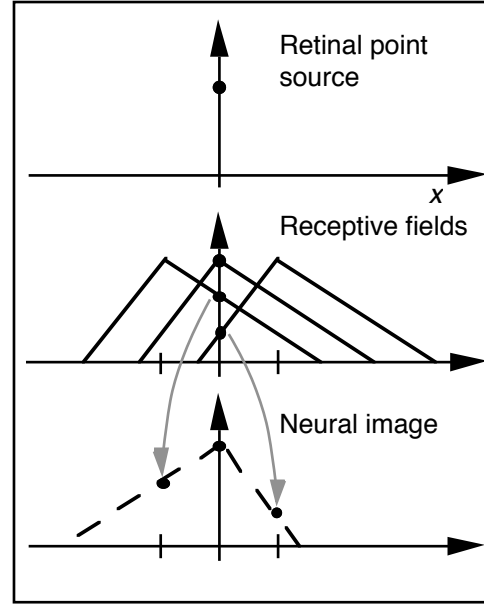


Fig. 2. Neural image for a point source of light on the retina.

If we let  $x_j$  be the center of the receptive field of the  $j^{\text{th}}$  neuron in the array, then the weighting function for that cell will be

$$w_j = w(x - x_j) \quad (3)$$

and the corresponding response  $r_j$  is found by combining (2) and (3) to give

$$r_j = \int_{rf} w(x - x_j) \cdot i(x) dx \quad (4)$$

It is important to emphasize at this point that although the neural and light images are distinctly different entities, they share a common domain. In other words, both kinds of image are functions of  $x$ , the visual direction. Thus, implicit in our reasoning is that when the  $j^{\text{th}}$  output neuron responds at level  $r_j$  it is sending a message to the brain that a certain amount of light has been received from visual direction  $x_j$ . Because of the regional specialization of the retina, the difference between visual directions of neighboring cells is small in the fovea and large in the periphery. Consequently, on a global level the neural image is spatially distorted causing the fovea to be highly magnified in comparison with the periphery. Such distortion is a prominent feature of the primate visual system which is accentuated further in higher visual centers of the brain<sup>7</sup>. This

complication will be avoided in the present analysis by assuming local uniformity of scale.

**2.2.1 Cross-correlation theorem** The result embodied in (4) can be placed on more familiar ground by temporarily ignoring the fact that the neural image is discrete. That is, consider substituting for  $x_j$  the continuous spatial variable  $u$ . Then equation (4) may be re-written as

$$r(u) = \int_{rf} w(x \square u) \cdot i(x) dx \quad (5)$$

which is recognized as a cross correlation integral<sup>8</sup>. Since this more general equation contains the previous discrete result as a special case, one could view the neural image as the cross-correlation of receptive field with the retinal image, evaluated at those specific visual directions  $x_j$  which are represented in the neural image. Using standard pentagram (H) notation for cross correlation, the result is

$$r(x_j) = w(x) \mathbf{H}i(x), \quad x_j = x_1, x_2, \dots \quad (6)$$

Replacing the awkward cross correlation operation with convolution yields

$$r(x_j) = w(\square x) \mathbf{k}i(x), \quad x_j = x_1, x_2, \dots \quad (7)$$

In other words, the discrete neural image is interpolated by the result of convolving the retinal image with the reflected weighting function of the neural receptive field.

**2.2.2 Neural point spread function** Based on this last result, we may immediately determine the neural image for a point source of stimulation on the retina by letting  $i(x)$  be an impulse ( $\delta$ ) function. Then the sifting property<sup>8</sup> of the impulse function yields

$$\begin{aligned} r(x_j) &= w(\square x) \mathbf{k}\delta(x) \\ &= w(\square x) \\ &= n(x_j), \quad x_j = x_1, x_2, \dots \end{aligned} \quad (8)$$

This fundamental relationship between the neural p.s.f. and the receptive field is summarized by the following theorem: *The neural image formed by a homogeneous array of linear neurons in response to a point of light on the retina is equal to their*

*common receptive field weighting function, reflected about the origin, and evaluated at those values of visual direction represented by the array. This image is called the neural point-spread function.* In what follows, the discrete neural p.s.f. will be designated  $n(x_j)$  while its continuous interpolation will be designated  $n(x)$ .

### 2.2.3 Neural image for arbitrary visual patterns

The concept of a neural p.s.f. introduced above is useful for specifying the output neural image for arbitrary input.

Combining (1), (7) and (8) we obtain

$$r(x_j) = o(x) \mathbf{k}p(x) \mathbf{k}n(x), \quad x_j = x_1, x_2, \dots \quad (9)$$

To put this equation in a form which emphasizes the role of neural sampling, we multiply the right-hand side of (9) by a continuous "array function"  $a(x)$ , which is defined to be unity at each visual direction  $x_j$  represented in the neural image but is zero for all the other "in-between" directions. By this maneuver the result of (9) takes on its final form

$$\begin{aligned} r(x) &= [o(x) \mathbf{k}p(x) \mathbf{k}n(x)] \cdot a(x) \\ \dots \quad a(x) &= \begin{cases} 1 & x = x_1, x_2, \dots \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (10)$$

This equation, which is illustrated graphically in Fig. 3, exposes the conceptual simplicity of signal processing by the common class of optic nerve fibers of the human eye. The final neural image is seen to be the result of three sequential stages of processing. First the object is optically filtered (by convolution with the optical p.s.f.) to produce a retinal image. Next the retinal image is neurally filtered (by convolution with the neural p.s.f.) to form a hypothetical, continuous neural image. Finally, the continuous neural image is point-sampled by the array of ganglion cells to produce a discrete neural image ready for transmission up the optic nerve to the brain. Notice the change of viewpoint embodied in equation (10). Initially the output stage of the retina was portrayed as an array of finite, overlapping receptive fields which simultaneously sample and filter the retinal image. Now this dual function is split into two distinct stages, filtering followed by sampling with an array of point samplers.

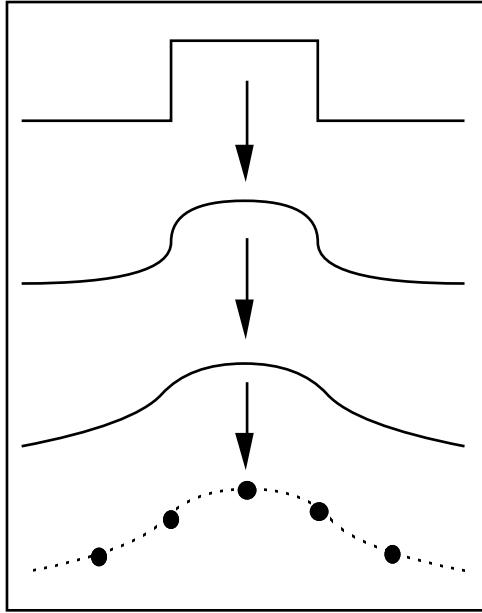


Fig. 3 Formation of the neural image

### 2.3 Frequency description of the neural image

A spectral description of the neural image may be obtained in a straightforward fashion by either of two methods. One is to find the Fourier transform of the right side of equation (10). Alternatively, if we make the simplifying assumption that the neural array is evenly spaced, then the discrete Fourier transform of (9) would suffice<sup>8</sup>. The latter method avoids the potential error of supposing that the spatial frequency spectrum may contain frequencies beyond that which can actually be supported by the discrete array. In either case, aliasing is a distinct possibility unless the combined filtering action of the optics and receptive fields eliminate all frequency content beyond the Nyquist limit. This issue is the topic of the following section.

## 3. FIDELITY OF THE NEURAL IMAGE

An important design criterion for man-made digital communications systems is to avoid the prospect of aliasing caused by undersampling of analog signals. Often it is preferable to discard high-frequency information by pre-filtering the signal rather than allow corruption of the remaining low-frequencies by aliasing. A cursory view of the front-end of the visual system, as diagrammed in Fig. 3, would seem to indicate that the design of the human eye also follows these sound engineering principles. Low-pass filtering by the eye's optical system and neural receptive fields

could both act as anti-aliasing filters. The question is, do they?

### 3.1 Aliasing in the neural image

If low-pass filtering by visual neurons is to be an effective anti-aliasing filter, then neural receptive fields must be relatively large compared to the spacing of the array. We can develop this idea quantitatively without detailed knowledge of the shape of the receptive field weighting function by employing Bracewell's equivalent bandwidth theorem<sup>8</sup>. This theorem, which is based on the central ordinate theorem, states that the product of equivalent width and equivalent bandwidth of a filter is unity. By definition, the equivalent width of a function is the width of the rectangle whose height is equal to the central ordinate and whose area is the same as that of the function. In the present context, the equivalent duration of the neural filter is the equivalent diameter  $d_x$  of the constituent receptive fields. To avoid aliasing requires that the spatial frequency bandwidth of the weighting function  $w(x)$  be less than the Nyquist limit as set by the characteristic spacing  $s$  of the array. If we adopt the equivalent bandwidth as a measure of the highest frequency passed to any significant extent by the filter, then by applying Bracewell's theorem we find that the requirement is for  $d_x > 2s$ . That is to say, aliasing will be avoided if the equivalent radius of the receptive field exceeds the spacing between fields.

A similar line of reasoning can be developed for two-dimensional receptive fields. Assuming radial symmetry of the fields, Bracewell's theorem states that the product of equivalent width  $d_x$  and equivalent bandwidth  $D_f$  is  $4/\sqrt{\pi}$ . Since the Nyquist requirement for anti-aliasing is that  $D_f < 1/2s$ , this means that we need  $d_x > 8s/\sqrt{\pi}$ . In other words, aliasing will be avoided if the equivalent radius of the receptive field exceeds  $4/\sqrt{\pi}$  times the spacing between fields, a slightly more stringent requirement than found above for the one-dimensional case.

### 3.2 Coverage factor rule

Neurophysiologists are well aware of the importance of aliasing for the fidelity of the visual system and so have devised a simple measure called the "coverage factor" to assess whether a given retinal architecture will permit aliasing<sup>9</sup>. Conceptually, the coverage factor of an array measures how much overlap is present. For a one-dimensional array, coverage equals the ratio of

width to spacing of the fields. For a two-dimensional, square array of radially-symmetric fields, coverage = area/spacing<sup>2</sup>. The utility of this measure here is that it encapsulates in a single parameter the importance of the ratio of receptive field size to receptive field spacing as a determinant of aliasing. The critical case derived above for a two-dimensional array is when  $d_x = 8s/\lambda$ , for which the coverage factor is  $16/\lambda$ . In other words, if the coverage is less than about 5 we can expect aliasing to result. Analysis of the cat retina<sup>9</sup> has confirmed that the  $\lambda$  class of ganglion cells is indeed well-designed to avoid aliasing but other, more sparsely populated classes have insufficient coverage to avoid aliasing. Similar analysis of the human retina is currently an area of active research.

### 3.3 Aliasing in human vision

Aliasing caused by neural undersampling can only occur if image frequencies beyond the Nyquist limit are present on the retina. It is technically possible to avoid low-pass filtering by the eye's optics with an interferometric visual stimulator. By using such a device it has been possible to demonstrate the existence of aliasing both foveally<sup>10</sup> and in the peripheral field<sup>11</sup>. This psychophysical evidence unequivocally shows that at least some optic nerve fibers are not protected from aliasing by neural filtering. When the eye's natural optical system is allowed to form the retinal image in the normal way, aliasing still occurs in the peripheral field but not foveally<sup>12</sup> as shown in Fig. 4. In the figure key, detection acuity means the highest frequency which is visible (as an alias) whereas resolution acuity means the highest frequency perceived veridically (i.e. does not alias). When compared to anatomical estimates of sampling density<sup>13</sup>, it is evident that the lowest frequencies aliased in the periphery are well below the Nyquist rate for cones but match reasonably well the Nyquist rate of retinal ganglion cells (RGC).

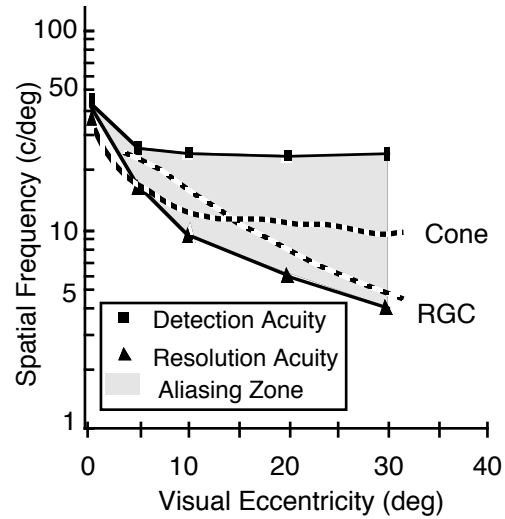


Fig. 4 Aliasing spectrum in human vision

The explanation of this state of affairs is that the eye's optical system does not pass spatial frequencies beyond about 60 c/deg, which is about the Nyquist limit for the foveal population of optic nerve fibers. Thus under normal viewing conditions, the fovea is protected from aliasing because of optical, rather than neural, filtering. In the periphery, however, such high quality optics are ineffective as an anti-aliasing filter. Evidently, neural filtering also is insufficient to prevent aliasing entirely, although it may substantially reduce the magnitude of aliased signals.

### 4. SIMULATIONS

The fact that a broad spectrum of frequencies are aliased over most of the visual field, spared by optical filtering only at the fovea, is probably unsettling to the seasoned design engineer. Indeed, vision scientists are asking the same question: why does the visual system permit such infidelity? To help gain some insight, two simulations of the peripheral neural image have been prepared. The first involved digitizing a photograph with the same sampling density (7 samples/deg) as used by the human retina at 30 deg in the peripheral field. The result is shown in Fig. 5. The sampling rate was 7 samples/deg, corresponding to a Nyquist rate of 3.5 c/deg, which is about the human resolution acuity indicated by Fig. 4. The original photograph was of a well known individual wearing a check jacket with a pattern finer than the 3.5 c/deg

Nyquist limit of the array of samplers. As a result, the checks are aliased into a coarse, zebra-like pattern. The middle line of text at the bottom of the photo corresponds to the optometrist's largest E (Snellen 20/200, a commonly accepted benchmark for legal blindness). The dominant frequency of these letters is 3 c/deg, just below the sampling rate. The text above is twice as large and the text below is half as large.

The second simulation was intended to show the consequences of anti-aliasing filtering of the same photo. The original photograph was optically defocused to the extent that the checks were no longer discernable, indicating that the high-frequency components of the image were effectively removed. This blurred picture was then sampled in the same way as before and the result is shown in Fig. 6.

Perhaps the most striking impression to emerge from this simulation is the devastating impact of anti-alias filtering on the informational content of the image. It would appear that the price which must be paid to avoid erroneous representation of fine patterns in the neural image is greatly outweighed by the loss of vital information available in the object. If this is true, then it could be argued that the reason the visual system permits the infidelity of aliasing is that it is much too costly to prevent it by pre-filtering.

Of course, aliasing could also be avoided by increasing the sampling density. This is probably not a viable option for the visual system, however, for at least two reasons. First, a massive increase in the number of nerve fibers, and thus the physical size of the optic nerve, would be required to sample the whole visual field at high density. Second, any increase in optic nerve capacity would require a corresponding increase of brain mass. Instead, the visual system has opted for compromise design based on an inhomogeneous retina in which sampling density is low, and aliasing tolerated, over most of the visual field. In one tiny region of retina sampling density is dramatically increased to match the bandwidth of the retinal image and so provide our most reliable and acute vision.

## 5. ACKNOWLEDGEMENTS

This research was supported by National Institutes of Health grant EY5109 and by AFOSR grant 870089 to the Indiana Institute for the Study of Human Capabilities. I thank D. Still, K. Haggerty

and J. Kubley for simulations of peripheral aliasing and technical assistance.

## 6. REFERENCES

1. W.N. Charman, "The retinal image in the human eye," *Prog. Retinal Res.* **2**, 1-50, 1983.
2. Still, D.L., Thibos, L.N. "Peripheral image quality is almost as good as central image quality," *Invest. Ophthalm. Vis. Sci.* **30** (suppl.), 52, 1989.
3. J.D. Gaskill, *Linear Systems, Fourier Transforms and Optics*, John Wiley & Sons, New York, 1978.
4. S.L. Polyak, *The Retina*, Univ. Chicago Press, Chicago, 1941.
5. W.R. Levick, L.N. Thibos, "Receptive fields of cat ganglion cells: classification and construction," *Prog. Retinal Res.* **2**, 267-319, 1983.
6. R.W. Rodieck, "The primate retina," *Comp. Primate Biol.* **4**, 203-278, 1988.
7. R.L. De Valois, K.K. De Valois, *Spatial Vision*, Oxford University Press, New York, 1988.
8. R.N. Bracewell, *The Fourier Transform and Its Applications*, McGraw-Hill, New York, 2<sup>nd</sup> ed., 1978.
9. A. Hughes, "New perspectives in retinal organization," *Prog. Retinal Res.* **4**, 243-313, 1985.
10. D.R. Williams, "Aliasing in human foveal vision," *Vision Res.* **25**, 195-205, 1985.
11. L.N. Thibos, D.J. Walsh, F.E. Cheney, "Vision beyond the resolution limit: aliasing in the periphery," *Vision Res.* **27**, 2193-2197, 1987.
12. Thibos, L.N. , Still, D.L., "What limits visual resolution in peripheral vision?" *Invest. Ophthalm. Vis. Sci.* **29** (suppl.), 138, 1988.
13. L.N. Thibos, F.E. Cheney, D.J. Walsh, "Retinal limits to the detection and resolution of gratings," *J. Opt. Soc. Amer.* **A4**, 1524-1529, 1987.