A cellular and molecular model of response kinetics and adaptation in primate cones and horizontal cells

Hans van Hateren

Department of Neurobiophysics, University of Groningen, Groningen, The Netherlands

A model for the sensitivity regulation in the primate outer retina is developed and validated using horizontal cell measurements from the literature. The main conclusion is that the phototransduction of the cones is the key factor regulating sensitivity. The model consists of a nonlinearity cascaded with three feedback control loops. The nonlinearity is caused by the hydrolysis of cGMP by activated phosphodiesterase. The first feedback loop is divisive, with calcium regulating the photocurrent in the cone outer segment. The second feedback loop is also divisive, with voltage-sensitive channels regulating the membrane voltage of the cone inner segment. The final feedback loop is subtractive, where the membrane voltage of the horizontal cell is subtracted from that of the cone before the cone drives the horizontal and bipolar cells. The model describes adequately the major characteristics of the horizontal cell responses to wide field, spectrally white stimuli. In particular, it shows (1) sensitivity and bandwidth control as a function of background intensity; (2) the major nonlinearities observed in the horizontal cells; and (3) the transition from linear responses toward contrast constancy (Weber's law) for background illuminances ranging from 1-1000 td.

Keywords: horizontal cells, macaque, cone, sensory transduction, light adaptation, computational model

Introduction

The light intensity of visual stimuli varies enormously in natural environments. Visual systems generally cope with these variations by regulating sensitivity to light at an early stage of visual processing. In the visual system of the macaque, Smith, Pokorny, Lee, and Dacey (2001) showed that such a regulation could be measured at the level of the horizontal cells in the outer retina. These cells are directly postsynaptic to the cones. They are thought to provide feedback to the cone pedicles, and thus regulate the temporal, spatial, and spectral properties of the output of the cones to the bipolar cells.

A fully adequate model of sensitivity regulation in the outer retina is lacking, whereas it is clearly important for modeling visual processing upstream in the retina and beyond. It should necessarily form the first stage of any model of the visual system that could function well in natural, outdoor-lighting conditions. The present study, therefore, aims to develop such a model, using the extensive set of horizontal cell measurements recently published (Lee, Dacey, Smith, & Pokorny 2003; Smith et al., 2001). It will be shown that the sensitivity regulation observed in horizontal cells with wide field, spectrally white stimuli, is fully consistent with the regulation expected from the known processes of the phototransduction machinery of the cones.

Model

Both phototransduction and the neural circuit of cone and horizontal cell have been extensively investigated, and much is known about the physiological processes that determine their dynamics (for reviews of phototransduction, see Arshavsky, Lamb, & Pugh 2002; Burns & Baylor, 2001; Fain, Matthews, Cornwall, & Koutalos 2001; Pugh & Lamb, 2000; for a review of horizontal cells, see Kamermans & Spekreijse, 1999). Below I will represent a series of these processes, first as differential equations describing their dynamics, and subsequently as a system model consisting entirely of first-order low-pass filters and static nonlinearities. Because many readers may not be fully familiar with all of these fields, I will present the key topics in a relatively self-contained manner. The variables in the model have different physical dimensions, ranging from trolands through ion currents and concentrations to electrical currents and voltages. To keep the equations as concise as possible, I will not explicitly include dimensional conversion constants, and I will allow gains that are not dimensionless. Wherever the dimensions in an equation appear not to match, the reader should assume an implicit conversion constant of unit dimension.

Low-pass filters and static nonlinearities

In the sections below I will show that the (coupled) differential equations describing the system can be translated into a system model with only two types of components: first-order low-pass filters and static nonlinearities. The fact that the behavior of these components is straightforward and well understood greatly simplifies understanding and analysis of the system dynamics. I will therefore first introduce these basic components.

The standard equation describing a low-pass filter with time constant $\tau_y$, transforming an input signal $x(t)$ into an output signal $y(t)$, is

$$\tau_y \frac{dy}{dt} = x - y$$  \hspace{1cm} (1)
with the dot denoting time differentiation:
\[ \dot{y} = \frac{dy}{dt}. \]  

(2)

Figure 1A shows how this filter is depicted in the models below. The solution of Equation 1 is
\[ y(t) = \int_{-\infty}^{t} h(t-t_1) x(t_1) dt_1, \]

which is the convolution of the input function \( x(t) \) with the impulse response of the filter,
\[ h(t) = \frac{1}{\tau_y} e^{-t/\tau_y} \text{ for } t \geq 0, \quad h(t) = 0 \text{ for } t < 0. \]

(4)

The low-frequency (DC) gain of this filter equals 1, thus low frequencies are not changed by the filter.

An alternative form of a low-pass filter that often occurs when dealing with chemical reactions and membrane processes is given by a variant of Equation 1:
\[ \tau_y \dot{y} = \tau_y x - y \quad \text{or} \quad \dot{y} = x - y/\tau_y. \]

(5)

This equation is associated with an impulse response
\[ h_i(t) = e^{-t/\tau_y} \text{ for } t \geq 0, \quad h_i(t) = 0 \text{ for } t < 0. \]

(6)

The low-frequency gain now equals \( \tau_x \). An example of such a process, where gain and time constant are covarying, is the voltage response to a current injected into a cell: When the membrane resistance is lowered, both the membrane time constant and the gain (voltage response to a given current) decrease. Below, such processes are described as the cascade of a separate gain \( \tau_x \) and a standard low-pass filter as in Figure 1A.

The second component used in the models below is the static (memory-less) nonlinearity depicted in Figure 1B: It gives, for each point in time, the output \( y(t) \) as a function \( f(x(t)) \) of the input \( x(t) \). The function \( f(\cdot) \) itself is not a function of time, only of the present value of its input \( x \). A special case is the function \( f(x) = g x \), with \( g \) a constant describing a (linear) gain.

**Outer segment: Phosphodiesterase activation and cGMP hydrolysis**

For the equations describing the phototransduction cascade I will keep the notation as short as possible by mostly using single letter symbols for the variables and constants involved. Furthermore, I will keep the number of variables, such as scaling constants, to a bare minimum, and include only those that affect the calculations when fitting to the measurements on horizontal cells. Cascaded scaling factors or gains will therefore be merged into a single factor or gain wherever possible. Many of the equations are based on modeling work by Calvert, Govardovskii, Arshavsky, and Makino (2002), Detwiler, Ramanathan, Sengupta, and Shraiman (2000), Hamer (2000), Koutalos and Yau (1996), Pugh and Lamb (2000), Schnapf, Nunn, Meister, and Baylor (1990), and in particular Nikonov, Engheta, and Pugh (1998), and Nikonov, Lamb, and Pugh (2000). These studies also give extensive references to the original experimental studies that have led to the understanding of the dynamics of the various parts of the system.

Figure 2A shows on the left the reaction scheme leading from the absorption of light \( I \) by the visual pigment \( R \) via a G-protein \( G \) to the production of activated phosphodiesterase \( E^* \). In the first step, the visual pigment absorbs light, and is converted into an active form \( R^* \). The active form is removed with a rate constant \( k_R \). It is assumed here that the light intensity is sufficient low such that the concentration of \( R \) is not significantly affected (new \( R \) is produced fast enough, and there is no significant bleaching of the pigment). The reaction is then described by
\[ \dot{R}^* = A_R I - k_R R^*, \]

(7)
with $I$ the retinal illuminance in trolands, and $A_R$ a scaling constant. This can be rewritten as

$$\tau_R \dot{R}^* = \tau_R A_R I - R^*, \quad \text{with } \tau_R = 1/k_R, \quad (8)$$

which can be recognized as a low-pass filter with time constant $\tau_R$ and gain $\tau_R A_R$.

In the second step, $R^*$ activates a G-protein into its active form, $G^*$, which subsequently forms an active complex $E^*$ with phosphodiesterase (PDE). Although the latter reaction can be modeled as a separate low-pass filter, its time constant is assumed to be so short that it can be neglected (Lamb & Pugh, 1992). It is again assumed that light intensities are such that neither $G$ nor PDE is significantly reduced. Therefore, the production of $E^*$ is described by

$$\dot{E}^* = A_E R^* - k_E E^*, \quad \text{or}$$

$$\tau_E \dot{E}^* = \tau_E A_E R^* - E^*, \quad \text{with } \tau_E = 1/k_E, \quad (9)$$

with $k_E$ the rate of $E^*$ inactivation. This is also a low-pass filter, with time constant $\tau_E$ and gain $\tau_E A_E$. For simplicity, this gain and the one in Equation 8 are merged into the gain, $k_β$, describing the activity of $E^*$ (see below). Equations 8 and 9 can therefore be represented by the system model on the right side of Figure 2A.

Figure 2B shows on the left the reaction scheme where $E^*$ reduces the concentration of cGMP (symbolized by $X$ below), which is itself produced at a rate $α$. The activity of PDE, $β$, is described by a constant dark activity (here $c_β$ often called $β_{dark}$ in the literature) plus a term depending on the amount of activated PDE:

$$\beta = c_β + k_β E^*, \quad (10)$$

with $k_β$ a constant. The concentration of cGMP is then described by

$$\dot{X} = α - β X . \quad (11)$$

At first sight, this looks like a complicated nonlinear equation, because the input, $β$, is multiplied by the output, $X$. However, it is possible to reformulate this equation as a static nonlinearity followed by a first-order low-pass filter with variable time constant. This can be seen be rewriting Equation 11 as

$$\tau_X \dot{X} = \frac{1}{β} X - X \quad \text{with } \tau_X = 1/β . \quad (12)$$

In this equation, the input $1/β$ is low-pass filtered into an output $X$, with a time constant $\tau_X$ depending on the input, and a gain $α$. Assuming for the moment that the gain $α$ is constant (it is in fact under calcium control; see below), Equations 10 and 12 can be represented by the system model on the right side of Figure 2B. The importance of $1/β$ for setting the sensitivity and time constant of the system has been discussed for rods (see Nikonov et al., 2000).

**Outer segment: The calcium feedback loop**

Figure 3 shows on the left the reaction scheme of the calcium feedback loop. An increase in cGMP concentration ($X$) will open more cyclic nucleotide-gated (CNG) ion channels in the plasma membrane of the outer segment, resulting in an inward current consisting partly of Ca$^{2+}$. The resulting increase in calcium concentration is counteracted by a Na$^+$/Ca$^{2+}$-K$^+$ exchanger, and possibly by calcium buffering. The increased calcium concentration reduces the rate with which guanylate cyclase (GC) synthesizes cGMP, and thus counteracts the initial rise of cGMP. In a secondary loop, more important in cones than in rods (Rebrik & Korenbrot, 2004), Ca$^{2+}$ decreases the sensitivity of the CNG channels to cGMP.

![Figure 3](image-url)

The Ca$^{2+}$ concentration, called $C$ here, is described by

$$\dot{C} = \eta I_{os} - k_C C, \quad \text{or}$$

$$\tau_C \dot{C} = \tau_C \eta I_{os} - C \quad \text{with } \tau_C = 1/k_C, \quad (13)$$

where $\eta$ is a scaling constant describing which proportion of the total photocurrent, $I_{os}$, is carried by calcium, and $k_C$ is the removal rate of calcium, presumably due to the exchanger. Thus $C$ is a low-pass filtered version of the photocurrent, with time constant $\tau_C$ and gain $\tau_C \eta$. For the calculations, this gain will be merged into the scaling constant $a_C$ determining $α$, as described below.

The reaction scheme contains two nonlinearities due to cooperativity. The opening of the CNG channels depends on the cGMP concentration as

$$I_{os} = X^{n_X}, \quad (14)$$

with $I_{os}$ the photocurrent into the outer segment, and $n_X$ typically 3 (Koutalos & Yau, 1996). Scaling is incorporated into the scaling constant $a_C$ below, and into the scaling constant $a_{os}$ determining the voltage response of the inner
segment (see next section). The activity with which GC produces cGMP is, for the values of $C$ relevant for the fits, described by

$$\alpha = \frac{1}{1 + (a_C C)^{n_C}},$$  \hspace{1cm} (15)

with $a_C$ a scaling constant, and $n_C$ typically taken to be 2 (Koutalos & Yau, 1996). The numerator of Equation 15 is scaled to 1 without loss of generality, because all scaling needed for the model fits is incorporated into the scaling constants $a_C$ and $a_{is}$.

Equations 12-15 are represented by the system model on the right side of Figure 3. To emphasize that it is a negative feedback loop and to enable an easy comparison with existing models, it is drawn as a divisive gain control: The input $Q$ (=1/β) is divided by $1/\alpha$, which is itself an expansive function of the calcium concentration, and therefore provides strong negative feedback. The broken line in the Figure represents the reduction of the channel sensitivity caused by calcium: either through the calcium concentration $C$ or, possibly more accurately, through the calcium current if the interactions are restricted to local domains. To keep the model as simple as possible, I am not modeling this interaction in detail. Instead, I assume it accounts for the fact that the standard value $n_X = 3$ cannot produce acceptable fits of the model to the measurements. It produces major deficiencies in the sensitivities as a function of contrast and mean intensity. Leaving $n_X$ as a free parameter in the fits leads to much lower values, typically $n_X = 1$. This low value of $n_X$ may be the result of the direct calcium feedback onto the CNG channels: Channel openings caused by an increase in $X$ are counteracted by the subsequent increase in calcium influx. It is well known in engineering that negative feedback loops can apparently linearize nonlinearities in the forward path. For example, a fast divisive feedback loop with a squaring operation in the feedback path will have output = input/output$^2$, or output = input$^{1/3}$, thus converting an actual $n_X = 3$ into an apparent $n_X = 1$. I will assume the value $n_X = 1$ below.

Similarly, I found that the fits, although not unreasonable for values $n_C = 2$ or 3, were significantly better for $n_C \approx 4$. An apparent value of $n_C = 4$ was recently reported and discussed for rods (Burns, Mendez, Chen, & Baylor, 2002). I will assume $n_C = 4$ below.

**Inner segment and cone pedicle**

The inner segment contains a range of ion channels, partly voltage sensitive, shaping the response (Yagi & MacLeish, 1994). To model this without introducing many (poorly known) parameters, I will follow here the approach of Detwiler, Hodgkin, and McNaughton (1980). It is assumed that the overall steady-state conductance of the membrane, $g_{is}$, is given by a (nonlinear) function of the receptor potential $V_{is}$:

$$g_{is} = g_{is}(V_{is}),$$  \hspace{1cm} (16)

where $V_{is}$ is defined relative to the resting potential (i.e., the potential when $I_{os} = 0$). The voltage response to an abrupt change in $I_{os}$ is given by the instantaneous conductance of the membrane, $g_i$, thus

$$V_{is} = \frac{I_{os}}{g_i}. $$  \hspace{1cm} (17)

Finally, it is assumed that the instantaneous conductance approaches the steady-state conductance according to first-order kinetics:

$$\tau_{is} g_i = g_{is}(V_{is}) - g_i. $$  \hspace{1cm} (18)

The precise form of $g_{is}(V_{is})$ is not known for primate cones, but I found that those aspects of horizontal cell responses that are presumably generated at the inner segment (in particular, response sagging, rebounds after pulses and steps, and reduced sensitivity at low frequencies) are well modeled by assuming

$$g_{is}(V_{is}) = a_{is} V_{is}^{\gamma}, $$  \hspace{1cm} (19)

where $a_{is}$ is a scaling constant, and $\gamma$ is a constant that is approximately 0.7 according to the fits below.

The change of $V_{is}$ in response to a change in $I_{os}$ is in reality not as instantaneous as suggested by Equation 17, because the membrane capacitance has to be charged. This can be represented by an additional low-pass filter with time constant $\tau_m$, assumed to be approximately constant. The gain of this filter is merged into the scaling constant $a_{is}$. The above equations lead to the system model shown in Figure 4. Although the signal transfer from the cone inner segment to the cone pedicle may produce some additional filtering (Hsu, Tsukamoto, Smith, & Sterling, 1998), it is assumed here that the signal arriving at the cone pedicle is that of the inner segment. An alternative interpretation is that any additional filtering can be thought to be incorporated into the parameters $\gamma$, $a_{is}$, $\tau_{is}$, and $\tau_m$ used here for the model of the inner segment.

![Figure 4. Model of the filtering properties of the inner segment.](https://example.com/f4.png)
The cone-horizontal cell feedback loop

For the feedback from the horizontal cell to the cone pedicle, I assume the subtractive scheme shown in Figure 5A. This is consistent with recent proposals for the feedback mechanism (Kamermans, Fahrenfort, Schultz, Janssen-Bienhold, Sjoerdma, et al., 2001), and also with the spatially local character of the gain control (see Discussion). The driving force \( V_s \) for transmitter release \( I_t \) is determined by the difference of the voltage of the inner segment, \( V_{in} \), and the voltage \( V_h \) of the horizontal cell multiplied by a gain \( g_h \). Thus

\[
V_s = V_{in} - g_h V_h. \tag{20}
\]

In the calculations below, \( g_h \) is fixed to 1, because it can be merged with the forward gain \( g_s \). Going around the loop there are three low-pass filters, which together with the gain \( g_s \) determine the characteristics of the resonant oscillations observed in horizontal cells. The minimum number of low-pass filters required to obtain oscillations is two, but I found that the shape of the resonance peak and the associated oscillations is better described with three filters than with two or more than three. Although the filters can be arranged in any order, I tentatively consider the filters with relatively short time constants, \( \tau_1 \) and \( \tau_2 \), as being involved in the processes of transmitter release (e.g., related to the rate of presynaptic calcium extrusion; Morgans, El Far, Berntson, Wässle, & Taylor, 1998), synaptic diffusion (e.g., related to the rate of glutamate removal; Gaal, Roska, Pi-caud, Wu, Marc, et al., 1998), or postsynaptic transduction. The longer time constant \( \tau_3 \) may then be interpreted as the effective time constant of the horizontal cell. The properties of the feedback loop, in particular the total gain, will depend on the spatial properties of the stimulus with respect to the (broad) receptive field of the horizontal cell. Because the spatial extent of all stimuli used in the measurements considered in this article was kept constant (5° diameter), I assume a fixed total gain for each cell. The influence of the spatial layout of the stimulus will be considered in a forthcoming article.

Nonlinear synaptic gain and illuminance-dependent dynamics

The release of transmitter and its postsynaptic transduction are known to be nonlinear processes. The cone-horizontal cell feedback loop, possibly together with more local ionic feedback circuits, will tend to (apparently) linearize this process (Kraaij, Spekreijse, & Kamermans, 2000). The model as presented up to this point, with a linear gain \( g_s \), produces fits acceptable fits to all horizontal cell measurements considered below. There are two nonlinearities that improved the fits sufficiently to justify their inclusion into the model. For high-contrast steps at high illuminance, there is an indication of a transiently reduced gain \( g_s \) (see Comparison with measurements, Figure 8). To model this, I included a nonlinearity similar to the one used by Kamermans, Kraaij, and Spekreijse (2001): The effective transmitter release as a function of \( V_s \) is, for the values of \( V_s \) relevant for the fits, described by a Boltzmann function:

\[
I_t(V_s) = \frac{g_t}{1 + \exp(-a(V_s - V_k)/V_n)}, \tag{21}
\]

with \( g_t \), \( V_k \), and \( V_n \) constants. The more negative \( V_s \) becomes, the more the transmitter release will shut down. Because the stimuli considered in this article are spectrally white and have a broad spatial extent (5°), the modulation of \( V_k \) is generally small (e.g., see Figure 6B, panel 11). Therefore, the effect of the nonlinearity of Equation 21 is limited. This would be different, however, when the stimulus has a narrow spatial extent, resulting in a larger difference between the response of the excited cones and the horizontal cell, and thus in a larger modulation of \( V_k \). Similarly, a non-white stimulus can increase the modulation of \( V_k \) as well. The nonlinearity of Equation 21 is then more important (Kamermans & Spekreijse, 1999).

A second nonlinearity becomes clear when inspecting the high-frequency oscillations seen in response to pulses (Figure 7). The frequency of these oscillations decreases with decreasing background intensity. This indicates that the time constants or gain of the feedback loop are intensity dependent. This also influences the high-frequency part of the curves giving the sensitivity as a function of illuminance and frequency (Figure 14). I found that the shift in
Example of model performance

The model is shown in its complete form in Figure 6A. Figure 6B shows how a simple stimulus, a 100-ms step of contrast 2 at a background illuminance of 100 td, is transformed by the subsequent processing steps in the model, using the parameters fitted to Figure 7 below. The circled numbers in the panels correspond to the numbers in Figure 6A. The intensity step is first low-pass filtered by $\tau_R$ and $\tau_E$, yielding a signal proportional to the concentration of activated PDE (panel 2). The rate of cGMP hydrolysis, $\beta$, is elevated by $c_\beta$ relative to the $E^*$ curve (Equation 10), hardly visible in the graph at these illuminances (panel 3). The inverse of $\beta$ (panel 4) is the signal that acts as the input to the calcium feedback loop. It is regulated by $1/\alpha$ (panel 8), a low-pass filtered version of the loop output. This signal is slightly delayed relative to $1/\beta$ because of the low-pass filtering, and therefore boosts high temporal frequencies (panel 5). Furthermore, the feedback effectively reduces the dynamic range needed by the signal (cf. panels 6 and 4). The feedback loop of the inner segment (panel 10) reduces low frequencies (panel 9). In panel 9 the membrane voltage of the cone is shown relative to its membrane potential in the dark (i.e., $V_{is} - V_{is, \text{dark}}$). Similarly, panel 14 shows the membrane voltage of the horizontal cell relative to its dark.
Figure 7. Black circles: data from Figures 4 and 5 of Smith et al. (2001), fitted by the model (continuous red lines). The dashed blue lines show the responses of the model using the generic parameter values (Table 1). The left column shows responses to 10-ms pulses (starting at $t=0$ ms; see black bars at time axes) of varying contrast (0.1, 1, 2, 4, 8, 16) at background illuminances of 100, 10, and 1 td. The right column shows this for 100-ms steps (starting at $t=0$ ms; see black bars at time axes). Traces are offset from 0 mV for the sake of clarity. Parameters of the fitted curves: $\tau_R = 0.49$, $\tau_E = 16.8$, $c_\beta = 2.8 \cdot 10^{-5}$, $k_\beta = 1.63 \cdot 10^{-4}$, $\tau_C = 2.89$, $a_C = 9.08 \cdot 10^{-2}$, $\gamma = 0.678$, $\tau_{is} = 56.9$, $a_{is} = 7.09 \cdot 10^{-2}$, $g_I = 151.1$, $V_I = 19.7$, and $\mu = 0.733$; for dimensions of the parameters, see Table 1. Parameters not mentioned are set at the generic values given in Table 1 (as they are for the figures below). Note that the order of the parameter values of $\tau_R$ and $\tau_E$ is arbitrary; the shortest of the values will always be arbitrarily assigned to $\tau_R$ below.

Figure 8. Effect of using a linear (dashed green line) or nonlinear (continuous red line) synaptic gain for a 100-ms step of contrast 16 at 100 td (i.e., a step from 100 to 1700 td; step starts at $t=0$ ms; see black bar at time axis). Parameters for the nonlinear case are identical to those of Figure 7; for the linear case $g_s = 8.81$, with other parameters identical to those of Figure 7.
value (i.e., \( V_{\text{h}} - V_{\text{h, dark}} \)). It should be noted that the scaling of \( V_{\text{h}} \) relative to \( V_{\text{h}} \) is not fully determined by the measurements on \( V_{\text{h}} \) considered here, because it depends on fixing \( g_b = 1 \) in Equation 20. In the cone-horizontal cell feedback loop (panels 11-14), the membrane potential of the horizontal cell is subtracted from that of the cone. The resulting \( V_{\text{h}} = V_{\text{h}} - V_{\text{h}} \) is small, only in the order of a mV. It is transient, because the signal in the horizontal cell is delayed, due to the low-pass filtering, relative to that in the cone. The resulting oscillations (panels 11, 12) arrive, however, strongly attenuated in the horizontal cell (panel 14). Note that the signal presumably going to the bipolar cells (bc in Figure 6A) is considerably more transient (panel 13) than that of the horizontal cell. How much of this shows up in the bipolar cells depends on the amount of low-pass filtering and further processing occurring in the bipolar cells.

Model implementation

I implemented the model as a series of auto-regressive moving average (ARMA) filters (e.g., Fante, 1988; also see below). A Fortran90 code producing the example of Figure 6B can be obtained by clicking here (Supplementary material). I verified the correctness of the model implementation by comparing its responses to small signals with the transfer function obtained from a small-signal analysis of the model (see Supplementary material), and by comparing its responses to large signals with a fourth-order Runge-Kutta solution of the system of differential equations involved. The implementation using ARMA filters is advantageous because it is both highly accurate and very fast, and therefore allows extensive fitting of the model to the horizontal cell data. The fitting was performed using a simplex algorithm (Press, Teukolsky, Vetterling, & Flannery, 1992) for minimizing the RMS-deviation between model responses and measurements. When responses with very different amplitudes \( r_{\text{max}} \) were used for simultaneous fitting, the RMS-deviations were scaled by \( 1/r_{\text{max}}^{0.25} \) to prevent the largest responses to completely dominate the fit. This was a pragmatic choice that still assigns more weight to large responses because they are less affected by noise and provide more information on the nonlinearities in the model. The measured data from horizontal cells in Figures 7-14 were obtained from the graphs in the on-line versions of Smith et al. (2001) and Lee et al. (2003), using specialized software (g3data).

For a first-order low-pass filter with time constant \( \tau \), the output \( y(n) \) to an input \( x(n) \), with a time step \( \Delta t \), is given by the following ARMA filter (Brown, 2000):

\[
y(n) = f_1 y(n-1) + f_2 x(n) + f_3 x(n-1)
\]

(23)

with

\[
f_1 = \exp(-1/\tau')
\]

\[
f_2 = \tau' - (1 + \tau') \exp(-1/\tau')
\]

\[
f_3 = 1 - \tau' + \tau' \exp(-1/\tau')
\]

(24)

In the calculations I used \( \Delta t = 100 \mu s \); with this time step, the model computes approximately 100 times faster than the real cone (3 GHz PC, Intel Fortran compiler on Linux). I verified that results were indistinguishable from those obtained with time steps of 10 \( \mu s \) or 200 \( \mu s \). For time steps significantly longer than 200 \( \mu s \), the implicit extra delay of one time step in the high-gain feedback loops of Figure 6A leads to spurious oscillations.

Comparison with measurements

Below I present fits of the model to a range of measurements on macaque H1 horizontal cells, as published in Lee et al. (2003) and Smith et al. (2001). The behavior of the model is complex, due to the nonlinearities and feedback loops, and the model contains many parameters. There is always the possibility, therefore, that good fits are obtained by overfitting (i.e., with very different, nonphysiological parameter values for different stimuli or different cells). To show that this is not a significant problem here, I determined a set of typical parameter values (listed in Table 1), which together determine a ”generic model” producing generic responses. For most measurements presented below (black symbols), I will show two model calculations: the generic response (dashed blue line) and the fitted response (continuous red line). I often fitted with only a subset of parameters left free, fixing the others because which parameters are relevant depends on the type of stimulus. For example, the parameters determining the behavior of the model for large changes of background intensity should not be left free to vary for a medium contrast stimulus at one particular background intensity. Details on the fitted parameters and their values are given in the Figure captions and in Table 1.

Because the properties of horizontal cells, such as their sensitivity, vary, the generic response will generally deviate from the measurements. But as will become clear below, the generic responses always capture the basic qualifying characteristics of the measured responses.

Apart from the parameters of the model of Figure 6A, there is one additional parameter that I used as a free parameter in the fits where applicable, namely an overall delay. This accounts for any delay not incorporated into the model (such as transmission and diffusion times and possibly instrumental delay). The delay was forced to be identical for the responses to all stimulus conditions used for a particular cell. It was always small, typically around 3 ms.
Table 1. Parameters and variables used in the model; see Figure 6 and Equations 7-22. Generic values are used for calculating the generic curves in the figures. The range shows the minimum and maximum values obtained from all fits presented here. Notes: the smallest of the values at $\tau_R$ and $\tau_E$ is arbitrarily assigned to $\tau_R$; generic values of $\tau_X$ ($=1/(c\beta+k\beta I)$, with $I$ in td) are 340, 230, 53, and 6.1 ms for illuminances of 1, 10, 100, and 1000 td, respectively. au = arbitrary unit.

The light intensities in this study are expressed in units of trolands (td). This is a measure of retinal illuminance, defined as the area, in mm$^2$, of the pupil of the eye times the scene luminance, expressed in candela/m$^2$. At a luminance of 100 cd/m$^2$, the pupil has a diameter of approximately 3 mm, thus then 100 cd/m$^2$ corresponds to 640 td. Such a luminance corresponds roughly to the mean luminance outdoors on a dull cloudy day, and is an order of magnitude higher than the typical mean luminance indoors.

Pulse and step responses

Figure 7 shows responses to 10-ms pulses and 100-ms steps, with various contrasts superimposed on 100-, 10-, and 1-td backgrounds. Note the nonlinearity of response size as a function of contrast and of background
intensity. The model fits (continuous red lines) were made to all stimulus conditions simultaneously, and are generally good, considering the wide range of stimulus conditions. The generic model responses (dashed blue lines) capture the qualitative characteristics of the measured responses. It is possible to identify the parts of the model responsible for various aspects of the response shapes. The response size as a function of contrast and background intensity is mainly determined by the model components up to the generation of the photocurrent, $I_{os}$. The response sagging during the step responses at 100 td, and the response rebounds after pulses and steps, are mainly due to the properties of the model components representing the inner segment. The high-frequency oscillations are due to the cone-horizontal cell feedback loop.

At 100 td these oscillations are less prominent at high-contrast steps than at low-contrast steps, which points to the existence of a nonlinearity in the cone-horizontal cell feedback loop. This was modeled in Figure 5B. Figure 8 shows the effect of the absence of this nonlinearity (as in the model of Figure 5A). The effect of the nonlinearity can be understood as follows. During a high-contrast step, $V_s$ becomes strongly negative, and thus brings the transmitter release, Equation 21, into a part of the curve with decreased slope. This decreases the small-signal gain, and this decreased gain in the feedback loop leads to a less steep rise of the response and less prominent oscillations. The fact that the oscillations in the model response are not as strongly reduced as in the measured response suggests that there are additional nonlinearities present, possibly due to voltage-sensitive channels in the horizontal cell membrane active at large hyperpolarizations.

I found that the best fits to the stimuli here and below were obtained with a low value of the time constant of the calcium feedback loop, with $\tau_C$ typically 3 ms. Forcing $\tau_C$ to be much larger worsened the fits considerably, and led to strongly biphasic cone photocurrents and membrane voltages in response to pulses and steps. This biphasic behavior is not consistent with the horizontal cell measurements considered here, which show only a mild rebound (Figure 7), with dynamics differing from those predicted from a slow calcium feedback. In rods, the photocurrent is generally not biphasic, unless the calcium dynamics are manipulated (Torre, Matthews, & Lamb, 1986). Although there are reports in the literature of strongly biphasic photocurrents and membrane voltages in primate cones (Schnapf, Nunn, Meister, & Baylor, 1990; Schneeweis & Schnapf, 1999), these findings may well be a consequence of a disturbed calcium dynamics due to the experimental techniques used. Such a disturbance of the cones is unlikely to have occurred in the horizontal cell measurements considered here, because cones were not directly manipulated, and the preparation left the retina mostly intact (Dacey, 1999; Smith et al., 2001).

**Responses to sinusoids**

Figure 9 shows responses to sinusoids of various frequencies at contrasts 0.25, 0.5, and 1, all at a 1000-td back-

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure9.png}
\caption{Black circles: data from Figure 6 of Lee et al. (2003), fitted by the model (continuous red lines). Dashed blue lines: generic model. Responses are shown to sinusoids of 0.61, 4.88, 9.76, and 30.3 Hz, of contrast 0.25, 0.5, and 1, at a mean illuminance of 1000 td. Horizontal dashed lines show the steady-state membrane potential. Fitted parameters: $\tau_R = 3.46$, $\tau_E = 9.01$, $c_\beta = 3.44 \cdot 10^{-3}$, $k_\beta = 6.44 \cdot 10^{-5}$, $\tau_c = 2.41$, $a_c = 9.51 \cdot 10^{-2}$, $\gamma = 0.488$, $\tau_u = 73.2$, and $a_u = 8.68 \cdot 10^{-2}$; fixed parameter: $g_t = 150$.}
\end{figure}
ground level. In particular at high contrasts there are several distortions visible in the experimental data, which are also produced by the model. The compressive/expansive distortion at 0.61 Hz is mainly produced by the static nonlinearity, $1/\beta$. The low-intensity part of the sinusoidal stimulus produces a small $\beta$, which is subsequently blown up by $1/\beta$ to the high peak in the response. The peak height is limited by the minimum value of $\beta$, $c_\beta$ in Equation 10. The high-intensity parts of the sinusoid produce a large $\beta$, which is subsequently compressed by $1/\beta$, which acts then as a compressive nonlinearity. The detailed shape of the distortion at 0.61 Hz is also determined by the calcium feedback loop, which in effect relinearizes the response to some extent: the high peak (low intensity, large $1/\beta$) produces high levels of calcium, reducing the gain of the forward path (divisive gain control in Figure 6A). This brings the response considerably closer to the steady-state level (dashed line). On the other hand, the low response (due to high intensity, small $1/\beta$) produces low levels of calcium, increasing the gain of the forward path, also resulting in a response closer to the steady state than would have resulted without the calcium feedback. Because the former effect is stronger than the latter, the distortion is reduced.

For higher frequencies, the distortion due to $1/\beta$ gradually disappears, because the first two low-pass filters, $\tau_R$ and $\tau_E$, reduce the depth of modulation of $\beta$ and thus $1/\beta$. At 4.88 and 9.76 Hz, another distortion becomes clearly visible: The falling flank of the response becomes steeper than the rising flank. This distortion is mainly due to the low-pass filters in the calcium feedback loop: The maximum reduction due to the control signal $1/\alpha$ (Figure 6A) is only reached right after the peak in $1/\beta$, which results in a steep falling flank right after that peak.

Finally, a third distortion can be seen at 30.3 Hz, where the rising flank is steeper than the falling flank. This is due to the filtering by the cone-horizontal cell feedback circuit. This circuit has a resonance frequency around 30-40 Hz, with strong phase changes near the resonance peak. This strongly shifts the phase of the distortion products, harmonics at multiples of the fundamental frequency, already produced by the cone. The first harmonic thus gradually shifts phase relative to the fundamental when going through the frequency range surrounding the resonance frequency, which happens to result in a steepening of the rising flank at 30.3 Hz. The distortions described here appear not to be specific to macaque H1 cells, because very similar distortions were measured in cat horizontal cells (Lankheet, van Wezel, & van de Grind, 1991).

It may be noted that some of the remaining deviations between measurements and fits in Figure 9 are related to small vertical offsets between them. Indeed, the fits improve when allowing for small errors, in the order of a mV or less, in the estimates of the steady-state potentials (horizontal dashed lines). Such errors could arise, for example, by small drifts in the intracellularly recorded membrane potential during the experiment.

### Sinusoid on sinusoid

In Lee et al. (2003) an experiment was performed to test the speed of the sensitivity regulation. It consists of a high-frequency test sinusoid superimposed on a low-frequency vehicle wave of high modulation depth. The local response to the test modulation then gives a measure of the sensitivity at a particular phase, and therefore illumination, in the vehicle wave. Figure 10 shows measurements and model responses. As can be seen, the sensitivity regulation is almost instantaneous, with only a small delay mainly due to the low-pass filters in the calcium control loop. The test response can be quantified by extracting the amplitude of the fundamental frequency for each response cycle (Lee et al., 2003). Figure 11 shows for two cells how this response varies as a function of vehicle contrast. Clearly, the responses are well captured by the model.

![Figure 10. Black circles: data from Figure 1B of Lee et al. (2003), fitted by the model (continuous red lines). Dashed blue lines: generic model. A test wave of 19.5 Hz with an amplitude of 127.5 td was superimposed on a vehicle wave of 0.61 Hz and contrast 0.825 at a mean illuminance of 1000 td. Horizontal dashed line: steady-state membrane potential. Fitted parameters: $\tau_\alpha=5.10$, $k_\beta=1.14\cdot10^{-4}$, $a_c=3.53\cdot10^{-2}$, $\gamma=0.729$, $c_\beta=22.8$, and $a_\gamma=6.68\cdot10^{-2}$; fixed parameters: $\tau_\beta=1$, $c_\beta=3\cdot10^{-3}$, $\tau_\gamma=2$, and $g=100$.](image)

### Responsivity and Weber's law

By dividing the test response amplitude (in mV) by the amplitude of the test stimulus (in troland), a measure is obtained (in mV/td) of the responsivity of the system. Studying the responsivity over a range of background intensities reveals where the system is on the scale from linear to the point where it follows Weber's law. For a linear system, the responsivity should be independent of background intensity: a stimulus of given strength in td should always produce the same response in mV. For a system following Weber's law (i.e., a system displaying contrast constancy), the responsivity should decline inversely with the background illumination: The response to 1 td at a background of 100 td (contrast 0.01) should be 10 times smaller than the response to 1 td at a background of 10 td (contrast 0.1). Intermediate stages between linear and Weber are possible as well.

---

*Lack of double-spacing in this output is due to limitations of the text extraction process.*
Figure 11. Black circles: data from Figure 4 of Lee et al. (2003), fitted by the model (continuous red lines). Dashed blue lines: generic model. The data points show the amplitude of the first harmonic of the test wave in an experiment similar to the one in Figure 10, for vehicle contrasts of 0.2, 0.34, 0.6, and 0.85. A and B are from two different H1 cells. Fitted parameters of A: $\tau_R=4.66$, $\tau_E=9.27$, $k_\beta=7.83\cdot10^{-5}$, $aC=1.09\cdot10^{-1}$, and $aS=3.33\cdot10^{-2}$; fixed parameters of A: $c_\beta=3\cdot10^{-3}$, $\tau_C=4$, and $g_t=150$. Fitted parameters of B: $\tau_R=2.87$, $\tau_E=2.99$, $k_\beta=7.34\cdot10^{-5}$, $aC=1.54\cdot10^{-1}$, and $aS=6.09\cdot10^{-2}$; fixed parameters of B: $c_\beta=3\cdot10^{-3}$, $\tau_C=4$, and $g_t=150$.

Figures 12A and B show the results of an experiment designed to test this (Lee et al., 2003). As can be seen, responsivity indeed declines as a function of background intensity, although not as fast as Weber’s law would prescribe. This can be seen better in Figure 12C, where responsivities (at the test frequency of 19.5 Hz) as a function of local illuminance have been averaged for a series of cells (black symbols, with error bounds shown by the dotted black lines). The continuous red line shows the fitted response; the generic response (dashed blue line) lies within the error bounds of the data. The Figure shows that the system behaves linearly at low intensities, and only curves toward Weber’s law at higher intensities. On this log-log plot, Weber’s law would correspond to a slope of −1. Both data and model fall in general short of that, only reaching slopes close to −1 at illuminances near 1000 td. At higher retinal illuminances, several thousands of trolands and up, pigment bleaching becomes important and starts contributing to contrast constancy (Hood & Finkelstein, 1986; Pugh & Lamb, 2000).

Figure 12. Black symbols: data from Figure 5A, 5B, and 5C of Lee et al. (2003), fitted by the model (continuous red lines). Dashed blue lines: generic model. Stimuli similar to those in Figure 11 were used, where the response amplitude was now divided by the illuminance amplitude of the test wave to yield a measure of test responsivity (mV/td). The experiment was performed at background illuminances of 10, 100, and 1000 td for two different H1 cells (panels A and B). C. Similar curves as in A and B were collected for 15 H1 cells, and the momentary test responsivity was plotted as a function of the momentary illuminance of the vehicle wave. The symbols show the average of 15 cells; dotted black lines ±1SD. Fitted parameters of A: $\tau_R=2.15$, $\tau_C=7.59$, $k_\beta=2.20\cdot10^{-4}$, $aC=2.08\cdot10^{-1}$, and $aS=1.27\cdot10^{-1}$; fixed parameters of A: $c_\beta=3\cdot10^{-3}$, $\tau_C=2$, and $g_t=150$. Fitted parameters of B: $\tau_R=2.08$, $\tau_C=10.9$, $k_\beta=3.86\cdot10^{-4}$, $aC=1.19\cdot10^{-1}$, and $aS=1.74\cdot10^{-1}$; fixed parameters of B: $c_\beta=3\cdot10^{-3}$, $\tau_C=2$, and $g_t=150$. Fitted parameters of C: $k_\beta=2.01\cdot10^{-4}$, $aC=1.22\cdot10^{-1}$, and $aS=8.05\cdot10^{-2}$; fixed parameters of C: $\tau_R=4$, $\tau_C=6$, $c_\beta=3\cdot10^{-3}$, $\tau_C=2$, and $g_t=150$. 
Peak response to increments and decrements

Figures 13A and 13B show responses, at 1000-td background illumination, to incremental and decremental pulses of several durations and contrasts. Note the asymmetry in shapes of responses to decrements and increments at larger durations, and the changes in shape for different contrasts. Again, these nonlinear aspects are well captured by the model. Keeping the parameter values obtained from the fit to these responses fixed, Figure 13C shows how the model predicts the measured peak responses as a function of pulse duration. Apart from the 6 data points explicitly fitted in Figure 13A and 13B, all other data points are accurately predicted by the model.

Sensitivity as a function of frequency

Figure 14 shows for three cells the sensitivity as a function of frequency and background illumination. The model curves were obtained by simulating the response to low-contrast sinusoids, and subsequently extracting the amplitude and phase of the fundamental frequency in the response. The sensitivity is then defined as the response amplitude per unit illuminance modulation. The main features of the measured curves are well captured by the model: (1) a decrease of responsivity as a function of background level for low frequencies, and no decline at high frequencies; (2) an increase of the cut-off frequency as a function of background level; and (3) a resonance and sharp phase decline at 30-40 Hz, in particular visible at the highest background level.

The model responses of Figure 14 are obtained using the full nonlinear model, but are in fact identical to those obtained using a linearized, small-signal version of the model (see Supplementary material). The small-signal analysis also shows that the cGMP hydrolysis by PDE (i.e., $1/\beta$) is the key factor determining the illuminance dependence of responsivity and cut-off frequency.

Discussion

The model developed here for describing sensitivity regulation in the outer retina performs quite well. It produces the main features of the responses to a wide variety of stimuli, at backgrounds ranging from 1 to 1000 td. When going from low to high illuminances, it produces the measured transition from linear behavior to contrast constancy, as well as a considerable increase in the speed of the response. Moreover, the characteristic distortions of measured responses to sinusoids, indicative of nonlinearities in the processing, are generated in the right way. The model is built from modules that all have a clear physiological interpretation, in particular in terms of the established dynamics of the phototransduction cascade, and the interactions between cones and horizontal cells.

Notwithstanding the overall success of the model, there are small differences between the modeled and measured responses for most stimuli. This is not too surprising: Several of the components were only modeled in a stylized way. For example, the calcium dynamics may be more complicated due to calcium buffering, and the dynamics of the inner segment membrane over its entire voltage range will not be as simple as modeled. Many of the parameters involved are in fact likely to change somewhat as a function of ion concentrations and membrane potential. For example, the first low-pass filter in the phototransduction cascade, $\tau_{Ta}$, is known or, in cones, at least suspected to depend on the calcium concentration (Pugh & Lamb, 2000). The membrane properties, such as input resistance and time constant, of both cone and horizontal cell are likely to change depending on membrane potential, and thus to change depending on stimulus contrast and background.
illuminance. For a more precise fitting of the model to the measurements, many of such parameter changes would probably need to be included. However, it appears that such changes can be considered as second-order effects for the present purpose. The model, as it stands, gives an adequate description and explanation of the major response properties, and it is simple and accurate enough to serve as a useful preprocessing stage for models of upstream processing in the retina and beyond.

The validity of the model is not expected to extend much beyond the illuminance range of the measurements used for validation here (i.e., approximately 1-2000 td). At the high end of this illuminance range, pigment bleaching will become increasingly important, changing the gain of the system and adding a process with particularly slow dynamics. At the low end of the illuminance range, several of the parameters may change: For example, $\tau_R$ may increase, the calcium dynamics may become slower and more complicated, the production rate of cGMP, $\alpha$ as described in Equation 15, may need an extra term $\alpha_{\text{min}}$ (Nikonov et al., 2000) at large calcium concentrations, and the filtering properties of the inner segment may change as well when it is steadily depolarized. Several of these processes could be added to the model under guidance of existing experimental data, but ideally an extended model should be validated to a set of experiments similar to those used here, performed at a still larger range of illuminances.

Much of the model draws on the remarkably detailed picture of the phototransduction cascade in rods that has unfolded in the past decades. Much less is known about the phototransduction cascade in cones, but it is generally assumed that it consists of essentially the same processes as in rods. The model fits obtained here suggest several properties that are perhaps specific to cones or even to cones of primates: The calcium feedback loop is fast ($\tau_C \approx 3 \text{ ms}$) and powerful ($n_C \approx 4$), whereas its forward gain is effectively linearized ($n_X \approx 1$), possibly due to the calcium gain control on the CNG channels. For the intensity range considered, the model did not require a calcium-dependent change in the inactivation rate of the cone visual pigment.

The cone-horizontal cell feedback circuit is similar to circuits proposed and analyzed earlier (for a review, see Kamermans & Spekreijse, 1999). In particular, it is consistent with current ideas that the feedback may be ephaptic (i.e., produced through extracellular currents and voltages in the microenvironment of the cone-horizontal cell–bipolar cell synapse; Kamermans, Fahrenfort et al., 2001). It would be consistent, though, with other mechanisms as well, as long as the feedback is mainly subtractive (i.e., linear). This is an essential requirement to be consistent with the spatially local character of the gain control (Lee, Dacey, Smith, & Pokorny, 1999). The idea is that the major nonlinearities are confined to the cone, and are thus established once the signal has arrived at the intracellular voltage...
of the cone pedicle. From that point on, the processing is essentially linear (spatiotemporal subtractive feedback), and will thus not influence indicators of local nonlinearities (e.g., spatial distortion products as in MacLeod, Williams, & Makous, 1992) or local gain control (e.g., indicated by the center response amplitude during intensity modulated rings, as in Lee et al., 1999). Note that the nonlinearities in the cone-horizontal cell feedback circuit proposed here (Figure 5B) can be neglected for low-to-medium-contrast stimuli, at least for wide field, spectrally white stimuli, and are anyway small compared with the $1/\beta$ nonlinearity of the cone.

There is a long history of attempts to model sensitivity regulation at an early processing stage in the vertebrate visual system (e.g., Sperling & Sondhi, 1968; Tranchina, Sneyd, & Cadenas, 1991; Lankheet, van Wezel, Prickaerts, & van de Grind, 1993; He & MacLeod, 1998; Snippe, Poot, & van Hateren, 2000). Usually, these models contain gain control loops and static nonlinearities in various combinations, not unlike the present model. As far as I am aware, the specific form shown in Figure 6A has not been investigated before. The main feature is that it contains a strongly compressive nonlinearity at an early stage, which is subsequently relaxed by an again strongly nonlinear feedback loop containing an illuminance-dependent time constant. This particular combination of nonlinearities and gain control is well suited for handling the intensities likely to be encountered in the natural environment. The first nonlinearity together with the gain control compresses the large dynamic range of such stimuli into a rather compact range suitable for coding as a membrane potential, and the increased time constant at low intensities averages photon noise at light levels where that is needed.

The cone-horizontal cell feedback loop produces a highly transient transmitter release at the cone pedicles. This transience is only partly visible in the responses of the horizontal cell, because this cell low-pass filters its input. For the forward path, via the bipolar cells to the inner retina, this transience is important, and in fact depends on the spatial aspects of the stimulus through the spatiotemporal filtering performed by the horizontal cells.

The measurements used here for testing the model have been used for developing a quasilinear model by Smith et al. (2001), adding a feedforward gain control loop in Lee et al. (2003). Several of the model components used in the present model have counterparts in their model. The considerable, illuminance-dependent change in the time constant $\tau$ of their model is reflected in the changing $\tau_X$ in the present model. The model for the inner segment used here behaves, for small signals, approximately as a lead-lag filter (see Supplementary material), similarly to the lead-lag filter used by Smith et al. (2001). The second-order resonator introduced in their model behaves not unlike the implicit resonance of the cone-horizontal cell feedback loop introduced here. Finally, the feedforward gain control of Lee et al. (2003) is to some extent related to the calcium feedback gain control applied here. Nevertheless, the present model supersedes the earlier attempt, because it encompasses responses over the entire range of background levels of the measurements at once, including all major nonlinearities, and is physiologically realistic.

### Acknowledgments

I wish to thank Herman Snippe for numerous discussions, and Herman Snippe and Doekele Stavenga for helpful comments on the manuscript.

Commercial relationships: none.

Corresponding author: J. H. van Hateren.

Email: j.h.van.hateren@rug.nl.

Address: Dept. of Neurobiophysics, University of Groningen, Nijenborgh 4, NL-9747 AG Groningen, The Netherlands.

### References


Rebrik, T. I., & Korenbrod, J. I. (2004). In intact mammalian photoreceptors, Ca²⁺-dependent modulation of cGMP-gated ion channels is detectable in cones but not in rods. Journal of General Physiology, 123, 63-75. [PubMed]


