Visual Information Processing in Primate Cone Pathways—Part II: Experiments

Samir Shah and Martin D. Levine, *Fellow, IEEE*

***Abstract*—This paper presents the results of experiments performed to validate the *computer retina* model presented in a companion paper [15]. Experiments commonly performed by electrophysiologists on biological retinas are simulated and the *computer retina* outputs compared with published recordings of cells in biological (principally primate) retinas. Experiments with more complex stimuli further reveal how the *computer retina* enhances spatio-temporal contrast information and adapts to a wide range of illumination levels much like the primate retina.**

I. Introduction

ALTHOUGH commercially available machine vision sensors are beginning to approach the photoreceptor densities found in primate retinas, they are still outperformed by biological visual systems in terms of dynamic range, and strategies of information processing employed at the sensor level. This has sparked recent interest in studying biological retinas in order to learn more about their information processing strategies with the hope of using this knowledge to design better machine vision sensors. Toward that goal, a model of information processing in the cone pathways of the primate retina has been developed and is presented in detail in a companion paper [15]. The *computer retina* model is based closely on biology but has been significantly simplified to allow for the possibility of implementation in silicon.

In this paper, the response of the *computer retina* model to various simple visual stimuli is presented and compared to responses in biological retinas in order to validate the model. In addition, experiments with more complex images are used to illustrate what the outputs of the primate retina might look like if such bulk recordings were possible. The aim is to show that the retinal model developed in the previous paper [15], although quite simple, responds in a manner similar to the primate retinal cone system. In particular, the model qualitatively accounts for the gain and sensitivity changes as well as visual acuity changes found in the retina with illumination levels varying over seven orders of magnitude.

II. A Model Summary

Fig. 1 shows the flow of information in the *computer retina* model described in [15]. The information processing strategies

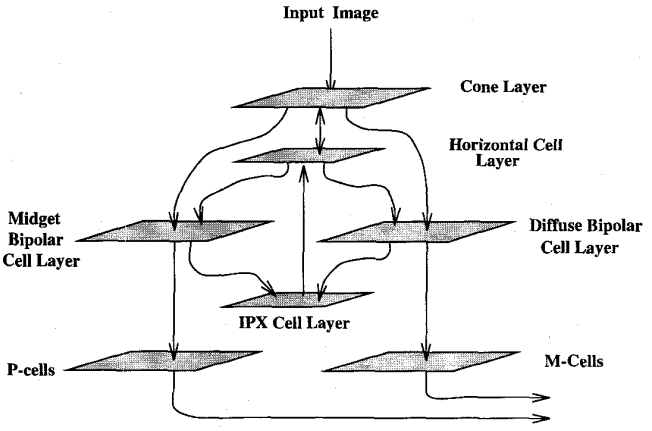

Fig. 1. Information flow in retinal model: The diagram shows the flow of information between each layer of "cells" in the retinal model described in [15].

utilized by the retina and incorporated into the model may be summarized as follows:

- Nonuniform sampling at the photoreceptor level, modeled by a log-polar mapping, significantly reduces the amount of data that needs to be processed at subsequent levels.
- The cone layer is responsible for converting the illumination intensity information into signals used by the rest of the retinal network. A measure of the local neighborhood spatiotemporal ambient intensity is used to locally set the sensitivity of each photoreceptor element and allows the *computer retina* to adapt to scenes containing a large dynamic range of retinal illumination intensities. This feature models the effects of pigment bleaching and horizontal cell feedback on the adaptation of cone sensitivity.
- Diffusion is used to model the coupling or spread of signals between neighboring retinal cells and to generate Gaussian-weighted receptive fields. The diffusion operation requires information only from immediately neighboring cells, and thus, is relatively simple to implement.
- The horizontal cell layer generates a spatial and temporal lowpass version of the visual signal. Horizontal cells have much larger receptive fields and integration time constants than cones.
- A difference operation between the *computer retina* cone and horizontal cell signals, each with different spatiotem-

Manuscript received January 21, 1994; revised February 26, 1995. This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

The authors are with the Centre for Intelligent Machines, McGill University, Montréal, Québec, Canada H2V 4A2 (e-mail: samir@cim.mcgill.ca; levine@cim.mcgill.ca).

Publisher Item Identifier S 1083-4419(96)02301-1.

poral characteristics, leads to the extraction and enhancement of spatiotemporal contrast in the visual signal at the bipolar and ganglion cell levels.

- Two output channels emerge at the bipolar and ganglion cell levels with visual information coded at different scales and with distinct spatiotemporal properties. The smaller receptive field sizes and greater number of midget bipolar and P-type ganglion cells make them sensitive to high spatial contrast. In comparison, diffuse bipolar and M-type ganglion cells have much larger receptive fields, are less numerous, and are better tuned to respond to temporal frequency modulation. The spatiotemporal properties of the *computer retina* P-cells and M-cells appear qualitatively similar to those of cells in Macaque monkey retinas.
- The diffusivity or coupling between cells in the cone and horizontal cell layers in the *computer retina* is changed as a function of the local neighborhood illumination level. Altering the cell coupling in such a manner affects the size of receptive fields at the ganglion cell level and serves as an adaptation mechanism that attempts to maintain both high contrast sensitivity and high visual acuity.
- Cells in the interplexiform (IPX) layer react vigorously in regions of high spatiotemporal contrast in the input image and reduce the RF size of horizontal cells. This adaptation mechanism reduces the relative size of the surround portion of bipolar and ganglion cell RF's in regions of high spatiotemporal contrast (spatial edges, flickering lights, etc.) which improves the localization of spatiotemporal edges.

III. Methods

A. Model Implementation

The model of the retinal cone system as described in the companion paper [15] was implemented in C and executed on Sun4 Sparc machines.[1] The *computer retina* program takes in as input a sequence of images which are then processed by the model. Several sets of outputs are produced, one for each of the retinal neuron layers. The output data may be either displayed as images or analyzed using a mathematical package such as MATLAB.[2]

The *computer retina* program simulates both the temporal and spatial characteristics of retinal processing. The spatial domain is discretized into pixels with each pixel width assumed to represent the width and spacing of foveal cones in the human retina. The temporal domain is also discretized. The sequence of images read in as visual input is assumed to have a 3 ms interval between each successive frame. This value was chosen to simplify the implementation of a 3 ms delay (τ_{delay}) of the horizontal cell signal to the diffuse bipolar terminal required in the model [13]. For some experiments, thousands of image frames had to be processed by the *computer retina* because these experiments required the input image stimuli to be presented for several seconds of real-time. Typical execution times on a sequential SPARC-10 machine for 100×100 images (without using log-polar sampling) were about 5 s/frame.

The *computer retina* program allows flexibility in controlling the features of the model that are enabled or disabled. For example, the use of log-polar sampling can be enabled and disabled at will. For most of the experiments described in this paper, log-polar (foveated) sampling was disabled. With this feature disabled, there was no variation of receptive field (RF) sizes with eccentricity and the fovea encompassed the entire field of view. This simplified analysis and comparison of the model cell outputs to responses of biological retinal cells, most of which were measured in the fovea. Other *computer retina* program directives allow enabling/disabling of features such as the effects of horizontal feedback on cones, and adaptation of cone and horizontal cell receptive field sizes.

B. Input Stimuli

In order to test the adaptive aspects of the model, images with a large dynamic range of intensities (up to 7 orders of magnitude) are required. Commercial camera systems and frame grabbers are inadequate, as most only use 8b to represent the output intensity level. To represent 6–7 orders of intensity magnitude, at least 20b are required. Most image databases are also restricted to 8b intensity dynamic ranges, and thus, are not directly suitable.

To circumvent these problems, most of the images used in these experiments were synthetically generated. Two main types of experiments were conducted and the stimuli required for each type were generated in different ways.

To compare the outputs of this model with published data on biological retinas, some of the experiments commonly performed by electrophysiologists were replicated. Typically, these experiments measure the outputs of a few retinal cells in response to very simple stimuli such as flashed backgrounds, spots of light, step edges, and temporally modulated sine gratings. Such simple images, with dynamic ranges of up to 12 log units of intensity, were easily generated in software and used as input to the *computer retina*.

In addition to experiments with simple stimuli, the outputs produced by the model in response to more complex stimuli (more realistic images) are also shown. For these tests, some images were generated using a very simple ray tracing image generation package.[3] This package used 16 b to quantize the image intensities allowing four orders of magnitude of image intensities to be represented in a single image. To further increase the input illumination dynamic range, the intensity values in synthetic images and images from image databases were exponentially scaled by various factors to crudely approximate the effects of increased illumination levels.

C. Units

All intensity values in this paper are expressed in photopic trolands (td) as this is the unit most commonly used in the biological literature. Photopic trolands are a measure

[1] Sun4 Sparc is a registered trademark of Sun Microsystems, Inc.

[2] MATLAB is a registered trademark of The MathWorks, Inc.

[3] The ray tracing package (*MakeImage*) used in the experiments was developed by Hiro Yamamoto of Canon, Inc.

of retinal illuminance and are defined as the product of luminance (cd/cm^2) and the pupil diameter in (mm). Time is expressed in seconds (s) or milliseconds (ms). Cell output values are normalized so that a maximal output is one. Spatial frequencies are represented in cycles/pixel or cycles/degree when compared with the primate data. Temporal frequencies are expressed in Hertz (Hz).

D. Model Evaluation

The results of four basic types of experiments are presented in this paper. These are:

- Flashed background experiments
- Experiments with sinusoidal gratings
- Experiment with step edge stimuli
- Experiments with more complex real and synthetic images.

The first two types replicate some of the typical experiments performed by electrophysiologists on biological retinas, and thus, allow comparison of the *computer retina* outputs to biological retinal cell responses. The last two types of experiments show the behavior of the *computer retina* in response to more complex stimuli and illustrate the data reduction achieved by using log-polar sampling.

Wherever possible, comparisons of the *computer retina* outputs to those in biological retinas are made using published data on the primate retina. Although there have been a fair number of recordings of the outputs of primate cones and ganglion cells[4] [1], [3], [5], [12], [14], the recording of most other retinal cells is extremely difficult to study in mammals because of their small size. What is available are recordings of these cells in other animals. Fortunately, for many visual properties, interspecies comparisons are possible when the underlying behavior and structure of the retina are similar [4], [17]. Many of the visual properties of interest here can be qualitatively compared with fish, cat, and monkey retinas. Therefore, comparisons of the model outputs are made with data from animals whose visual systems are believed to be similar to primates, when primate data was not available.

In making comparisons of the *computer retina* responses with biological retinal cells, the differences in the polarity of some retinal cell outputs as compared to the model responses should be noted. In all biological vision systems the output potential of photoreceptors (rods and cones) and horizontal cells hyperpolarize (become more negative) in response to increased light stimulation. In contrast, our model cone and horizontal cell outputs are designed to become more positive to signal increased output levels. In biological retinas, bipolar cells are of two types: those that hyperpolarize in response to increased stimulation in their receptive field centers, and those that depolarize (become more positive). The bipolar cells outputs in the *computer retina* are positive when the receptive field center is brighter than the surround and are negative when the surround is brighter than the center.

[4]Many of the ganglion cell characteristics are actually inferred from recordings of cells in the Lateral Geniculate Nucleus (LGN), where the bulk of the ganglion cell outputs terminate.

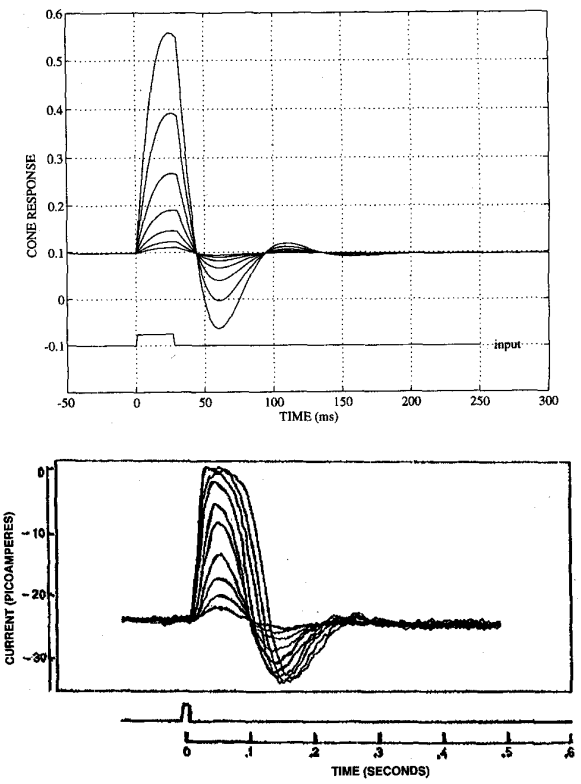

Fig. 2. Cone response to short full-field flashes: The top figure shows the model cone response to 30 ms full-field flashes at flash intensities of 10, 20, 40, 80, 160, 320, and 640 trolands added onto a background of 100 trolands. The bottom figure shows the membrane current in a monkey cone cell as the flash intensities were progressively doubled. From Schnapf [14]. Copyright © 1987 Scientific American. All rights reserved.

IV. Flashed Background Experiments

We conducted flashed background experiments similar to those conducted by Normann, Werblin, and Schnapf in order to show the temporal responses of our *computer retina* cells to full-field flash inputs [11], [14], [19]. In these experiments, a uniform background intensity (which fills the entire visual field) was presented to the retina interspersed with short intervals in which a background of a different intensity was "flashed" or substituted. The response of various cells are recorded for the entirety of the experiment. Four types of experiments were performed with flashed background stimuli. These were used to determine the following:

- The cone "impulse" responses to short dim flashes
- The effect of horizontal cell feedback on cone output
- The response of various cells to long full-field flashes
- The intensity-response curves for various retinal cells.

The following sections describe the results of the above tests.

A. Cone "Impulse" Response

Fig. 2 shows the superimposed responses of a model cone cell when the *computer retina* was subjected to short 30 ms "flashes" of increasing intensity (top figure) and the responses of a monkey cone in a similar test (bottom figure) [14]. The

monkey cone response is characterized by a peak due to the flash followed by an overshoot when returning to the resting potential. This overshoot is due to the effects of feedback from horizontal cells to the cone. The model response also exhibits this overshoot with the horizontal cell feedback enabled. The model cone output reaches a peak 30 ms after presentation of the " flash" while a monkey cone typically takes 50 ms. This discrepancy indicates that the cone time constant ($\tau_{cone} = 10$ ms) used in the model may be too small. However, the chosen cone and horizontal cell time constants are consistent with the values used by Richter [13]. In all other aspects, the model's cone response appears to behave similarly to the monkey cone cell.

B. Effects of Horizontal Cell Feedback on Cone Output

To isolate the responses of cones alone, without the horizontal cell feedback, electrophysiologists often treat the retina with an aspartate solution. The presence of aspartate blocks the horizontal cell signal from influencing the cone output [6], [11]. In the *computer retina* program, it is of course simple to turn off the horizontal cell feedback.

Fig. 3(a) shows the model cone response to a flashed background test both with and without horizontal cell feedback. Data on the effects of horizontal cell feedback on primate cone responses are not available. However, for comparison, data for the *gecko* "cone-like" rod photoreceptors is shown in Fig. 3(b).

Although the temporal characteristics of the gecko rod are much slower than a primate cone, and the input stimuli used are slightly different, the effects of horizontal cell feedback on the gecko photoreceptor responses are comparable to that observed in the model cone response. It can be seen that when horizontal cell feedback is present, the small delay in the feedback signal with respect to the cone serves to significantly increase the peak amplitude and sharpen the cone temporal response to a "flash." The initial cone peak response is quickly diminished by the effects of network feedback from the horizontal cells. This network adaptation mechanism acts to quickly return the cone potential to the near the middle of its response range, thereby preventing cone saturation.

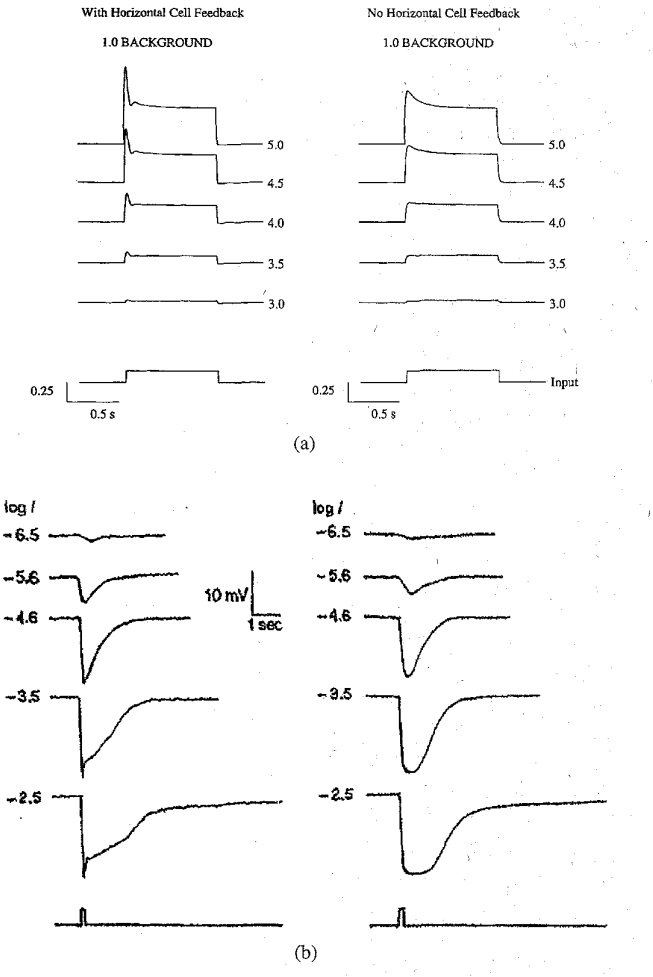

Fig. 3. Effect of horizontal cell feedback on cone response: (a) shows the normalized model cone responses to 900 ms flashes of varying intensity with horizontal cell feedback enabled (left) and disabled (right). The intensity of the flash is indicated beside each curve in log troland units. The background intensity was 1.0 log troland units. It can be seen that horizontal cell feedback causes the cone potential to very quickly return from a peak response to one nearer to the middle of its response range, thus "sharpening" the cone transient response. These responses may be compared with those obtained from recordings of *gecko* "cone-like" rod photoreceptors in (b) with horizontal feedback present (left) and absent (right) (from Kleinschmidt [8] with permission by the Rockefeller University Press). The *gecko* rods hyperpolarize in response to increased illumination stimulation, and thus, their responses are inverted with respect to the model cone responses. Although the temporal characteristics of *gecko* photoreceptors are much slower than the primate cones and the input stimulus is different from that used in (a), the effect of the horizontal cell feedback in sharpening the photoreceptor transient response is comparable.

C. Cell Responses to Full-Field Flashes

Shown next are the outputs of the model cone, horizontal, and bipolar cells in response to flashed background experiments performed at various ambient illumination levels. These outputs are compared to those recorded in fish retinas by Normann [11] and Werblin [19].

Fig. 4 shows the model cone outputs (a) and the *Necturus* (a type of fish) cone recordings (b) when exposed to long (900 ms–2 s) test flashes of various intensity at two different background intensities [11]. The *Necturus* cone recordings were made in an aspartate treated retina, and thus, were isolated from the effects of horizontal cell feedback [11]. To allow for a comparison, a separate experiment recorded the cone outputs in the absence of horizontal cell feedback. The model cone responses are similar to those recorded in the *Necturus*. The main difference is that the model cones have a much sharper decay from the initial peak response to a flash, and thus, reach a stable plateau much faster (within 250 ms) than the *Necturus* cones (about 1 s). This decay to a stable plateau is a result of cone temporal adaptation mechanisms which adapt the cone sensitivity to the new ambient intensity. The observed difference is thus due to the much shorter time constant chosen for cone temporal adaptation in the *computer retina* (modeling pigment bleaching in the primate system) than exists in the *Necturus* retina.

Figs. 5 and 6 show the data for horizontal and midget bipolar cells, respectively, for flashed background experiments

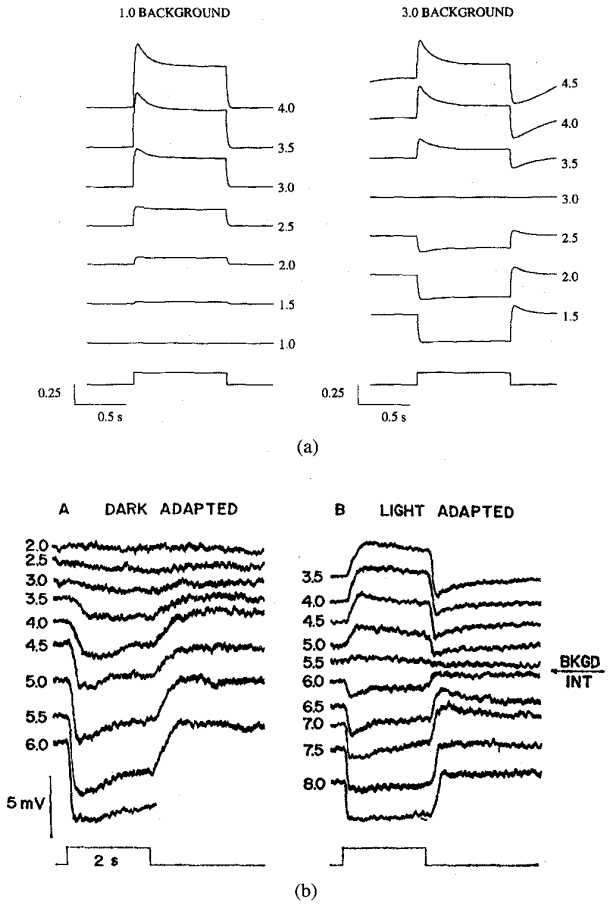

Fig. 4. Cone temporal response to full-field flashes: Fig. 4(a) shows the temporal response of model cone cells when 900 ms test flashes were substituted for the background every 8 s at two different log background intensities (1.0 and 3.0 log td) indicated at the top of each set of curves. The intensity of the substituted flash is shown to the right of each trace in units of log trolands. The initial cone response to a "flash" slowly decays to a new resting potential as the cone adapts to the new illumination level. The model cone cell response curves are similar to cone responses obtained by Normann in the aspartate treated *Necturus* retina shown in Fig. 4(b) [11] (with permission of Rockefeller University Press © 1974). Note that the cone potential in the *Necturus* retina hyperpolarizes (becomes more negative) in response to brighter illumination levels, and thus, is opposite in polarity to the model cone responses. Increasing flash intensities are toward the bottom of the figure for the *Necturus* and the opposite of those for the model cone curves.

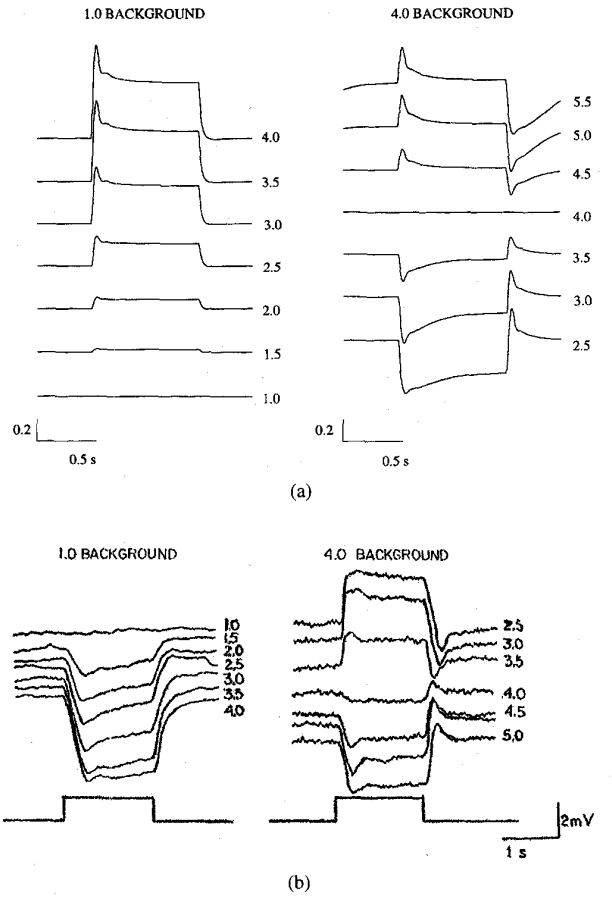

Fig. 5. Horizontal cell response to full-field flashes: Fig. 5(a) shows the temporal response of model horizontal cells to 900 ms test flashes substituted for the background every 8 s at two separate background intensities (shown at the top in units of log trolands). The intensity of the substituted flash is indicated to the right of each curve in log trolands as in the cone experiment. Fig. 5(b) shows horizontal cell recordings obtained by Werblin in the mudpuppy retina in a similar experiment [19] (with permission of Rockefeller University Press ©1974). Note that mudpuppy horizontal cells hyperpolarize in response to increased stimulation while *computer retina* horizontal cell outputs increase in response to increased stimulation. As a result, the polarity of the model horizontal cell response is the opposite of that of a mudpuppy horizontal cell.

as performed on the cones. In this case, horizontal cell feedback to the cones was enabled to allow comparison of model cell outputs against those of nonaspartate treated mudpuppy retinas [19] The model cell outputs are comparable to those of the mudpuppy but have a much sharper transient responses due to the shorter time constants used in the model cells. Although primate horizontal and bipolar cell recordings are not available, Richter estimates that the time constant in mammals may be anywhere from 5 to 10 times shorter than in fish [13].

D. Intensity-Response Curves

Fig. 7 shows the intensity-response curves for the model cone, horizontal, midget bipolar and diffuse bipolar cells. These curves were obtained by plotting the peak amplitudes of the cell output vs. the log intensity of the flashes as obtained from the flashed background experiments described in the previous section. A separate curve was obtained at each of the background intensities shown. The model cone and horizontal cell plots show that these cells have a dynamic range of about 2.5–3.0 log units centered around any given background intensity. Flashed intensities outside this range either elicit negligible cell response or a saturated response. Increasing the background intensity shifts the operating point of cone and horizontal cells and shifts their entire IR-curve to the right without a change in shape. Both the midget and diffuse bipolar cells also exhibit similar behavior. The model midget bipolar cell has a dynamic range of approximately 1.5 log units while the diffuse bipolar cell has a very narrow 1.0 log unit response range around a given background intensity. For background intensities close to the "cone threshold" ($<10^2$ td),

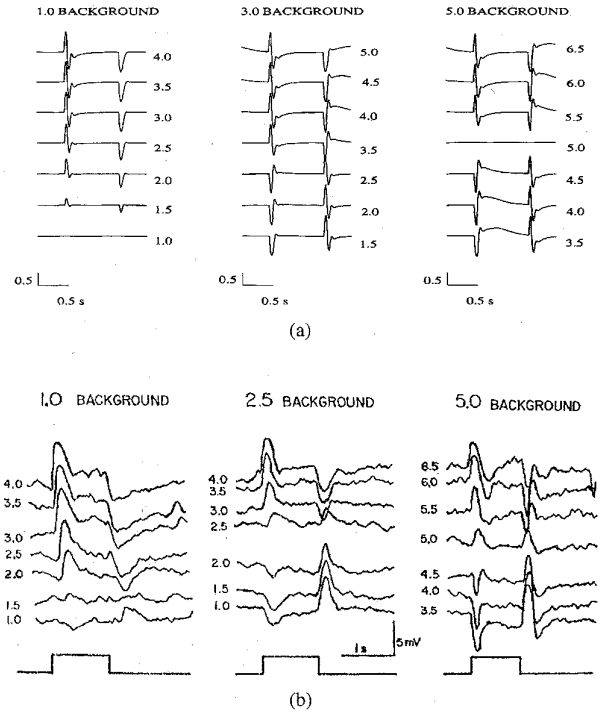

Fig. 6. Midget bipolar cell response to full-field flashes: Fig. 6(a) shows the time course of model midget bipolar cells to 900 ms test flashes substituted for the background every 8 s at three separate background intensities (indicated at the top in units of log trolands). The intensity of the substituted flash is shown to the right of each curve in log trolands. Fig. 6(b) shows recordings from a depolarizing bipolar cell in the mudpuppy retina obtained by Werblin in a similar experiment [19] (with permission of Rockefeller University Press © 1974).

most of the response range of bipolar cells is used to signal intensities brighter than the background (positive contrast). It is only at much brighter intensities that the model bipolar cells respond equally well to negative contrast as well as positive contrast. Fig. 7 shows the corresponding intensity-response curves obtained by Normann [11] and Werblin [19] in the *Necturus* (cone) and mudpuppy (horizontal and bipolar cell) retinas. Note that the slopes of the fish cone and horizontal cell curves are the negative of those for the corresponding *computer retina* cells because of the hyperpolarizing nature of these biological cells. The mudpuppy horizontal cells appear to have a dynamic range of slightly over 3.0 log units while the bipolar responds to only a 1.5 log unit range of intensities around the prevailing background intensity.

The flashed background experiments show that the outputs of the *computer retina* approximate to the first order the temporal nature of responses found in biological retinas. Where differences exist, most may be attributed to a difference in the time constants between fish retinas and those used in the model implementation which were based on estimates for a mammalian retina.

V. Sinusoidal Grating Stimulus Experiments

Sinusoidal grating stimuli are often used in visual psychophysics and electrophysiology since they allow researchers to characterize the response of the visual system over a wide spectrum of spatial and temporal frequencies. Typically, a sinusoidal grating stimulus consists of a pattern in which the intensity varies sinusoidally along one of the spatial coordinates. In addition, the stimulus may be modulated temporally so that the intensity also varies sinusoidally with time. The parameters that specify the sinusoidal grating pattern include its spatial frequency k (cycles/pixel), its temporal frequency of modulation ω (Hz), background intensity I_b and the contrast or depth of modulation c of the grating ($0 < c < 1.00$).

In electrophysiological experiments on primates, the sinusoidal grating stimuli are presented while the outputs of cells in the parvocellular and magnocellular layers of the lateral geniculate nucleus (LGN) are recorded. The parvocellular and magnocellular layers are where the outputs of retinal P and M ganglion cells terminate, respectively. The LGN cells have very simple center-surround receptive-field structures, just like the P and M retinal ganglion cells, and are thus believed to have properties, in most respects, similar to the ganglion cells which provide their input.

In these experiments, the recorded monkey P-cell and M-cell data may be expressed in any number of ways. The most common measures used are contrast sensitivity, gain, and contrast gain. Contrast sensitivity for individual cells is typically defined as the reciprocal of the contrast required in the input stimulus to produce a criterion response (a minimum level of output that must be achieved in order to consider the input stimuli distinguishable from background noise) in the cell output. Gain and contrast gain in the Fourier domain are defined here to be consistent with Shapely [16] and Purpura [12]:

$$\text{Gain} = \frac{\text{Amplitude of cell response}}{(\text{Stimulus contrast})(\text{Mean retinal illumination})} \quad (1)$$

$$\text{Contrast gain} = \frac{(\text{Amplitude of cell response})}{(\text{Stimulus contrast})}. \quad (2)$$

Only measurements of contrast gain or absolute gain are given for the *computer retina* ganglion cell responses. Determination of the contrast sensitivity was omitted as it requires a lengthy iterative procedure to determine the smallest input contrast required to elicit a criterion output response. Measuring it would have increased the execution time of the experiments by at least an order of magnitude. In the experiments presented here, it is assumed that the contrast gain and contrast sensitivity curves are similar enough to make strictly qualitative comparisons between the computer retina and biological retinas. Such an assumption is valid if one assumes P-cells and M-cells behave linearly over most of their dynamic range [9].

Three types of experiments were conducted using sinusoidal grating stimuli to characterize the following features of the *computer retina*:

- Changes in spatial frequency sensitivity caused by illumination level,
- Changes in temporal frequency gain as a function of illumination level,

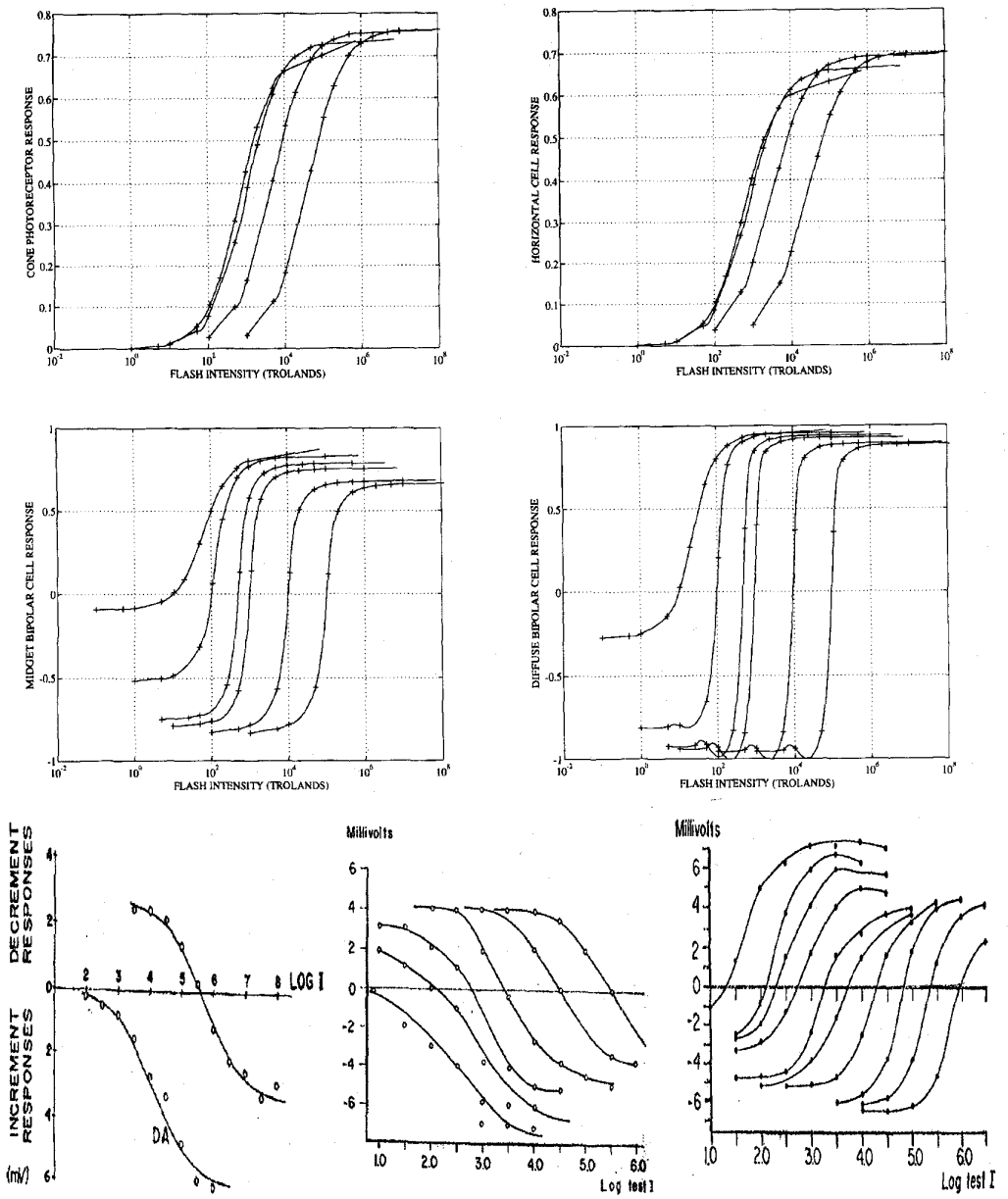

Fig. 7. Intensity-response curves: The top four plots in the figure show intensity-response curves for the model cone, horizontal, midget bipolar, and diffuse bipolar cells as obtained from flashed background experiments. The bottom plots show (from left to right), the corresponding data for similar experiments in the *Necturus* (cone) [11] and mudpuppy (horizontal and bipolar cells) [19] (with permission of Rockefeller University Press © 1974). Note that the slopes of the fish cone and horizontal cell curves are the negative of those for the corresponding *computer retina* cells because of the hyperpolarizing nature of these biological retinal cells.

- Changes in contrast gain as a function of input contrast and temporal frequency.

The results from these experiments are presented next.

A. Spatial Frequency Sensitivity

In this experiment, the spatial frequency sensitivity of the model P-cells and M-cells was determined at several illumination levels. At each background illumination level, the peak responses of the P-cell and M-cell outputs were recorded as the spatial frequency of the grating was slowly varied over three orders of magnitude. The input stimulus was presented with a temporal modulation of 2 Hz while the grating contrast was kept fixed at $c = 0.20$ throughout the experiment. The ganglion cell output levels were divided by the contrast c to obtain a measure of contrast gain.

The Fig. 8(a) and (b) shows the contrast gain of model P-cells and M-cells as a function of spatial frequency at several illumination levels. The spatial frequencies have all been converted from cycles/pixel to cycles/degree to aid comparison with *Macaque* monkey data from Derrington [5] shown at the bottom of Fig. 8(c) and (d). The figure illustrates several

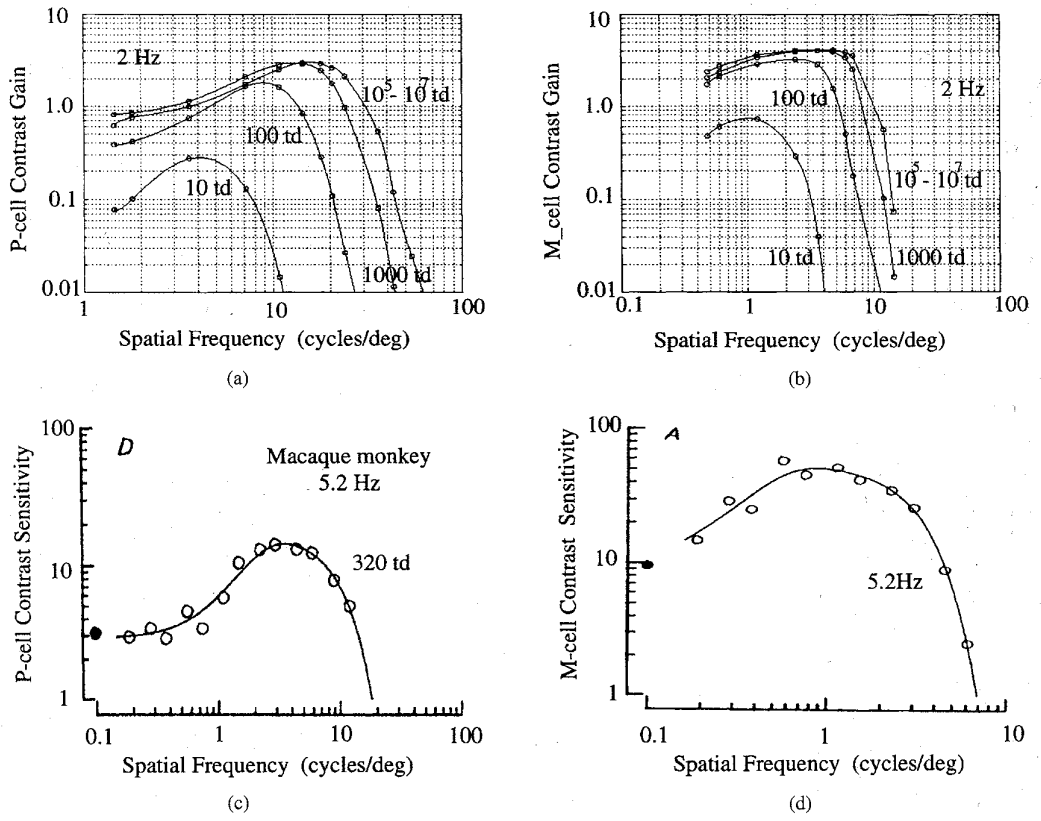

Fig. 8. Ganglion cell spatial frequency tuning. The top figures show the contrast gain of the *computer retina* P-cells (a) and M-cells (b) as a function of the sinusoidal grating spatial frequency at several illumination levels (10, 100, 1000, 10^5, and 10^7 trolands). The temporal frequency of modulation was maintained at 2 Hz and the grating contrast was fixed at $c = 0.20$ throughout the experiment. The spatial frequencies have been scaled from cycles/pixel to cycles/degree to allow comparison with primate data. Comparable data on the typical contrast sensitivity of individual *Macaque* monkey parvocellular (P) and magnocellular (M) LGN cells is shown in the bottom fiures (from Derrington [5] with permission of The Physiological Society © 1984). The grating stimuli used in the *Macaque* monkey experiment (c) was modulated at a temporal frequency of 5.2 Hz at a background illumination of approximately 320 td.

aspects of the behavior of the *computer retina* P-cells and M-cells which may be compared with known facts on the primate retina.

- The frequency response of primate P-cells and M-cells is spatially bandpass at photopic illumination levels [21], [17].
- Increasing illumination levels are accompanied by increases in visual acuity, and thus, an ability to resolve higher sinusoidal grating frequencies [18], [20].
- At low illumination levels, an increase in illumination level is accompanied by an increase in contrast sensitivity (contrast gain) [21].
- At still much higher illumination levels, the contrast sensitivity (contrast gain) as well as spatial acuity reach a plateau and become insensitive to any further increases in background illumination level [21].
- The M-cell system in the primate retina has a higher contrast gain than the P-cell system [1], [17].
- P-cells are able to respond to much higher spatial frequencies (40–60 cycles/degree) than M-cells (10–20 cycles/degree) as a result of their much smaller receptive fields [9].

All of the above properties are qualitatively exhibited by our *computer retina*. As the illumination levels increase, the P-cell and M-cell contrast gain increases and the outputs reach a peak at much higher spatial frequencies of modulation. The M-cell outputs exhibit higher contrast gain (larger peak outputs) and are tuned to lower spatial frequencies than P-cells at all illumination levels. The model P-cells have a peak sensitivity to spatial frequencies of 12–15 cycles/degree and are able to signal frequencies up to 50–60 cycles/degree and is consistent with primate data. The model P-cells and M-cells also exhibit a slight spatial bandpass characteristic.

It should be noted that the shape of the spatial frequency response is dependent upon the temporal frequency of modulation as well. At very low temporal frequencies, both P and M cell responses exhibit a strong spatial bandpass characteristic. At moderate or high temporal frequencies of modulation, the spatial bandpass characteristic becomes much less pronounced and eventually becomes a lowpass characteristic. This effect is also observed in the primate and cat retinas [5], [7].

B. Temporal Frequency Sensitivity

Fig. 9 (top) shows the gain of the *computer retina* P and M ganglion cells as a function of temporal frequency when stimulated by sinusoidal gratings. For this experiment, the spatial frequency of the sinusoidal grating was held fixed at three cycles/degree (0.0208 cycles/pixel), and the grating

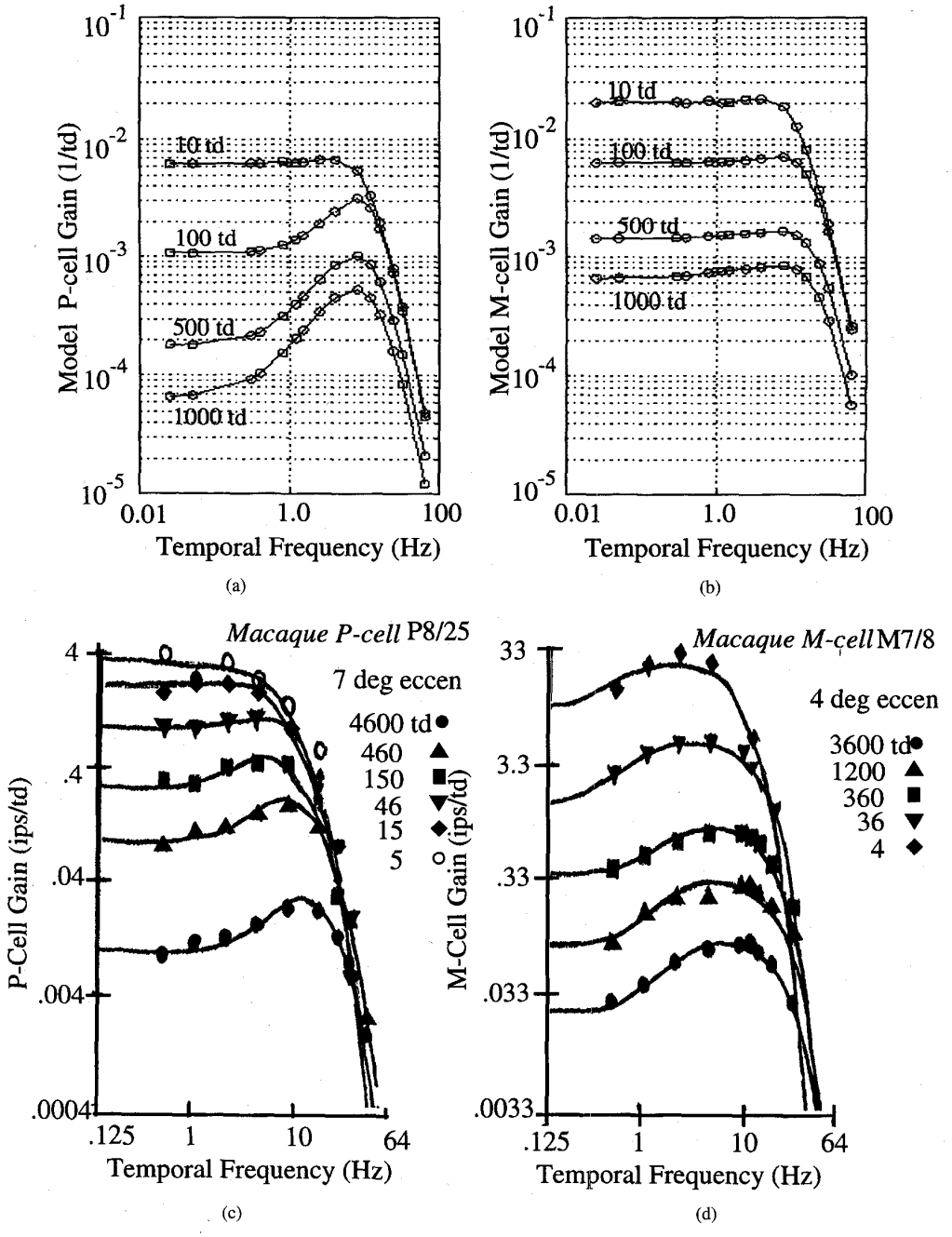

Fig. 9. Ganglion cell temporal frequency gain: Response curves in the top two figures shows the gain of foveal model P-cells (a) and M-cells (b) as a function of sinusoidal grating temporal frequency at several background illumination levels (indicated on the curves). A grating of fixed contrast ($c = 0.20$) and spatial frequency 3 cycles/degree was used throughout the experiment. The gain was computed as the mean ganglion cell output level divided by the product of the stimulus contrast c and the background illumination intensity. The data may be compared with recordings of P-cells (c) and M-cells (d) made in the *Macaque* monkey LGN by Purpura, © 1990/Cambridge University Press. In Purpura's data, a grating of spatial frequency 3 cycles/degree and a grating contrast of 64% was used for the P-cell experiments while gratings of 1.6 cycles/degree and contrasts between 4% and 25% were used for the M-cell experiments.

contrast was set to $c = 0.30$ throughout the experiment. The grating temporal frequency was varied slowly and held at each frequency for 4 s during which time the mean P-cell and M-cell outputs were recorded and divided by the mean retinal illumination level and the contrast c to obtain the gain values. This experiment was repeated for several illumination levels to examine the effects of illumination level on the temporal frequency gain. The model response may be compared to data from recordings in the *Macaque* monkey LGN made by Purpura [12] shown in bottom of Fig. 9.

The general temporal frequency properties of the *computer retina* P-cells and M-cells may be summarized as follows. Fig. 9 shows that the M-cell system responds with a much larger gain to temporal stimulation and is thus able to respond better than the P-cell system at both low and high temporal frequencies. This is in agreement with generally cited prop-

erties of primate retinal P-cells and M-cells [9]. The gain of the *computer retina* M-cells is about five—eight times larger than that for the P-cells and is comparable to the factor of six—seven difference found in monkey cells. The peak P-cell temporal frequency is around 8–9 Hz and within the range of values (8–12) Hz found in primates [12]. In addition, the temporal frequency gain of both M-cells and P-cells is inversely proportional to the mean retinal illumination level at low temporal frequencies, and thus, is consistent with Weber's law. At higher temporal frequencies, the gain of the system depends increasingly on the temporal frequency and time constants of various cells in the retina and becomes relatively insensitive to the illumination level.

The *computer retina* model's temporal frequency gain compares reasonably well with the Macaque monkey data in most respects. The main difference is in the shift of the peak temporal frequency (the temporal frequency which gives a peak ganglion cell output) as a function of illumination level. In the monkey data, as the illumination levels rise, the peak temporal frequency also increases fractionally. This seems to indicate that in addition to modulation of receptive field sizes, the dynamic properties (time constants of various cells) of the retina also adapt to the illumination level. In our model, for the sake of maintaining simplicity, the time constants for all the cells are fixed with respect to illumination level, and thus, the *computer retina* peak temporal frequency does not change with illumination level. The monkey data is accounted for by Purpura by adapting the time constants of several lowpass temporal filters (through which the input signal passes) as a function of the illumination level [12].

C. Contrast Gain

In order to measure the effect of input stimulus contrast on the contrast gain of P-cells and M-cells, an additional experiment was performed and the responses compared to data on a similar experiment performed on monkey retinal ganglion cells. In this experiment, the outputs of model P-cells and M-cells were recorded as a function of temporal frequency at several grating contrasts. The contrast gain was computed as defined in (2). The parameters held fixed in this experiment were the retinal illumination level (1000 trolands) and the grating spatial frequency (0.009 cycles/pixel or 1.3 cycles/degree). Fig. 10 (top) shows the contrast gain of model M-cells and P-cells as a function of temporal frequency for several grating contrasts. Comparable data on retinal ganglion M-cells and P-cells in Macaque monkeys are reported by Benardete [1] and are shown at the bottom of Fig. 10.

This experiment shows that the P-cell contrast gain is largely unaffected by the stimulus grating contrast while the M-cell contrast gain diminishes noticeably with increased contrast input at the cell's peak temporal frequency. Such behavior is consistent with the notion that P-cells, which have a smaller contrast gain than M-cells, respond linearly with respect to contrast [9]. M-cells, however, are easily saturated by moderate to high contrasts in the input signal due to their high contrast gain, and thus, exhibit diminished contrast gain with increasing stimulus contrast. Again, the *computer retina* outputs are similar to the first order of approximation to the behavior observed in monkey retinal ganglion cells.

VI. Step Edge Responses

In order to illustrate the adaptation features of the model when confronted with changing illumination levels, the *computer retina's* behavior in the neighborhood of a spatial step edges was explored at several background illumination levels. The spatial step edge stimuli consisted of images in which the right half (of the image) was twice as bright in intensity as the left half at any given background intensity. The experiment was conducted so that a uniform intensity background was presented for 2 s in order to allow the retinal responses adapt to this intensity. This was followed by the presentation of the step edge stimuli for 0.5 s with the left half of the image intensity matching the adapting background intensity. The outputs of all cells along a line crossing perpendicular to the step edge (in the x-direction) were recorded. This procedure was repeated at several background intensities. The background intensities ranged from 100 trolands to 10^6 trolands in 1.0 log unit steps. The responses of the cone, horizontal, and midget bipolar, and interplexiform cells, as well as additional data from this experiment, are shown in Fig. 11.

The plots (a)–(h) in Fig. 11 illustrate several aspects about the model retina:

- Plot (a) shows the illumination intensity profile across the step edge at pixel position $x = 20$. The right side of the image is twice as bright as the left half in each trial.
- Plot (b) shows the peak cone response approximately 40 ms after presentation of the step edge. For retinal illumination levels below 10^4 trolands (td), increasing background illumination levels result in a steady increase in the cone potential. However, for intensities above 10^4 td, mechanisms modeling the effects of pigment bleaching and horizontal cell feedback act to keep the cone potential near the middle of its operating range. The magnitude of the peak cone response to the step edge depends on the contrast of the step edge and very little on the illumination level.
- Plot (c) shows that the larger receptive fields of the horizontal cell layer result in a blurred version of the cone output.
- Plot (d) shows the peak midget bipolar cell output approximately 40 ms after presentation of the edge. Note that for decreasing illumination levels, the width of the step edge response becomes wider as cone and horizontal cell coupling increases. In all other aspects, the bipolar cell response is insensitive to global changes in illumination level and instead responds to the contrast in the visual signal.
- Plot (e) shows the steady-state output of the interplexiform cell layer in the model. Interplexiform cells receive their input from bipolar cells, and thus, their outputs are large in regions where the bipolar cell outputs are large (like near the step edge).
- Plot (f) shows how the degree of horizontal cell coupling (expressed by the width of its Gaussian shaped receptive

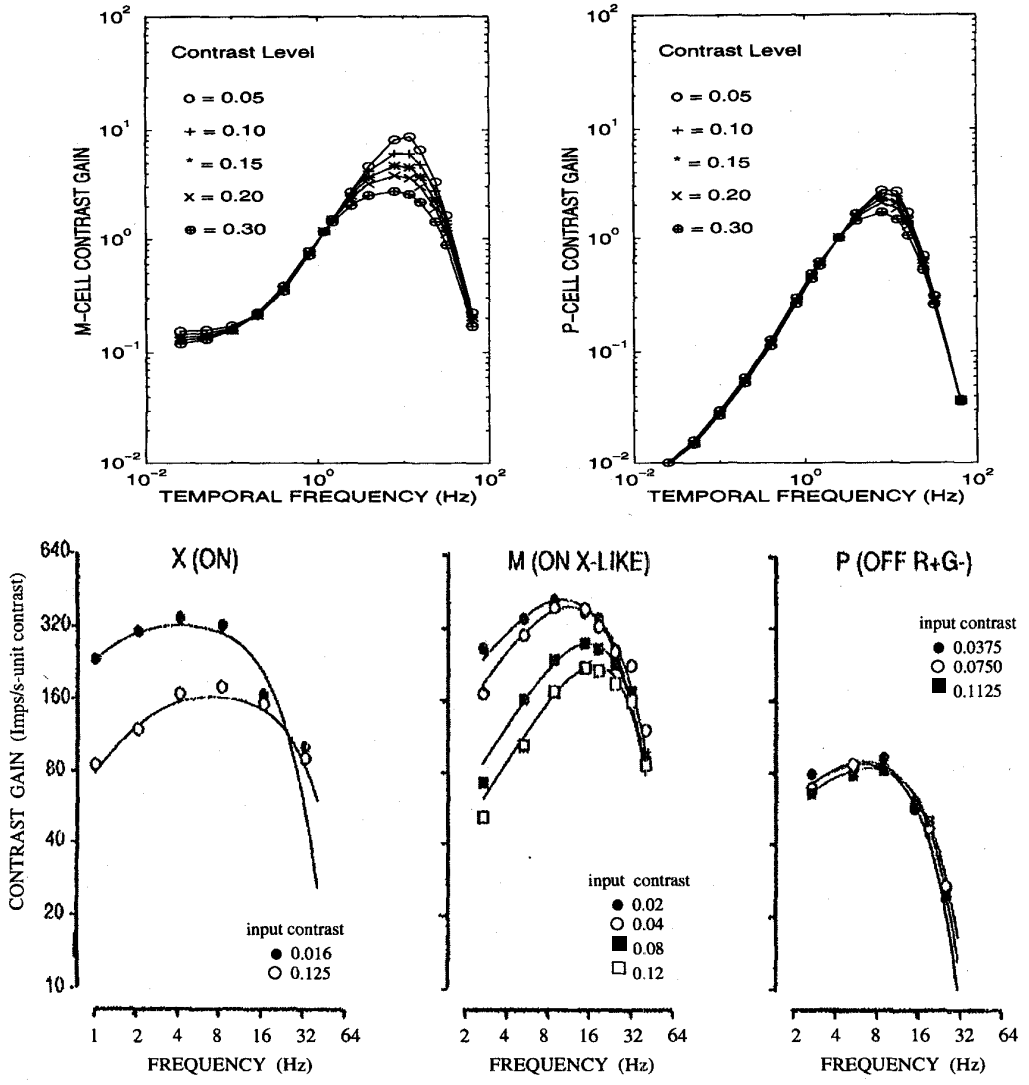

Fig. 10. P-cell and M-cell Contrast Gain: P-cell and M-cell contrast gain: The top figures show the contrast gain of model M-cells and P-cells as a function of temporal frequency at several grating contrasts. The tests were performed at a background illumination level of 10^3 td and a spatial frequency of 0.009 cycles/pixel (1.3 cycles/degree). The bottom figures show similar data for a cat X-cell, and a monkey (X-like) M-cell, and a monkey P-cell from Benardete [1], © 1992 Cambridge University Press. The M-cell is more sensitive to changes in contrast than the P-cell. Increased levels of contrast lead to diminished contrast gain in the M-cell system due to the nonlinear saturation function. This effect is much less pronounced in P-cells because of their smaller contrast gains.

field) is reduced by increasing levels of illumination level. Dopamine released by interplexiform cells in the neighborhood of the step edge further reduces the horizontal cell coupling in cells near the step edge.

- Plot (g) shows the cone and horizontal cell receptive field (RF) sizes (again expressed in the widths of the equivalent Gaussians describing their RF's) as function of background illumination level. At low illumination levels, cone and horizontal cell receptive fields expand in size in order to improve absolute contrast detection. At higher illumination levels, they shrink in order to improve acuity.
- Plot(h) shows the midget bipolar cell response 250 ms after presentation of the step edge. Note that, except in the vicinity of the edge, the bipolar cell response on the left side and right side of the image has now faded to zero. The figure also shows that the width of the step edge response gets narrower with increasing illumination levels and is a result of the shrinking receptive field sizes.

The step-edge experiment has illustrated some of the adaptive properties of the retinal model more clearly. The remaining experiments, in the next section, show the response of the *computer retina* to real images from image databases and a ray tracing package.

VII. Complex Images

The previous sections illustrated the response of the *computer retina* in response to very simple stimuli. We now show the model response to more complex images. The images used for this section are of two types. Some of the images were generated using a simple ray-tracing package which permitted intensity values to range over four orders of magnitude. Others

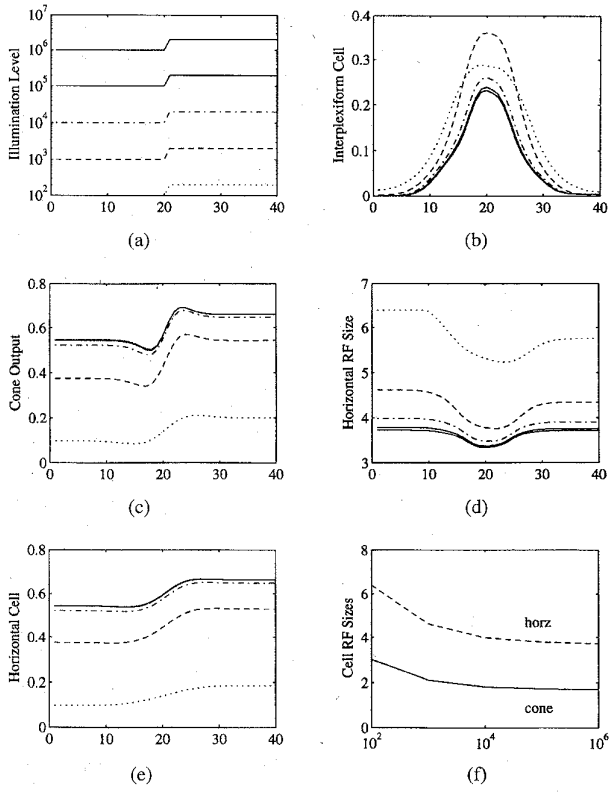

Fig. 11. Step edge responses: The figures show the results of the spatial step edge experiment: (a) the step edge intensity profile, (b)–(d) the peak cone, horizontal and midget bipolar cell outputs, (e) the steady-state interplexiform cell outputs, (f) the Gaussian width of the Horizontal cell receptive field (RF) along the step edge profile, (g) variation of cone and horizontal cell RF's with illumination level, and (h) the steady-state midget bipolar cell outputs. Explanations of these figures are found in the text.

were obtained from image databases and their intensities were exponentially scaled to expand the dynamic range of intensities in the input signal. Because of the considerable storage and computational requirements to generate temporally varying images, the images used in these experiments were all temporally stationary.

In order to display the outputs of various layers of cells on the computer screen and to make figures for this paper, the *computer retina* cell output ranges were scaled to fit intensity values from 0–255. For the bipolar and ganglion cells, which have outputs that may be positive or negative depending on the input contrast, the zero level of the outputs was set to a pixel intensity of 127, with the total output dynamic range covering intensities from 0–255.

A. Response to Large Dynamic Range Images

Fig. 12 shows the steady-state outputs of the *computer retina* to an image containing three objects (a complex block, a sphere, and a cylinder) on a table surface. The intensities in this generated image varied over three—four orders of magnitude.

Even though the intensities in the input image range over three—four orders of magnitude, in any given local neighborhood of the image, the cones adapt their sensitivity based on the local ambient intensity level. This adaptation mechanism tends to return the cone potential to near the middle of its range. As a result, the steady-state response of the cone layer to a stationary image is for the most part equal to the cone half-maximum response except in regions of sharp intensity gradient. The P-cell and M-cell steady-state responses are only sensitive to spatial edges in the image. All other features fade away in the steady-state response. This is similar to the "fixation blindness" observed in the primate retina where stabilized images on the retina fade away after a very short time [10]. Note that there is only one M-cell for every nine P-cells, and thus, the outputs of the M-cell layer sample the image at a much coarser resolution than the P-cell layer.

B. Foveated Sampling and Data Reduction

For all previous experiments in this paper, the changes in the processing of visual information associated with retinal eccentricity have been ignored. All comparisons made with biological data were from data in the fovea where the ratio of cone cells to ganglion cells is one-to-one. For these experiments, the fovea was assumed to cover the entire image. This greatly simplified the comparison and analysis of the experimental data as the effects of nonuniform sampling and data reduction that occur in the periphery could be ignored.

However the *computer retina* is also able to crudely approximate the effects of increasing receptive field sizes and coarser resolution sampling occurring in the periphery. The *computer retina* program makes use of log-polar mappings which separate the image data into two regions: the fovea and the periphery.[5]

In the fovea, there is a one-to-one mapping of the cone inputs to ganglion cell outputs. In the *computer retina* periphery, a log-polar sampling scheme, as described in [15], is used to model the nonuniform sampling occurring at the receptor level, as well as the convergence of information from the cones to the ganglion cell level.

The primate retina achieves tremendous data reduction through the use of a nonuniform sampling scheme that samples the image with increasingly coarser resolution with increasing retinal eccentricity. Further data reduction is realized by converging increasing numbers of cone inputs into ganglion cell outputs. The primate retina has just over a million ganglion cell outputs for about five million cone inputs, and thus, realizes a five fold reduction in the number of outputs to be processed at higher levels.

Fig. 13 illustrates the data reduction realized by the *computer retina* when using a log-polar transformation on the periphery data. The input image size is 512×512 (262 144 pixels). The foveated *computer retina* generates separate outputs for the fovea and periphery for both P-cells and M-cells. The foveal data is represented in the original Cartesian domain while the peripheral information is represented in the log-polar domain. In the log-polar domain images, increasing log eccentricity is in the horizontal direction to the right, while the angular position in the retina is represented along the vertical axis. We arbitrarily selected the fovea diameter for

[5]The computer program used to perform the log-polar mappings was developed by M. Bolduc, McGill University, as part of his Master's thesis.

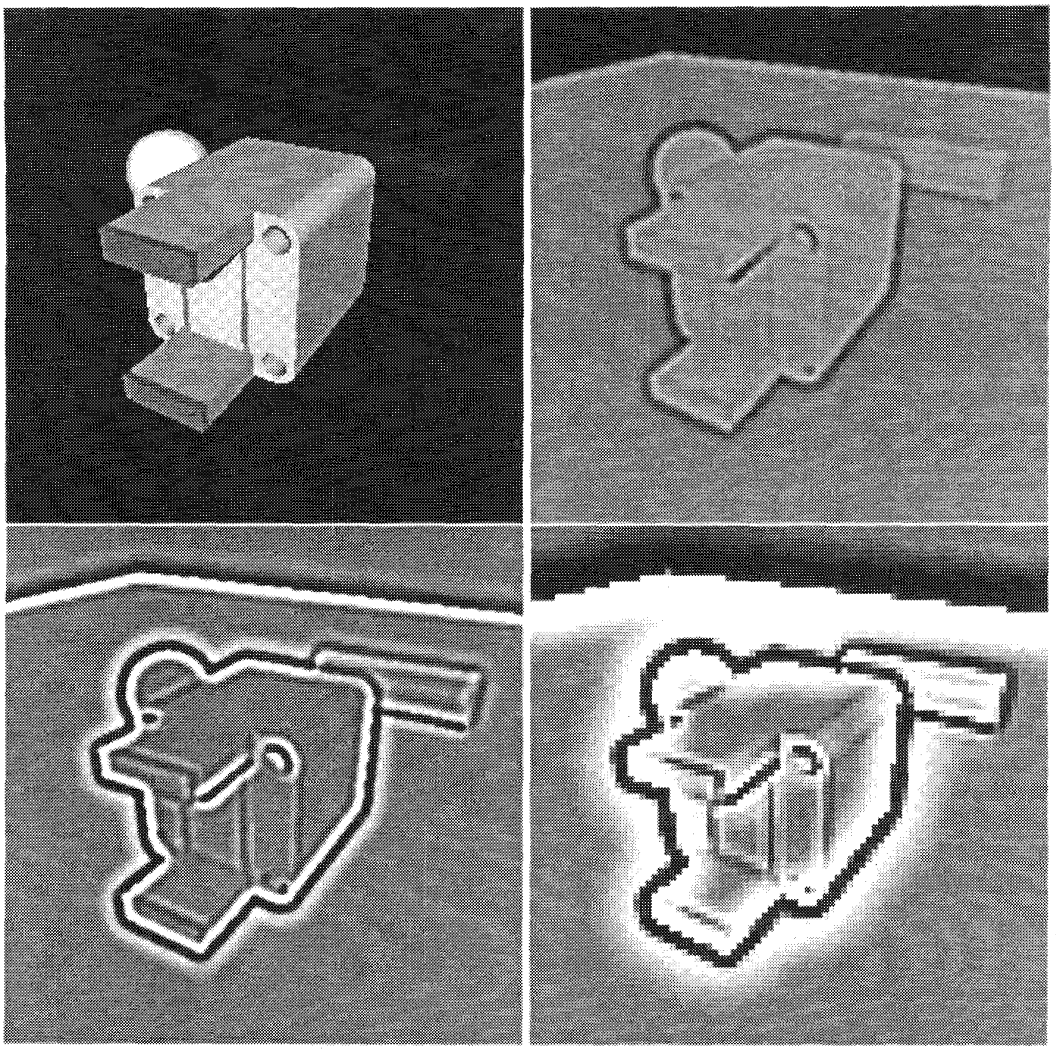

Fig. 12. Steady-state response to blocks image: The top left figure shows the input image with the intensities normalized to fit into a greyscale of 0–255. The actual intensities in the input stimulus ranged from $1 - 5 \times 10^4$ trolands. The very dim cylinder and table are difficult to see in original image. The top right figure shows the steady-state cone response to the image after 350 ms. The cones locally adapt their sensitivity so that both bright regions and dim regions in the image are visible. The P-cell and M-cell cell outputs are shown in the bottom left and bottom right figures, respectively. All outputs are normalized to fit a greyscale of 0–255.

the *computer retina* to be approximately 15% of the original image and required that the size of receptive fields on fovea and periphery boundaries match each other. The size of the outputs for each layer were as follows: P-cell fovea (79×79), M-cell fovea (26×26), P-cell periphery (121×305), M-cell periphery 40×102), for a total of 47 902 pixels, including a 15% overlap region[6] between the fovea and periphery data structures. In the above example, the *computer retina* achieves a six fold reduction in data required to be processed by all subsequent stages of visual processing by the use of a log-polar mapping scheme. The choice of the fovea and periphery sizes above was simply for convenience and does not necessarily reflect the real retina. However, the example illustrates the significant reduction in visual processing requirements that result from using a log-polar sampling scheme.

In Fig. 14, an inverse-mapping of the log-polar transformed data is provided for the convenience of the reader. The figure shows the foveal and peripheral information mapped back into the original retinal domain.

VIII. Summary

The experiments in this paper have illustrated the response of the *computer retina* for a variety of conditions. The experiments performed with sinusoidal grating stimuli and flashed background stimuli have shown, by direct comparison, that at least to a first order of approximation, the qualitative behavior of the *computer retina* cells is similar to that of fish and primate

[6] Near the fovea-periphery boundary the computation of the cell outputs requires information from both the fovea and the periphery. In order to decouple the computations in the fovea and the periphery, and thereby greatly reduce the computational complexity, the fovea and periphery data structures were extended by over 15% so that they overlapped each other. This allowed the computations in the fovea and the periphery to be performed independently of each other.

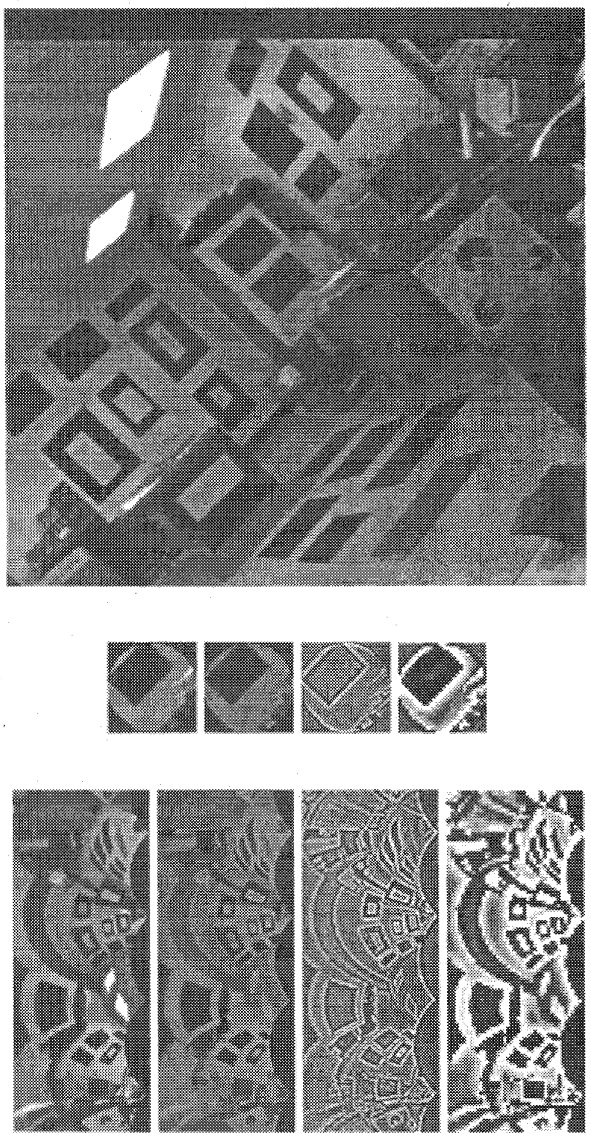

Fig. 13. Foveated Retina Outputs: The top figure shows the original image as presented to the *computer retina*. The middle row of figures shows the foveal representations of the input image, and the cone, horizontal cell, midget bipolar/P-cell, diffuse bipolar/M-cell layer outputs (from left to right). The foveal portion of the image is quite small (about 79 pixels in diameter) and is located at the center of the original image. The bottom row of figures shows the log-polar domain periphery representations in the same order as for the fovea. Note that the actual number of M-cell outputs is only one-ninth of the those for P-cells but the M-cell output image has been expanded nine-fold for easier comparison here.

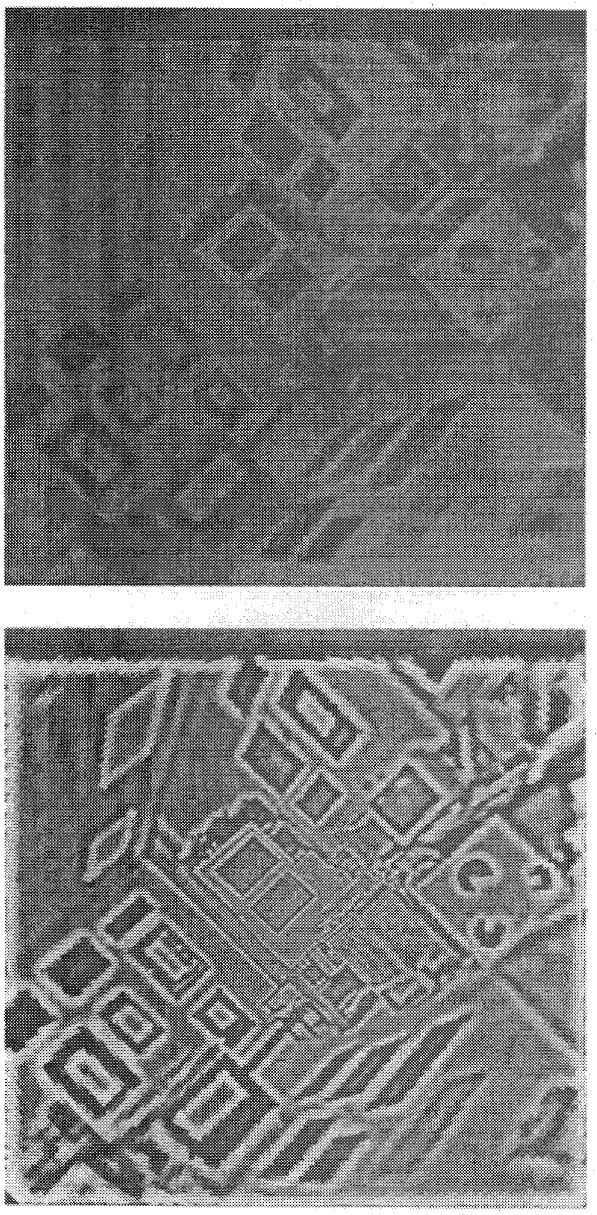

Fig. 14. Foveated Data Mapped back to the Retinal Domain Domain: For the reader's convenience, the figure shows data from the fovea and periphery data structures mapped back to the retinal domain for the cone (top) and P-cell layers (bottom). The inverse mapping illustrates how log-polar sampling models the foveation in the retina. Although the foveal region is sampled at high resolution, features are sampled more coarsely with increasing eccentricity in the periphery. At the output to the retina, P-cell and M-cell outputs enhance spatial and temporal contrast in the original image.

retinas. Both the temporal and spatial frequency responses of the *computer retina* cells match the data from experiments on fish and primate retinas, and thus, validates our model [1], [5], [11], [12], [14], [19]. The step edge experiment was presented to reveal how the model responds to a spatial step edge over a large range of background intensity levels. This experiment also illustrated the manner in which horizontal cell and cone photoreceptor coupling varies with the illumination level and the local contrast in the scene. The final experiments, with more realistic images, illustrated that our retinal model accounts for not only the adaptational aspects of the retinal function but also mimics the data reduction found in the human retina.

Although there are a few discrepancies in the behavior of the *computer retina* as compared to that of biological retinas, the essential information processing believed to be performed by the retinal cone system is also exhibited by the model. The model exhibits adaptation to wide range of illumination levels while maintaining a high contrast output at two different scales

and with two different spatio-temporal properties. It would be interesting to see this model implemented in silicon in order to gain further insight into the workings of biological visual systems and also to develop better machine vision sensors.

ACKNOWLEDGMENT

The authors would like to thank the CIAR and PRECARN for their support.

REFERENCES

[1] E. Benardete, E. Kaplan, and B. Knight, "Contrast gain control in the primate retina: P cells are not X-like, some M cells are," *Visual Neurosci.*, vol. 8, pp. 483–486, 1992.
[2] M. Bolduc, "A foveated sensor for robotic vision," Master's thesis, McGill University, 1994.
[3] R. Boynton and D. Whitten, "Visual adaptation in monkey cones: Recordings of late receptor potentials," *Science*, vol. 170, pp. 1423–1426, 1970.
[4] M. Crawford, R. Anderson, R. Blake, G. Jacobs, and C. Neumeyer, "Interspecies comparisons in the understanding of human visual perception," *Visual Perception: The Neurophysiological Foundations,* L. Spillman and J. Werner, Eds. New York: Academic, 1990, pp. 23–52.
[5] A. Derrington and P. Lennie, "Spatial and temporal contrast sensitivities of neurones in the lateral geniculate nucleus of Macaque," *J. Physiol.*, vol. 357, pp. 219–240, 1984.
[6] J. Dowling, *The Retina: An Approachable Part of the Brain.* Cambridge, MA: Belknap/Harvard Univ. Press, 1987.
[7] D. Fleet, P. Hallett, and A. Jepson, "Spatiotemporal inseparability in early visual processing," *Biolog. Cybern.*, vol. 52, pp. 153–164, 1985.
[8] J. Kleinshmidt and J. Dowling, "Intracellular recordings from gecko photoreceptors during light and dark adaptation," *J. Gen. Physiol.*, vol. 66, pp. 617–648, 1975.
[9] P. Lennie, C. Trevarthen, D. V. Essen, and H. Wassle, "Parallel processing of visual information," *Visual Perception: The Neurophysiological Foundations,* L. Spillman and J. Werner, Eds. New York, NY: Academic, 1990, pp. 103–128.
[10] M. Mahowald, "Silicon retina with adaptive photoreceptors," in *SPIE/SPSE Symp. Electronic Science and Technology: From Neurons to Chips,* Orlando, FL, Apr. 1991, pp. 52–58.
[11] R. Normann and F. Werblin, "Control of retinal sensitivity I: Light and dark adaptation of vertebrate rods and cones," *J. Gen. Physiol.*, vol. 63, pp. 37–61, 1974.
[12] K. Purpura, D. Tranchina, and E. Kaplan, "Light adaptation in the primate retina: Analysis of changes in gain and dynamics of monkey retinal ganglion cells," *Visual Neurosci.*, vol. 4, pp. 75–93, 1990.
[13] J. Richter and S. Ullman, "A model for the temporal organization of X and Y-type receptive fields in the primate retina," *Biolog. Cybern.*, vol. 43, pp. 127–145, 1982.
[14] J. Schnapf and D. Baylor, "How photoreceptor cells respond to light," *Scientific American,* vol. 156, pp. 40–47, 1987.
[15] S. Shah and M. Levine, "Visual information processing in primate cone pathways: Part I, A Model," *IEEE Trans. Syst., Man, Cybern. B*, this issue, pp. 259-274.
[16] R. Shapely and C. Enroth-Cugell, "Visual adaptation and retinal gain controls," *Progress in Retinal Research,* N. Osborne and G. Chader, Eds. Oxford: Pergamon, 1984, pp. 263–346.
[17] R. Shapely and V. Perry, "Cat and monkey retinal ganglion cells and their visual functional roles," *Trends Neurosci.*, vol. 9, pp. 229–235, 1986.
[18] J. Thomas, "Spatial resolution and spatial interaction," in *Handbook of Perception,* E. Cartertette and M. Friedman, Eds. New York: Academic, 1979, vol. 5, pp. 233–264.
[19] F. Werblin, "Control of retinal sensitivity II: Lateral interactions at the outer plexiform layer," *J. Gen. Physiol.*, vol. 63, pp. 62–87, 1974.
[20] G. Westheimer, "Visual acuity and spatial modulation thresholds," *Handbook of Sensory Physiology: Visual Psychophysics,* D. Jameson and L. Hurvich, Eds. Berlin: Springer-Verlag, 1972, pp. 171–187.
[21] J. Yellott, "Photon noise and constant-volume operators," *J. Opt. Soc. Amer. A*, vol. 4, pp. 2418–2446, 1987.

Samir Shah, for a photograph and biography, see this TRANSACTIONS, p. 274.

Martin D. Levine (S'59–M'66–SM'74–F'88) for a photograph and biography, see this TRANSACTIONS, p. 274.