A temporal model for early vision that explains detection thresholds for light pulses on flickering backgrounds

Snippe, H.P.; Poot, L.; Hateren, J.H. van

Published in:
Visual neuroscience

Publication date:
2000

Visual Neuroscience (2000), **17**, 449–462. Printed in the USA.

A temporal model for early vision that explains detection thresholds for light pulses on flickering backgrounds

H.P. SNIPPE, L. POOT, AND J.H. VAN HATEREN

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

(RECEIVED October 4, 1999; ACCEPTED January 18, 2000)

Abstract

A model is presented for the early (retinal) stages of temporal processing of light inputs in the visual system. The model consists of a sequence of three adaptation processes, with two instantaneous nonlinearities in between. The three adaptation processes are, in order of processing of the light input: a divisive light adaptation, a subtractive light adaptation, and a contrast gain control. Divisive light adaptation is modeled by two gain controls. The first of these is a fast feedback loop with square-root behavior, the second a slow feedback loop with logarithm-like behavior. This can explain several aspects of the temporal behavior of photoreceptor outputs. Subtractive light adaptation is modeled by a high-pass filter equivalent to a fractional differentiation, and it can explain the attenuation of low frequencies observed in ganglion cell responses. Contrast gain control in the model is fast (Victor, 1987), and can explain the decreased detectability of test signals that are superimposed on dynamic backgrounds. We determine psychophysical detection thresholds for brief test pulses that are presented on flickering backgrounds, for a wide range of temporal modulation frequencies of these backgrounds. The model can explain the psychophysical data for the full range of modulation frequencies tested, as well as detection thresholds obtained for test pulses on backgrounds with increment and decrement steps in intensity.

Keywords: Light adaptation, Contrast gain control, Computational model, Psychophysical detection thresholds

Introduction

The visual system functions at a wide range of light levels, from starlight to bright sunlight. To cope with this range, various processes of light adaptation have evolved. For instance, night vision is handled by rod photoreceptors, whereas during daylight the visual system uses cone vision. At photopic light levels, cone vision itself incorporates several processes for light adaptation (Valeton & van Norren, 1983).

Traditionally, psychophysicists have studied light adaptation using steps of light: the visual system adapts to a certain light level, and after full adaptation has been obtained the light is stepped to a new level. The dynamics of the adaptation process is gauged by measuring detection thresholds for a brief test probe presented at various times after the adaptation step (e.g. Crawford, 1947; Baker et al., 1959; Hayhoe et al., 1992; Poot et al., 1997). From the results of such experiments, models have been developed which combine a multiplicative light adaptation (in which the input is multiplied by a gain signal), a subtractive light adaptation (in which a signal is subtracted from the input), and a compressive (or saturating) instantaneous nonlinearity (e.g. Adelson, 1982; Kortum & Geisler, 1995; von Wiegand et al., 1995).

Address correspondence and reprint requests to: Herman P. Snippe, Department of Neurobiophysics, University of Groningen, Nijenborgh 4, NL-9747 AG Groningen, The Netherlands. E-mail: h.p.snippe@phys.rug.nl

Recently, an alternative paradigm to study the dynamics of light adaptation has attracted considerable interest (Boynton et al., 1961; Shickman 1970; Maruyama & Takahashi, 1977; Hood et al., 1997; Wu et al., 1997). In this paradigm the observer adapts to a background that is modulated (instead of stepped), and detection thresholds for brief test probes are measured for various phases of test presentation in the background modulation cycle. Conventional models for light adaptation have severe problems explaining results from this paradigm (see Hood et al., 1997 and Wu et al., 1997 for discussion). First, these models fail to describe the precise dynamics of test thresholds during the modulation cycle. Second, they do not describe the experimental result that, compared to the test threshold on a steady background, test thresholds are high throughout the background modulation cycle.

The goal of the present paper is twofold. First, we present new data for modulated backgrounds that have modulation frequencies which span the complete range of visual sensitivity from well below 1 Hz to well above flicker fusion. Second, we present a model that describes these data, as well as previous data obtained with steps in the light level (Poot et al., 1997).

A recent model for light adaptation (Wilson, 1997) can also explain detection data for test probes presented both on stepped and modulated backgrounds (see Hood & Graham, 1998). In the Wilson model, high detection thresholds on modulated backgrounds arise from push–pull connections through which ON and OFF ganglion cells inhibit each other (Hood & Graham, 1998).

This is a problematical aspect of the Wilson model, since push–pull connections have not been observed in the retina (review: Hood, 1998). In our model, the high thresholds for tests on modulated backgrounds are explained by a process of contrast gain control that scales sensitivity to the current contrast of the background dynamics (see Shah & Levine, 1996; Ahumada et al., 1998; Eisner et al., 1998 for other models of early vision that include a contrast gain control). Such an explanation is plausible since a retinal stage of contrast gain control is well established physiologically (e.g. Werblin & Copenhagen, 1974; Shapley & Victor, 1978; Lee et al., 1994; Benardete & Kaplan, 1999). Nevertheless, contrast gain control is lacking in most current models of early visual processing (e.g. Gaudiano, 1994; Dahari & Spitzer, 1996; Donner & Hemilä, 1996; Wilson, 1997; Gazères et al., 1998). On the other hand, models for contrast gain control (e.g. Victor, 1987; Wilson & Humanski, 1993; Foley, 1994; Lu & Sperling, 1996; Watson & Solomon, 1997; Carandini et al., 1997) have at best an impoverished description of the processes of light adaptation that precede the contrast gain control. The model presented here seeks to stimulate work that remedies this situation.

Psychophysical methods

A description of the methods used to obtain the psychophysical data has been published (Poot et al., 1997). Briefly, these methods were as follows.

Apparatus

Stimuli were presented monocularly through a two-channel Maxwellian-view system, thus excluding the influence of pupil size on the measurements. Two green (563 nm) Toshiba TLGD 190P light-emitting diodes (LEDs) were used as light sources. One LED provided a spatially homogeneous circular adaptation field of diameter 17 deg. The other LED was used for foveal presentation of a concentric, sharp-edged test stimulus with a diameter of 46 arcmin.

A Pentium PC controlled the LED intensities at a rate of 400 Hz, through a 12-bit digital-to-analog converter. Contrary to the description in Poot et al. (1997), the LED outputs were now linearized on-line using the photodiode-feedback design of Watanabe et al. (1992).

Stimuli

The retinal illuminance $I(t)$ of the adaptation field was harmonically modulated:

$$I(t) = I_0(1 + C \sin 2\pi ft). \quad (1)$$

The mean illuminance I_0 was 7500 Trolands (Td), and the temporal contrast C of the modulation was 0.8. Modulation frequencies f ranged from 0.39 Hz to 100 Hz; consecutive frequencies differed by a factor two (0.3 log unit). Results of the experiments at frequencies $f = 25$ Hz and $f = 50$ Hz were very different; therefore we studied one additional frequency: $f = 33.3$ Hz.

Test pulses (with duration 7.5 ms) were presented at moments that correspond to various phases ϕ of the modulation of the adaptation illuminance. Phase $\phi = 0$ deg is defined such that the middle of the test pulse coincides with the positive zero-crossing of the sine function in eqn. (1). Phase $\phi = 90$ deg corresponds to a test pulse presented when the adaptation field attains its maximum illuminance. For all background frequencies, test detection thresholds were determined at four phases in the modulation cycle: $\phi = 0$, 90, 180, and 270 deg. In addition, observer HS also measured detection thresholds at intermediate phase angles for a subset of modulation frequencies.

Interspersed between the experiments with modulated backgrounds, the detection threshold for the test pulse presented on a steady background of 7500 Td was measured. We refer to this threshold as the Weber value, since for steady backgrounds in the intensity range of the modulated backgrounds (1500–13,500 Td), it is directly proportional to the adaptation level, within experimental error (Poot et al., 1997).

Psychophysical procedure

Data were collected using a modified yes/no method (Poot et al., 1997), in sessions of duration 10–30 min. Each measurement session consisted of a number of runs. In each run, the modulation frequency f of the adaptation background and the phase ϕ of the presentation of the test pulse in the modulation cycle were kept fixed, as was the intensity of the test pulse. During a run, the modulated adaptation signal, eqn. (1), was on continuously. During the first 30 s of a run, no test pulses were presented; the observer adapted to the modulated adaptation signal. After this adaptation period test presentations began. Runs consisted of 20–50 trials. On each trial there was a 50% probability that the test pulse was actually presented, and after each trial the observer indicated, using a switch on a response box, whether the test pulse had or had not been presented (guessing if necessary). Detection thresholds for the test pulse were determined using the method of Poot et al. (1997), and correspond to 84% correct responses for an unbiased observer with identical proportions of false alarms and misses. Thresholds are based on 50–150 stimulus presentations per data point, and have an estimated uncertainty (standard deviation) of 7–10% (0.03–0.04 log unit).

Observers

The authors (age 25–39 years) were subjects in the experiments. Two observers use their spectacles to obtain good acuity; the third observer is emmetropic. Observer HS performed the most extensive measurements; the results for the two other observers support these data.

Psychophysical results

Detection thresholds as a function of the phase of the test presentation in the modulation cycle are shown in Fig. 1 for a range of modulation frequencies of the adaptation field. For most modulation frequencies the thresholds are elevated above the levels for steady backgrounds: the dashed lines show the threshold for a steady background at the time-averaged illuminance, and the curve at 0 Hz indicates the threshold for a background modulated very slowly. Further, thresholds as a function of phase ϕ deviate from harmonic functions, and there is a substantial range of frequencies (1.56–6.25 Hz) for which the thresholds peak during the upswing of the background illuminance, rather than at the moment that the background attains its maximum. The curves in Fig. 1 are calculated with the model described below.

Model

In this section, a model is presented that can explain these data. The model aims to describe the early stages of temporal processing

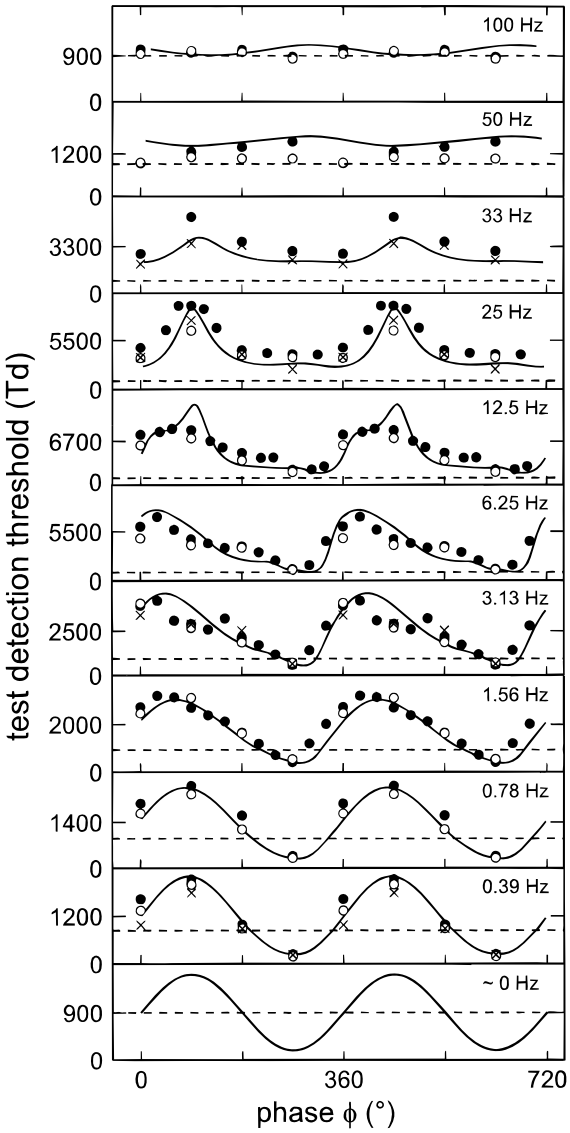

Fig. 1. Detection thresholds for test pulses as a function of the phase ϕ in the background modulation at which the test pulse is presented. Modulation frequencies f of the background range from 0.39–100 Hz. Psychophysical data are shown for three observers (symbols as in Fig. 4). To stress the similar dependence on phase ϕ for the different observers, the data plotted for observers LP and JH are their raw data multiplied by 0.69 and 0.61, respectively. This scaling equalizes their thresholds on a steady background of 7500 Td to the threshold (900 Td) obtained for observer HS on this background. The curves are predictions of the model described in Fig. 2. The lowermost curve, labeled ~0 Hz, equals the background modulation multiplied by the Weber fraction 0.12 obtained for observer HS on steady backgrounds. For the sake of clarity, two cycles of the background modulations are shown, with the same data points. Note the different scalings of the ordinates. In each of the graphs, the dashed line indicates the threshold (900 Td) on a steady background of 7500 Td. For the model calculations the parameters were $\tau_1 = 10$ ms, $k = 2.6$, $\tau_3 = 3$ ms, $n_3 = 12$, $q = 0.6$, $k_+ = 1.3$, $k_- = 0.8$, $\alpha = 0.35$, $g = 0.4$, $\beta = 3.3$, and $\lambda = 3.2$ s^{-1}.

of light inputs in the visual system; spatial and chromatic structure are left unspecified here. Although eventually it would be desirable to include spatial and chromatic effects in a model for early vision, for the present experiments a purely temporal model suffices.

As can be seen in Fig. 2A, the model is a sequence of three adaptation modules, with two instantaneous nonlinearities sandwiched in between. The nonlinearities are saturating, and describe the finite dynamic range at various levels in the visual system.

The three adaptation modules are, in order of processing of the input:

- a divisive light adaptation
- a subtractive light adaptation
- a contrast gain control

From psychophysical experiments on light adaptation (mainly using background steps), it has been concluded that light adaptation contains both divisive (also referred to as multiplicative) and subtractive components (e.g. Hayhoe et al., 1992; Graham & Hood, 1992*b*). Our model follows this tradition. It has been suggested previously that the elevation of test thresholds on modulated backgrounds above the test detection level for a steady background could arise from a contrast gain control process (Hood et al., 1997; Wu et al., 1997). This gain control would be activated by the temporal contrast of the background modulation, which would decrease the transmission gain for the test pulse, and hence its detectability. The third module in our model, contrast gain control, is a quantitative implementation of this suggestion.

We now discuss the detailed implementation of the three adaptation modules.

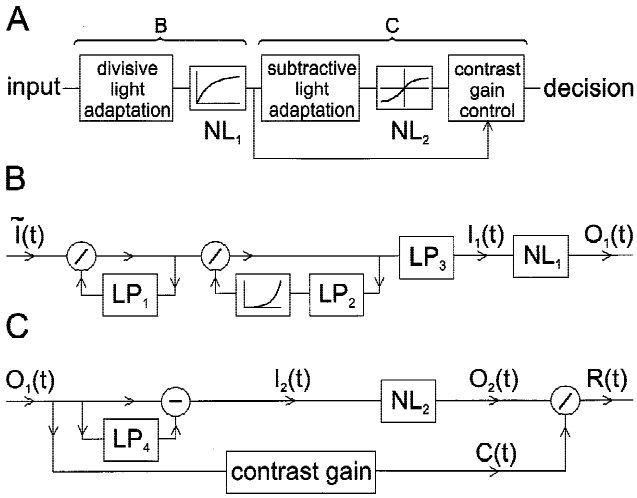

Fig. 2. Model structure. A: Model outline. The model consists of three consecutive adaptation modules, with two saturating nonlinearities (NL_1 and NL_2) sandwiched in between. The direction of the signal flow is indicated by arrows. Model input is a dynamic illuminance $I(t)$, measured in Trolands. Model output is a dynamic response $R(t)$, from which a detection threshold for a test pulse is derived. This test pulse is superimposed on the input during test trials. B: The feedback structure of the divisive light adaptation. Input $I(t)$ is divided by a constant of 1 Td, which yields a dimensionless input $\tilde{I}(t)$. LP = linear low-pass filtering. The graph in the second feedback loop represents an exponential nonlinearity. C: Subtractive light adaptation and contrast gain control. Subtractive light adaptation is generated by subtracting from the output of the divisive light adaptation module a low-pass filtered version of itself. Contrast gain control operates through a divisive feedforward loop.

Divisive light adaptation

As is shown in Fig. 2B, divisive light adaptation is implemented as a sequence of two feedback processes. It has been argued (e.g. Sperling & Sondhi, 1968; Wilson, 1997) that divisive light adaptation consists of feedback gain controls, that is, that the control signal is generated from the output of the adaptation process rather than from the input. The advantage of feedback gain control is that the dynamic range necessary for the control signal is much smaller than in a feedforward structure. Also, the biophysics of phototransduction indicates feedback control processes for divisive light adaptation (e.g. Bownds & Arshavsky, 1995; Detwiler & Gray-Keller, 1996; Koutalos & Yau, 1996).

In the first gain control loop of the model, the input is divided by a low-pass filtered version of the output. This control loop is fast; the time constant τ_1 of the first-order low-pass filter is set at 10 ms. For input signals of low frequency (e.g. $f = 1$ Hz), the filtering in the control path has no consequence. For these input frequencies, the control loop behaves as an instantaneous square-root compression device. This square-root behavior follows from the steady-state solution of the feedback loop. The input ("x") divided by the output ("y") equals the output: $x/y = y$, thus $y = \sqrt{x}$. For high frequencies (roughly $f \geq 10$ Hz), the temporal processing in the feedback path does play a role and the output is no longer simply the square root of the input. At these frequencies, the output signal is phase advanced relative to the input. Also, the output is not time symmetric: in each cycle the upstroke (increasing output) is sharper (and hence briefer) than the downstroke (Foerster et al., 1977). The temporal low-pass filtering in the feedback loop also results in transient overshoots and undershoots at steps (instead of sinusoidal variations) in the light level.

The second feedback gain control in Fig. 2B differs from the first in two important respects. First, there is now a rapidly expanding nonlinearity in the feedback path. Such a feedback structure has been used previously to describe retinal responses to flickering inputs (Tranchina & Peskin, 1988; Crevier & Meister, 1998). It is also a standard component in devices for automatic gain control (e.g. Ohlson, 1974). Here, the nonlinearity in the feedback path is modeled as an exponential function, with a multiplicative constant $k = 2.6$ in the exponent. For the complete loop, this feedback nonlinearity leads to a steady-state behavior that, at sufficiently high illuminances, is nearly logarithmic, instead of square root as in the first gain control. This is because in steady state the output ("z") of this loop equals the input y divided by the exponential of the output: $z = y/\exp kz$. Taking the logarithm of this expression yields

$$kz + \ln z = \ln y. \tag{2}$$

For high inputs, the first term in the left-hand-side of eqn. (2) dominates the second term, which leads to a logarithmic steady-state behavior $kz = \ln y$. At very low light levels (much lower than those used in the present experiments), the second term in the left-hand-side of eqn. (2) dominates the first term, which leads to the relation $\ln z = \ln y$, thus $z = y$. Therefore, at these low light levels the second gain control does not operate, leaving only the first, square-root gain control.

A second important difference between the first and the second control loop in Fig. 2B concerns the time constants of the filtering in the feedback. Whereas the first control loop is fast, the second control loop is assumed to be slow. In fact, we do not specify the exact value of the time constant τ_2 of the low-pass filtering in this feedback loop. We assume that this control loop remains virtually unmodulated even at the lowest frequency ($f = 0.4$ Hz) of the background modulation used in the experiments. Thus, the time constant in the second control loop must be considerably larger than 1 s. Adaptation processes at such long time scales are known to occur in actual photoreceptors (Baylor & Hodgkin, 1974; Laughlin & Hardie, 1978; Normann & Perlman, 1979; Valeton & van Norren, 1983).

The result $I_1(t)$ of the divisive light adaptation passes through an instantaneous, saturating nonlinearity (NL_1 in Fig. 2A), which yields the signal $O_1(t)$ on which the subtractive light adaptation operates:

$$O_1(t) = \frac{2}{\pi} \arctan I_1(t). \tag{3}$$

The nonlinearity in eqn. (3) is mathematically similar to the nonlinearity observed in the output of photoreceptors (e.g. Baylor & Hodgkin, 1974; Schnapf et al., 1990).

The nonlinearity NL_1 is preceded by a linear low-pass filtering (LP_3 in Fig. 2B), consisting of a cascade of 12 first-order filters with time constant $\tau_3 = 3$ ms each. It was necessary to choose a relatively high order for the low-pass filtering to explain the rapid decline of test thresholds for background modulations with frequencies above 25 Hz. It is well known that the time scale of filtering in the retina depends on the background light (e.g. Donner et al., 1995). However, in the present model, it was possible to ignore this dependence since the adaptation backgrounds were restricted to a relatively narrow photopic range in our experiments.

Subtractive light adaptation

As is shown in Fig. 2C, subtractive light adaptation consists of a feedforward self-inhibition scheme in which a low-pass filtered version of the output $O_1(t)$ of the divisive light adaptation is subtracted from $O_1(t)$ itself. Thus, the subtractive light adaptation consists of a linear high-pass filtering of $O_1(t)$. This high-pass filter suppresses the response to a constant input completely: after the subtractive light adaptation the steady-state response of the model to a constant adaptation field equals zero.

In the present implementation, the high-pass filtering multiplies the amplitude of its harmonic inputs of angular frequency ω with a factor ω^q. We found that a noninteger value $q = 0.6$ for the power exponent q yields the best model results. Physiologically, such power-law behavior can result from diffusion processes (Kelly, 1969), cable properties (Oldham & Spanier, 1974), or a superposition of exponential functions with a range of time scales (Thorson & Biederman-Thorson, 1974). Mathematically, it is well known that, for a causal system, this power-law behavior corresponds to a fractional differentiation (Oldham & Spanier, 1974). We explain the concept of fractional differentiation in the Appendix.

The result $I_2(t)$ of the high-pass filter passes through a saturating nonlinearity (NL_2 in Fig. 2A). Contrary to the first nonlinearity NL_1 in the model, where inputs are always ≥ 0, inputs to NL_2 can be negative. The nonlinearity NL_2 saturates for both positive and negative inputs, but the saturation is assumed to be asymmetric:

$$O_2(t) = \frac{2}{\pi k_+} \arctan(k_+ I_2(t)) \qquad \text{if } I_2(t) \geq 0,$$

$$O_2(t) = \frac{2}{\pi k_-} \arctan(k_- I_2(t)) \qquad \text{if } I_2(t) < 0, \tag{4}$$

with $k_+ = 1.3$ and $k_- = 0.8$. Both branches of eqn. (4) connect smoothly for $I_2 = 0$. However, since $k_+ > k_-$, saturation is more severe for positive signals. This saturation is crucial to explain the high detection thresholds observed at many background modulation frequencies in Fig. 1 during the upswing of the background illuminance. Due to the phase advance generated by the high-pass filter (as described in the Appendix and illustrated in Fig. 3), the output at this stage of the model attains its maximum well before the input illuminance for a wide range of frequencies. This leads to compression in NL_2, and hence to high detection thresholds when the test pulse is presented in the upswing of the background.

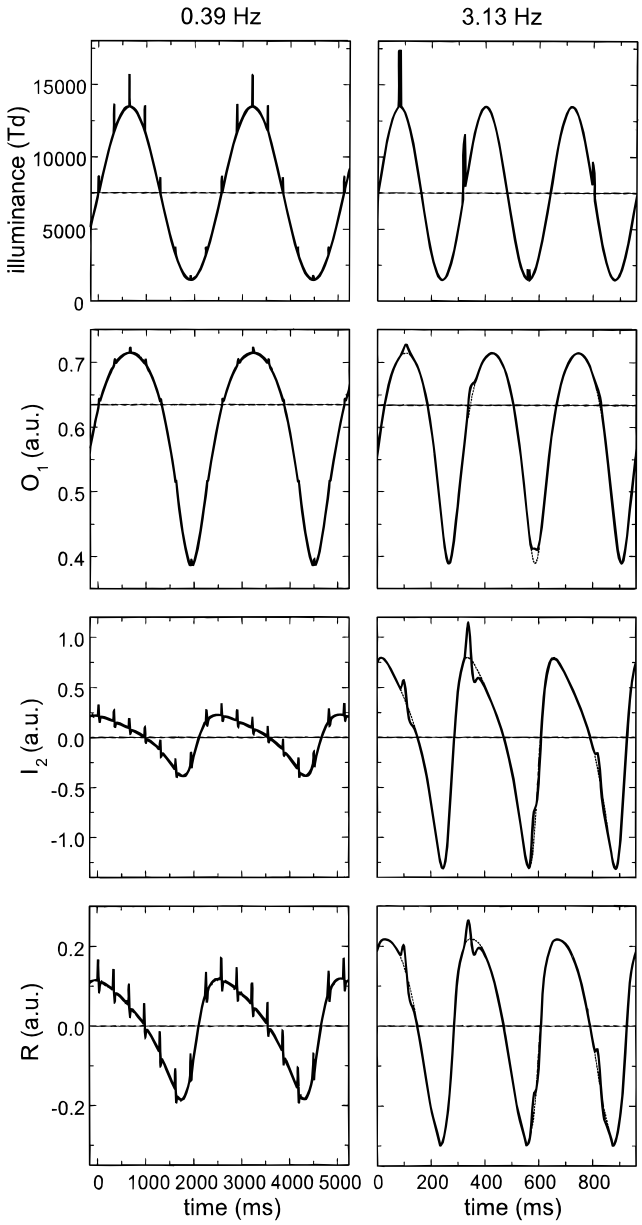

Fig. 3. Responses of the model at two frequencies of the background modulation. Input illuminance consists of a modulated background with superimposed test pulses. The amplitudes of the test pulses were adjusted such that they yield equal detectability. Intermediate responses O_1 and I_2, and final response R correspond to those designated in Fig. 2. For comparison, the thin dashed curves for 3.13 Hz show the response of the model to the background modulation without test pulses. See the text for further explanation.

Contrast gain control

Although compression in NL_2, as described above, can yield elevations of the test threshold, these elevations occur only at specific moments in the background modulation cycle (when the background response reaches an extremum). An instantaneous compressive nonlinearity (such as NL_2) cannot describe the consistent elevation of thresholds above the level obtained with steady backgrounds that is observed experimentally for tests presented at all phases in the background cycle. Such an elevation can be seen in Fig. 1 by comparing the thresholds with the dashed lines that indicate the threshold for a test presented on a steady background. This elevation of thresholds is generally observed in detection experiments with modulated backgrounds; it is known as the "dc-component" of the threshold function (Hood et al., 1997; Wu et al., 1997). In the present model, this component of threshold elevation is largely due to a divisive contrast gain control that is activated by the temporal contrast of the background. When background contrast is high, the contrast gain control signal is high, hence the test response is divided (attenuated) by a large number, leading to elevated thresholds for tests on modulated backgrounds. For flickering backgrounds, contrast gain control accounts for about 70% of the threshold dc-component in the present model; most of the remaining 30% of threshold elevation arises from compression in NL_2.

For the specific implementation of contrast gain control, a number of constraints have to be satisfied. First, for modulated backgrounds a positive contrast signal throughout the modulation cycle is desired. Second, threshold levels tend to increase with modulation frequency (up to $f = 12.5$ Hz; see Fig. 1), hence at equal physical contrast the contrast signal has to increase with increasing modulation frequency. Finally, contrast gain control can be fast. This is indicated both by psychophysics (e.g. Foley & Boynton, 1993; Poot et al., 1997; Wu et al., 1997; Wilson & Kim, 1998), and by retinal physiology (Victor, 1987; Shapley, 1997). A model for contrast gain control that satisfies these demands is presented in the Appendix.

As is shown in Fig. 2C, the output $R(t)$ of the model equals the output $O_2(t)$ of NL_2 divided by the contrast gain signal $C(t)$:

$$R(t) = \frac{O_2(t)}{C(t)}. \tag{5}$$

To illustrate the behavior of the model, Fig. 3 presents examples of signals produced by the model for two background modulation frequencies: 0.39 Hz (left) and 3.13 Hz (right). Input illuminances $I(t)$ consist of the superposition of the background modulation and a series of test pulses of duration 8 ms and with strengths equal to the detection thresholds calculated by the model. For the 0.39-Hz background, test pulses at the input are all virtually proportional to the background illuminance, indicating that Weber's law applies at this low background frequency. This is not the case for 3.13 Hz. The second row shows responses $O_1(t)$ (see Fig. 2B) of the model after the divisive light adaptation and the first nonlinearity NL_1. The dashed horizontal line indicates the steady-state response to a constant background of 7500 Td (the time-averaged background illuminance). Note the compression of the background signal. For the 0.39-Hz background, the responses

to the test pulses are all virtually identical, indicating that the part of the model shown in Fig. 2B produces the Weber behavior. The third row shows model responses $I_2(t)$ (see Fig. 2C) after the subtractive light adaptation. Note that the response to the background modulation is smaller at 0.39 Hz than at 3.13 Hz. Furthermore, the subtractive light adaptation induces a phase advance of the response to the background: the response to the test pulse presented in the upswing (positive zero crossing) of the background is now sitting on the maximum of the background response. Also note the biphasic nature of the response to the test pulses. The last row shows the model output $R(t)$ (see Fig. 2C). The difference in response to the background modulation at the two frequencies has been reduced relative to the stage I_2 (third row). This is due both to the compression in NL_2 (see Fig. 2C) and to the contrast gain control which is activated more strongly by the higher background frequency. These operations also result in equal detectability of all test pulses.

Detection of the test pulse

Detectability of a test pulse $p(t)$ superimposed on a background $I(t)$ was determined as follows. First, the model was run with solely the background signal $I(t)$ as input; this yields a model output $R(t)$. Next, the model was run using as input the sum $I(t) + p(t)$ of the background and a test pulse, yielding a model output $R_p(t)$. In calculating the response $R_p(t)$, we allowed the effects of the test pulse to enter the (feedback) gain paths of the luminance adaptation, but *not* to enter the contrast gain control path. The logic of this is that luminance adaptation is strongly localized in photopic vision (Burr et al., 1985; MacLeod et al., 1992; He & MacLeod, 1998; though see Tyler & Liu, 1996, for some counterevidence), while contrast gain control has a much larger spread, extending over several degrees (Shapley & Victor, 1979; Benardete & Kaplan, 1999). Thus, it is expected that the contrast gain that affects the test pulse (diameter 46 arcmin) mostly originates from adjacent parts of the retina where the test pulse is not present. We also performed simulations in which 10% of the contrast gain signal originates from the test area and 90% from the background field (this corresponds to assuming an area of the contrast gain with diameter about 2–3 deg). Results of these simulations correspond closely to the situation when the contrast gain signal originates completely from the background signal; we used the latter situation for simplicity.

From the model responses $R_p(t)$ and $R(t)$ the detectability d' of the test pulse is calculated as (Graham & Hood, 1992*a*; Watson & Solomon, 1997)

$$d' = \lambda \int_{-\infty}^{+\infty} |R_p(t) - R(t)|^{\beta} \, dt. \tag{6}$$

The exponent $\beta = 3.3$ was chosen identical to the steepness of the psychometric function in these experiments (Poot et al., 1997). Intensity of the test pulse for a given pulse-background condition was varied to find the pulse detection threshold, corresponding to $d' = 2$. The proportionality constant λ in eqn. (6) was determined by the detection threshold obtained for a test pulse presented on a constant background of 7500 Td. For observer HS this detection threshold was 900 Td for the 7.5 ms, 46 arcmin test pulse.

Implementation and fitting

The model was initially implemented as a MathCad program running on a PC, using Fourier methods for the calculations. To speed up the calculations, the model was subsequently implemented as a Fortran program running on a Hewlett Packard workstation. The latter implementation uses exclusively recursive filtering at a sampling frequency of 1000 Hz, and gives the same results as the MathCad program.

Parameter values used for the model results reported below are listed in the caption of Fig. 1. This set of parameter values was used for all the calculations presented in this article. The parameters were obtained by fitting simultaneously to the psychophysical results obtained with flickering (Fig. 1) and stepped (Figs. 5 and 6) backgrounds, and judging the quality of the fits by eye. We did not perform an exhaustive search through the parameter space of the model. Thus, it is likely that small differences between the data and the model could be reduced by a further optimization of the parameter settings in the model.

Model results

Steady backgrounds

A basic psychophysical result obtained for detection thresholds of test pulses presented on steady backgrounds of photopic intensity (e.g. Reeves et al., 1998) is Weber's law: test thresholds are proportional to the background illuminance. Thus, the ratio of test threshold (in Td) and the background illuminance is a constant, the Weber fraction. The present model complies with this result: for backgrounds in the range 10–100,000 Td, Weber fractions are constant to within 10%. For backgrounds above 100,000 Td, the model Weber fraction increases due to saturation at the nonlinearity NL_1. This saturation, however, could be prevented in a realistic way by including in the model the effects of photopigment depletion at these high illuminances (Burkhardt, 1994). For backgrounds below 10 Td, the Weber fraction steadily rises. Below about 0.1 Td, the model reverts to a de Vries-Rose behavior: thresholds are proportional to the square root of the background illuminance. This square-root behavior of the model at these low illuminances is entirely deterministic and caused by the first feedback loop of the divisive light adaptation in Fig. 2B.

Modulated backgrounds

Threshold predictions for test pulses on modulated backgrounds are indicated by the lines in Fig. 1. Although there certainly are differences with the psychophysical results, the model does correctly predict many of the features of the data:

- For modulation frequencies $f \leq 12.5$ Hz, threshold predictions during the upswing of the background modulation (e.g. at its positive zero crossing; $\phi = 0$ deg) are elevated with respect to predictions during the downswing (e.g. at $\phi = 180$ deg).
- Because of the threshold elevation during the upswing of the background, thresholds can reach their maximum well before the background modulation reaches its maximum. The model predicts both the maximum phase lead of the threshold maximum relative to the background (60–70 deg), and the modulation frequencies ($f = 3$–6 Hz) at which this maximum lead occurs.
- The phase lead of the threshold maxima disappears at high frequencies (e.g. at 25 Hz), but the threshold curve remains nonharmonic at this frequency: threshold maxima are sharper than threshold minima.

- Contrary to the behavior of the threshold maxima, thresholds attain their minimum when the background is minimal throughout the whole frequency range of background modulations.
- For the highest modulation frequency tested (100 Hz), thresholds for all test phases are virtually identical to the threshold on an unmodulated background of 7500 Td, as would be expected from the Talbot–Plateau law (Stockman & Plummer, 1998).
- The level of thresholds averaged over a modulation cycle (i.e. the dc-component of the threshold curve) gradually increases up to frequencies 15–20 Hz, and (on a log-frequency scale) rapidly drops at higher frequencies. This is shown more clearly in Fig. 4.

One feature that is present in the model predictions in Fig. 1, but not in the data, is a threshold peak near $\phi = 90$ deg for a modulation frequency of 12.5 Hz. However, this mismatch between the model and the psychophysical data may be less serious than it appears. Model predictions (not shown here) for a modulation frequency $f = 10$ Hz show a broad plateau that extends from $\phi = 0$ deg to $\phi = 120$ deg without an additional peak at $\phi = 90$ deg, which is very similar to the psychophysical data at 12.5 Hz. Further, the peak at $\phi = 90$ deg predicted by the model may also occur in the psychophysics, but at frequencies somewhat higher than 12.5 Hz. Such a peak is well developed in our data at 25 Hz. Also, Boynton et al. (1961) report (in their Fig. 2) a similar sharp threshold peak for a background frequency of 15 Hz.

Background steps

In this subsection, we show that the present model can explain not only test thresholds on modulated backgrounds, but also data obtained with steps in the background light. In Fig. 5, test thresholds are shown as a function of test timing τ with respect to 16-fold decrement (Fig. 5A) and increment (Fig. 5B) steps of the illuminance of the background field. The duration of the test pulse in this experiment is 10 ms; for this test duration the Weber fraction on a steady background equals 0.09 for observer HS. The data shown (not previously reported in this form) were obtained during data collection for a previous study (Poot et al., 1997). The timing convention used for the test pulse in Fig. 5 differs slightly from the convention used in this previous study. In Fig. 5, test timing is defined relative to the middle of the test pulse (which is identical to the timing convention in Fig. 1), whereas previously we timed the presentation of the test pulse with respect to its onset (a 5-ms timing difference for the 10-ms test pulse). Threshold predictions of the present model are indicated by the continuous lines. As can be seen in Fig. 5, the model predicts the major features of the data. First, for the decrement step the model correctly predicts an initial elevation of the test threshold at the moment of the decrement step of the background (although the size of the overshoot is slightly underestimated by the model). In the model, the threshold overshoot is due to the contrast gain control, which is activated by the temporal contrast of the background step. Second, for the increment step the model correctly describes a rapid initial increase of the test threshold to a high level at the moment of the background step, and a subsequent recovery of the threshold. The high threshold level at the moment of the background increment step is partly due to the temporal contrast of the background step, but the main contribution to the threshold peak is due to the saturation in the nonlinearity NL_2, since the preceding high-pass filter is strongly excited by the background step. The recovery of the threshold is

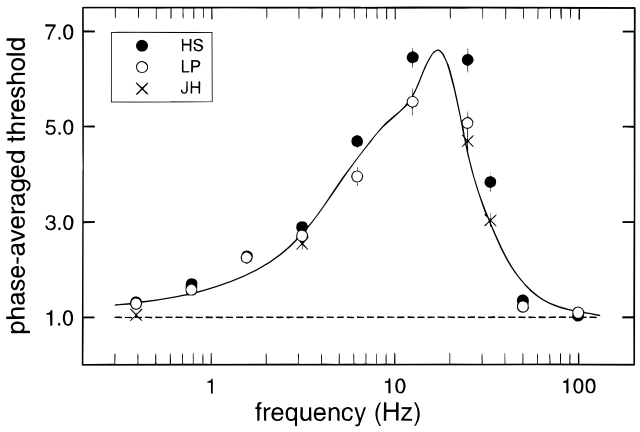

Fig. 4. Test thresholds averaged over a full modulation cycle of the background, in units of the threshold obtained for a steady background of 7500 Td (which equals the time-averaged illuminance of the modulated backgrounds) as a function of the background modulation frequency. Symbols are psychophysical data for three observers; the curve is the model prediction.

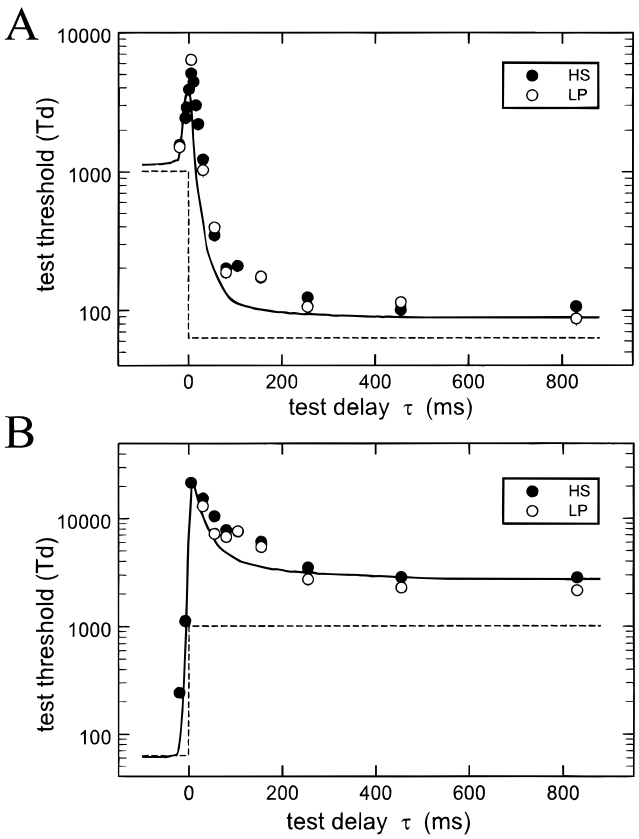

Fig. 5. Test thresholds as a function of the delay of test presentation with respect to a 16-fold decrement step (A) and a 16-fold increment step (B) of the background illuminance. Symbols are psychophysical data for two observers; the continuous curves are model predictions. As in Fig. 1, the data plotted for observer LP are the raw data multiplied by a factor 0.69, which equalizes her thresholds on steady backgrounds to those obtained for observer HS. The dashed curve is the background illuminance, scaled with the steady-state Weber fraction of observer HS (0.09 for the 10-ms test pulse).

somewhat faster for the model as compared to the data; a deficit in the model that also occurs after the background decrements in Fig. 5A. However, the model does correctly predict a flattening out of the thresholds to a level which is higher than the steady-state Weber threshold corresponding to the final background light level (this Weber threshold is indicated by the dashed curves in Fig. 5).

In our previous study, we showed that these long-term threshold elevations are not identical after increment and decrement steps in the adaptation level (Poot et al., 1997; see Yeh et al., 1996 for similar physiological results). After moderate (e.g. fourfold) increments in the background intensity, thresholds in our study were still substantially elevated above the new Weber level for tests presented 250–800 ms after the background step. However, at 100–200 ms after a similar (fourfold) decrement in the background intensity, test thresholds were nearly identical to the steady-state Weber level corresponding to the new (decremented) background illuminance. Only for the largest (16-fold) decrement steps studied was a reliable long-term threshold elevation observed. Fig. 6 shows that the present model can explain these results.

According to the model, the long-term threshold elevations after steps of background illuminance are due to a combination of the slow adaptation of the second feedback loop in the divisive light adaptation (Fig. 2B) and the subsequent nonlinearity NL_1. Consider the situation a few hundred milliseconds after an increment step of the adaptation background. Because adaptation in the second feedback loop is slow (many seconds), the feedback signal in this loop still corresponds to the old level of the adaptation illuminance, that is, it is too low. This has two consequences for the detectability of a test pulse. First, the test pulse is divided by an adaptation signal that is small (relative to the final steady state for the incremented background), which would normally lead to an improved detectability of the test pulse. However, because the response to the new *background* is also divided by the old (low) adaptation signal, the test response sits on a background response that has moved into the saturating part of the nonlinearity NL_1.

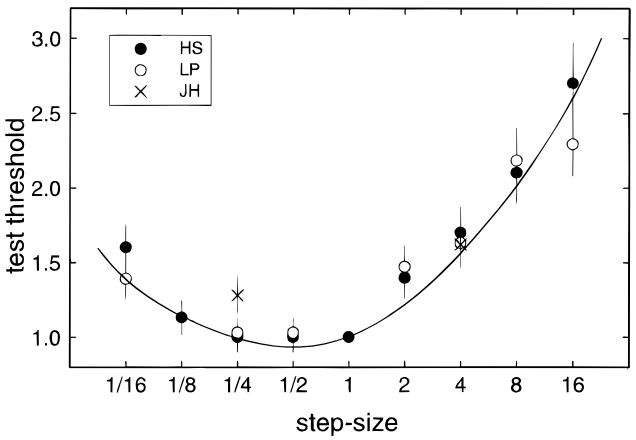

Fig. 6. Test thresholds obtained at 450–825 ms after a step in the background illuminance, as a function of the step size of the background. Step size is defined as the ratio of the background illuminance after the step and the background illuminance before the step. Hence, ratios larger than 1 are increment steps, and ratios smaller than 1 are decrement steps. Test thresholds are scaled with the steady-state (Weber) levels measured when the visual system is fully adapted to the background illuminance after the step. Symbols are psychophysical data for three observers; the curve is the model prediction.

This produces a compression and hence a lowered detectability of the test pulse. For the relatively intense (photopic) backgrounds used in the experiments, the second effect dominates after increment steps. This leads to threshold elevations compared to the (Weber) thresholds obtained when the second adaptation loop has finally reached its new steady-state adaptation level. For moderate decrement steps of the background the two effects very nearly balance: the decreased saturation due to the (temporary) leftward shift of the operating point in NL_1 almost precisely compensates for the divisive adaptation signal that is still too high. For large decrement steps this compensation is insufficient, and thresholds after the step are elevated compared to the final (new) steady state. In the quantitative implementation of these ideas, two aspects of the nonlinearity NL_1 were found to be important. First, its mathematical form. An arctangent produces good results (Fig. 6), better than other plausible saturating nonlinearities. For instance, a Naka–Rushton function would produce a threshold prediction in Fig. 6 with insufficient curvature, while the curvature for an exponential saturation would be too large. Second, the steady-state operating point of the nonlinearity NL_1 is critical. The location along the horizontal axis of the minimum of the model threshold function in Fig. 6 depends strongly on this set-point. Parameters in the model were chosen such that at high photopic adaptation illuminances (e.g. 2800 Td) the steady-state output of NL_1 is at 60% of its maximum. Steady-state outputs that are either lower (e.g. 50%), or higher (70%) produce thresholds that are inconsistent with the data in Fig. 6.

Discussion

Previous detection experiments on flickering backgrounds

The experiments with flickering backgrounds reported in the present paper are performed for flicker frequencies which span most of the range of frequency sensitivity of the visual system. Previous reports studied different subsets of this frequency range (Boynton et al., 1961; Shickman, 1970; Maruyama & Takahashi, 1977; Hood et al., 1997; Wu et al., 1997).

Boynton et al. (1961) used two flicker frequencies: 15 Hz and 30 Hz. At both frequencies, test thresholds were nearly in phase with the background, as we observed at 25 Hz (see Fig. 1). Shickman (1970) studied the frequency range of 3.1–10 Hz. At 100% background modulation contrast, thresholds were highest for test presentations in the upswing of the background, whereas minimum thresholds were obtained for tests that coincide with the background minimum. This is exactly the pattern of results that we obtained at comparable background frequencies (3.125 and 6.25 Hz; see Fig. 1). Further, as in our study, the "dc-component" of the threshold function increases with frequency in this range.

Recently, pulse-detection studies on modulated backgrounds were performed which were specifically designed to test current models for light adaptation (Hood et al., 1997; Wu et al., 1997). Wu et al. (1997) studied high modulation frequencies (20–70 Hz). In the frequency range 20–40 Hz, clear modulations in the threshold functions were observed. Thresholds were nearly in phase with the background at these frequencies, with a tendency toward an increasing phase of the threshold peaks with increasing frequency, precisely the pattern that we find at similar frequencies (12.5–33.3 Hz; see Fig. 1). At these frequencies, the dc-component of the thresholds observed by Wu et al. was still much elevated above the level for a steady background, as in our study. At still higher frequencies ($\geq$50 Hz), both the threshold modulation and the threshold elevation rapidly disappear, in complete agreement with our

study. Wu et al. took the precaution to present their background modulation in brief (Gaussian windowed) pulses, to avoid long-term effects of contrast adaptation (e.g. Magnussen & Greenlee, 1985). The similarity of their results and ours indicates that this precaution is unnecessary. Detection thresholds for brief tests presented on modulated backgrounds are very similar when the background contrast is either pulsed or present continuously.

Hood et al. (1997) used modulation frequencies 1–16 Hz. At the low frequencies (1–4 Hz), the pattern of their results is very similar to ours: threshold maxima occur during the upswing of the background illuminance and thresholds are minimal near the minimum of the background illuminance. Also, they observe a dc-component of the thresholds which steadily grows with increasing frequency, as we find as well (Fig. 4). At frequencies above 4 Hz, the results of their study and ours do not agree. Most dramatically, at frequencies 8–12 Hz, Hood et al. find threshold maxima during the *decreasing* part of the background modulation, a result not indicated by our thresholds at 6.25 and 12.5 Hz. Also the dc-component of the thresholds peaks at lower frequencies (about 8 Hz) in the Hood et al. study, compared to our results (12.5–25 Hz; see Fig. 4). There are at least three differences between our experiments and those of Hood et al. that could cause the different experimental results. First, the color of test and background used by Hood et al. was red, *versus* yellow–green in the present study. They performed a control experiment with green light (their Fig. 3) which produced results similar to those with red light, but only at a low background frequency (1 Hz) where there is no discrepancy with our results. Second, they used a spatially soft-edged test probe, whereas our test disc has hard (sharp) edges. Finally, the experiments of Hood et al. were done at adaptation levels (100–250 Td) that were much lower than our adaptation illuminance of 7500 Td. Which of these differences, if any, can explain the differences in experimental results is unclear at present and remains to be investigated in future experiments.

Detection thresholds after background steps

As described in the Model results, the present model can describe the dynamics of test thresholds obtained with steps in the adaptation background. Further, the model yields a realistic description of a long-term elevation (duration at least 1 s) of detection thresholds after increment and large decrement steps of the background (Poot et al., 1997). Similar threshold elevations after the offset of photopic backgrounds were reported by Reeves et al. (1998). At a delay of 0.2–1.6 s after the offset of the adaptation field, test thresholds in their experiments were virtually independent of the delay, and proportional to the square root of the adaptation illuminance over a range of several decades (from about 10 Td to 10,000 Td).

The present model can explain many aspects of these experimental results. First, Reeves et al. (1998) observed a Weber behavior of thresholds when the visual system was light adaptated (i.e. before offset of the adaptation field), as predicted by our model. Second, the model correctly predicts a time-independent threshold after large decrement steps of the background for the delay range used by Reeves et al. (see Fig. 5A). Finally, the model can explain the relatively shallow dependence of thresholds on the background illuminance which Reeves et al. observed after the offset of the background. The explanation is as follows. When light-adapted to a non-bleaching illuminance level I, both the output and the divisive feedback signal of the first feedback loop of the luminance gain module (Fig. 2B) are proportional to $\sqrt{I}$. Hence the second feedback loop receives an input proportional to $\sqrt{I}$. The input–output relation of this second feedback loop is shallow (due to the exponential function in the feedback path of this loop). Thus, this loop divides out most of its steady-state input, that is, the divisive feedback signal of this loop is also nearly proportional to $\sqrt{I}$. When the adaptation background drops to zero (dark) in the Reeves et al. experiment, the first feedback loop rapidly adapts (time constant 10 ms) to a small history-independent dark level. We did not include this level explicitly in the model, but this could easily be done by including a small additive constant in the feedback path of this loop (Lankheet et al., 1993). The second feedback loop in the model, however, is slow (time constant many seconds). Therefore, after the offset of the adaptation field the divisive signal in this feedback will initially remain at its light adapted level, thus dividing the response to the test signal with a factor that is nearly proportional to the square root of the adaptation illuminance. Reeves et al. (1998) present an alternative explanation of their results which includes a contribution from photon noise. The explanation of their results provided by our model, however, is entirely deterministic.

Noise

The present model does not explicitly contain any noise sources, but implicitly an additive source of noise is assumed in the decision process (i.e. a source of postsensory noise: Sperling, 1989). The strength of this noise source determines the value of the parameter λ in eqn. (6). Noise sources that occur before the final decision process could change the dependence of detection thresholds on stimulus parameters. Such noise sources could be either in the stimulus (e.g. photon noise), or they could arise during processing in the retina (sensory noise: Sperling, 1989). The present model assumes that the decision noise dominates these other noise sources. It remains for future work to determine the limits of this assumption.

A possible concern is whether the model is robust against internal noise, because the various derivative operations present in the subtractive light adaptation and the contrast gain control would amplify such noise. This problem, however, is checked by the low-pass filtering in the model. In an effort to minimize the number of free parameters in the model, we concentrated the low-pass filtering in the forward path into a single filter (LP_3 in Fig. 2B). In reality, this low-pass filtering is more distributed throughout the system. Thus, each of the derivative operations would be accompanied by some low-pass filtering; they would effectively be band-pass filters (Koenderink & van Doorn, 1987) without extreme sensitivity to high-frequency noise.

Relation with retinal physiology

The present model does not have, nor was it intended to have, a full one-to-one correspondence with retinal physiology. For instance, the output $R(t)$ of the model consists of an analog signal, not of spike trains as in the retina. To obtain good predictions for the psychophysical data, it was not necessary to include all that we know about retinal processing. However, despite this incompleteness of the model as a description of retinal processing, it is possible to point out certain relations with retinal physiology that the model does include. Although the physiological data below are obtained from a wide range of different animal species, this does not invalidate the comparison: it appears that the model relates to

physiological results that remain consistent for the retinas of a wide range of species.

The structure of the divisive light adaptation in the model is partly based on characteristics of the responses of cone photoreceptors. When fully adapted, cones have a steady-state voltage output that increases very slowly (about logarithmic) with the adaptation level: over a range of about four decades of illuminance, each tenfold increase of adaptation illuminance increases the cone steady-state output by about 10% of its dynamic range (turtles: Normann & Perlman, 1979; Burkhardt, 1994; monkeys: Valeton & van Norren, 1983). Our module for divisive light adaptation (Fig. 2B) implements this behavior, mainly through the exponential nonlinearity in the second feedback loop which leads to a roughly logarithmic input–output relation in the steady state.

Abrupt variations in input such as steps and pulses of light are transmitted with a gain that is much higher than the gain of the steady-state curve (Normann & Perlman, 1979; Valeton & van Norren, 1983). In the model, this increased gain is due to the low-pass filtering in the feedback loops, which cannot immediately follow abrupt variations in input, leading to overshoots relative to the steady-state behavior.

The first (rapid) feedback loop in the divisive light adaptation (Fig. 2B) implements a square-root transformation for most input frequencies. Evidence for a square-root transformation has been observed in turtle cones (Pluvinage & Green, 1990), as well as in horizontal cells in the cat (Lankheet et al., 1993; van de Grind et al., 1996). Lankheet and co-workers were able to quantitatively model their data using a feedback loop that is virtually identical to our first feedback loop (albeit with a longer time constant: 250 ms *vs.* 10 ms in our model). We have unsuccessfully tried to model the present results using a feedback time constant of 250 ms; a faster adaptation appears necessary to explain our results.

Further evidence for an early square-root transformation in cone vision can be observed in the responses of macaque cones. Schnapf et al. (1990) studied responses of dark-adapted cones to light pulses and steps. They deduced an output nonlinearity from the initial reaction of the cones to light pulses (their Fig. 2A). Schnapf et al. fit this nonlinearity with a weighted sum of a Naka–Rushton relation and an exponential saturation, but our nonlinearity NL_1 (an arctangent) fits their data also. After a step input, Schnapf et al. observed that 1 s after the step the initial cone response had dropped to a new level (their Fig. 7A). This new response level had a rather shallow dependence on the illuminance level I after the step (the triangles in their Fig. 7B). We find that these data can be fitted with good precision by a function of the form $\arctan[\sqrt{I}]$, as would be predicted from our model when the rapid first (square-root) feedback module has already attained adaptation after 1 s, but the second (slow) feedback module is still at its old (dark-adapted) level. Thus, at this time after the step the cone output nonlinearity receives a signal proportional to $\sqrt{I}$, leading to a rather shallow dependence of cone output on illuminance input I. Data reported by Baylor and Hodgkin (1974) in turtle cones (their Fig. 1) can be similarly explained by the present model.

As explained in the Model section, subtractive light adaptation was implemented through a feedforward self-inhibition scheme with an amplitude transmission that depends on angular frequency ω as ω^q. The power coefficient q equals the low-frequency slope on a log–log plot of transmission *versus* frequency. Measurements of this slope for ganglion cells in the cat were reported by Frishman et al. (1987) in their Table I. For spatially diffuse stimulation, which is the relevant comparison for our spatially homogeneous backgrounds, Frishman et al. report slope values at low frequencies of 0.43–0.70, depending on cell type (X or Y) and adaptation level, which is in good accord with the value $q = 0.6$ used in the present model.

Contrast gain control is a well-established feature of ganglion cells in the magnocellular pathway of the monkey (Lee et al., 1994; Benardete & Kaplan, 1999). Although, to our knowledge, for the monkey retina no experiments have been performed concerning the speed of this gain control, results from X-cells in the cat indicate that the control of retinal contrast gain is very fast, with a time constant of at most 15 ms (Victor, 1985, 1987). Contrast gain control in the present model was designed to be similarly fast. We cannot exclude the possibility that part of the threshold elevations observed with flickering backgrounds is due to a cortical contrast gain control (e.g. Carandini et al., 1997; Sengpiel et al., 1998). However, the flickering backgrounds used in the present experiments were spatially homogeneous, and therefore may have produced relatively little cortical activity. Thus, it is conceivable that the contrast gain control in the present model should be identified with a purely retinal process.

Which pathways does the model describe?

Early in retinal processing, neural signals are divided into an ON and an OFF pathway (Schiller, 1992). For the psychophysical predictions of the model, it was not necessary to introduce this division explicitly. However, model outputs with positive sign are most likely related to activity in the ON pathway, and negative outputs to activity in the OFF pathway (Watson & Solomon, 1997).

Another division in retinal processing is that into parvocellular and magnocellular pathways (Lee, 1996). Again, the present model does not include this distinction explicitly. The strong presence of a contrast gain control in the model, however, indicates that it is probably best to consider it a model of temporal processing by the magnocellular pathway (Benardete & Kaplan, 1999), since a retinal contrast gain control appears to be absent in the parvocellular pathway (Lee et al., 1994; Benardete & Kaplan, 1997). Possible reasons that detection in our experiments would be mediated through the magnocellular pathway include the relatively large size (46 arcmin) and the brief duration (7.5–10 ms) of the test pulse. At the moment, it must remain an open question whether detection in our experiments takes place through the magnocellular pathway at all times, however. For instance, the model responds strongly immediately after large increment steps in the adaptation field. At these moments, the saturating nonlinearities in the model compress the response to the test pulse, which yields high thresholds (Fig. 5B), consistent with the psychophysics. However, at present it cannot be excluded that immediately after large steps in the adaptation field psychophysical detection takes place in the parvocellular pathway, which has a smaller contrast gain (Lee, 1996) and thus would be saturated less than the magnocellular pathway. To settle this issue, it would be of interest to choose the characteristics of the test pulse such that its detection in these experiments would be mediated through the parvocellular pathway at all times. Preliminary results in our laboratory indicate that this may indeed be feasible.

Speed of divisive light adaptation

Divisive light adaptation in the present model consists of a sequence of three processes (see Fig. 2B): two feedback operations and a static nonlinearity. Two of these processes are fast: the nonlinearity NL_1 is instantaneous, and the first feedback loop contains a low-pass filter with a time constant τ_1 of only 10 ms.

Recent studies indicate processes of light adaptation that are similarly fast. Lee et al. (1997) measured light adaptation in horizontal cells of primates, by superimposing a (20 Hz) test signal on a high-contrast, 0.61-Hz background modulation (with a mean illuminance of 1000 Td). From the modulation of the responsivity of the horizontal cells that is induced by the background, Lee et al. conclude that "adaptation was complete within *ca.* 25 ms," in good accord with our estimate $\tau_1 = 10$ ms for the time constant of the fast feedback control. Psychophysical experiments by He and MacLeod (1998) using laser interferometry also indicate a fast process of light adaptation (7 ms full width at half-height, which corresponds to a time constant of 10 ms). Fast adaptation in the retina is also indicated by ERG measurements (Wu et al., 1995).

Relationship with natural stimuli

Subtractive light adaptation in the model is implemented through a high-pass filtering with an amplitude transmission ω^q. Here we show that this is consistent with optimal processing for the dynamic inputs encountered by the visual system when moving about in a natural environment.

For typical natural environments, the temporal power spectrum of a spatially localized input to the visual system (e.g. on the scale of foveal receptive fields in the retina) depends on the temporal frequency ω of the input as $1/\omega^\gamma$, with γ close to 1 (van Hateren, 1997). Thus, low frequencies dominate the input. To prevent these low frequencies from saturating the available dynamic range, they should be attenuated. The best way to do this, at least at high signal-to-noise ratios, is to whiten (decorrelate) the input (Srinivasan et al., 1982; Atick & Redlich, 1990; van Hateren, 1993). Whitening the input is achieved by a filter that has a power spectrum ω^γ and hence an amplitude spectrum $\omega^{\gamma/2}$, which for $\gamma \approx 1$ has a fractional exponent of about 0.5, close to the value $q = 0.6$ used in the model.

Plausibility of contrast gain control

Thresholds for tests presented on flickering backgrounds tend to be elevated relative to tests on a steady background; this elevation is known as the threshold dc-component (Hood et al., 1997; Wu et al., 1997). In the present model, most of this dc-component of test thresholds is due to a contrast gain control (some additional contributions to the dc-threshold component originate from the saturating nonlinearities in the model). Since a retinal process of contrast gain control is well established (review: Hood (1998), it is an *a priori* likely explanation of this threshold component. However, in their Fig. 9, Wu et al. (1997) show that for a 30-Hz background stimulus with a time-varying contrast envelope, test thresholds follow the background contrast without any noticeable delay. Wu et al. expect contrast gain control to produce a lag of the test thresholds behind the contrast envelope of the background, caused by the time necessary in the visual system to evaluate background contrast. They therefore conclude that contrast gain cannot be the explanation of the threshold dc-component. In the present model, however, we present an implementation (using temporal derivatives) which can yield an arbitrarily fast contrast estimate. Thus, we claim that the result obtained by Wu et al. does not rule out contrast gain control as an explanation of the dc-component of pulse-detection thresholds on modulated backgrounds. Rather, it constrains the gain control of temporal contrast to be fast, as it is in the present model.

The Wilson (1997) model

To our knowledge, at the moment only two models have been shown to be able to explain the behavior of test pulses presented on flickering backgrounds: the present model, and the Wilson (1997) model, as adapted by Hood and Graham (1998). Here we highlight two important differences between our model and Wilson's model. First, in our model divisive light adaptation precedes subtractive light adaptation (Fig. 2A), whereas in the Wilson model a subtractive operation (of the cones and the horizontal cells) precedes a major divisive adaptation at the level of the bipolar cells. A consequence of that ordering is that there is relatively little divisive gain control in the cones, which leads to a steady-state cone response in the Wilson model that increases rapidly with retinal illuminance (a $\frac{2}{3}$ power law). In the Appendix of his paper, Wilson shows how this could be reduced to a $\frac{1}{2}$ power law, but even this would represent a growth of the steady-state cone output that is much faster than the log-like increase of the steady-state cone output observed in physiological experiments (e.g. Normann & Perlman, 1979; Valeton & van Norren, 1983; Burkhardt, 1994). The divisive light adaptation in our model, on the other hand, was designed such that it generates this slow growth of the cone steady-state output. Second, an important feature of the Wilson model is a push–pull arrangement of the ON-center and OFF-center bipolar inputs to the ganglion cells, which is crucial to explain the elevation of test thresholds on modulated backgrounds (Hood & Graham, 1998), but which appears to be absent in the actual retina (Hood, 1998). On the other hand, our model explains the threshold elevations through a contrast gain control, which is well established in the magnocellular retinal pathway, but which is absent in the Wilson model.

Finally, simulations of the Wilson model performed in our laboratory show an unexpected behavior of its bipolar cells. ON-center bipolar cells in the Wilson model receive antagonistic inputs from the cones and the horizontal cells with (relative) weights respectively -7 and $+3$ [Wilson, 1997; his eqn. (4)]. OFF-center bipolar cells receive antagonistic inputs from the cones and the horizontal cells with weights respectively $+3$ and -7 (Dr. H.R. Wilson; Matlab code and personal communication, 1999). A consequence of this arrangement is that in steady-state ON and OFF bipolars have an identical response (which increases with the adaptation illuminance). At high modulation frequencies (which can be followed by the cones, but not by the horizontal cells), ON and OFF bipolars respond nearly in counterphase. This is the temporal relation between ON and OFF cells that would normally be expected. However, for frequencies below about 0.5 Hz, which can be followed by the horizontal cells, ON and OFF bipolars respond *in* phase. This is an unexpected prediction of the Wilson model that remains to be tested in physiological experiments.

Acknowledgments

This research was supported by the Netherlands Organization for Scientific Research (NWO) through the Research Council for Earth and Lifesciences (ALW).

References

Adelson, E.H. (1982). Saturation and adaptation in the rod system. *Vision Research* **22**, 1299–1312.

Adelson, E.H. & Bergen, J.R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A* **2**, 284–299.

Ahumada, A.J., Jr., Beard, B.L. & Eriksson, R. (1998). Spatio-temporal

discrimination model predicts temporal masking functions. *SPIE Proceedings* **3299**, paper 14.

Atick, J.J. & Redlich, A.N. (1990). Towards a theory of early visual processing. *Neural Computation* **2**, 308–320.

Baker, H.D., Doran, M.D. & Miller, K.E. (1959). Early dark adaptation to dim luminances. *Journal of the Optical Society of America* **49**, 1065–1070.

Baylor, D.A. & Hodgkin, A.L. (1974). Changes in time scale and sensitivity in turtle photoreceptors. *Journal of Physiology* (London) **242**, 729–758.

Benardete, E.A. & Kaplan, E. (1997). The receptive field of the primate P retinal ganglion cell, I: Linear dynamics. *Visual Neuroscience* **14**, 169–185.

Benardete, E.A. & Kaplan, E. (1999). The dynamics of primate M retinal ganglion cells. *Visual Neuroscience* **16**, 355–368.

Bownds, M.D. & Arshavsky, V.D. (1995). What are the mechanisms of photoreceptor adaptation? *Behavioral and Brain Sciences* **18**, 415–424.

Boynton, R.M., Sturr, J.F. & Ikeda, M. (1961). Study of flicker by the incremental threshold technique. *Journal of the Optical Society of America* **51**, 196–201.

Burkhardt, D.A. (1994). Light adaptation and photopigment bleaching in cone photoreceptors *in situ* in the retina of the turtle. *Journal of Neuroscience* **14**, 1091–1105.

Burr, D.C., Ross, J. & Morrone, M.C. (1985). Local regulation of luminance gain. *Vision Research* **25**, 717–727.

Carandini, M., Heeger, D.J. & Movshon, J.A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *Journal of Neuroscience* **17**, 8621–8644.

Crawford, B.H. (1947). Visual adaptation in relation to brief conditioning stimuli. *Proceedings of the Royal Society B* **134**, 283–302.

Crevier, D.W. & Meister, M. (1998). Synchronous period-doubling in flicker vision of salamander and man. *Journal of Neurophysiology* **79**, 1869–1878.

Dahari, R. & Spitzer, H. (1996). Spatiotemporal adaptation model for retinal ganglion cells. *Journal of the Optical Society of America A* **13**, 419–435.

Detwiler, P.B. & Gray-Keller, M.P. (1996). The mechanisms of vertebrate light adaptation: Speeded recovery versus slowed activation. *Current Opinion in Neurobiology* **6**, 440–444.

Donner, K., Koskelainen, A., Djupsund, K. & Hemilä, S. (1995). Changes in retinal time scale under background light: Observations on rods and ganglion cells in the frog retina. *Vision Research* **35**, 2255–2266.

Donner, K. & Hemilä, S. (1996). Modelling the spatio-temporal modulation response of ganglion cells with difference-of-Gaussians receptive fields: Relation to photoreceptor kinetics. *Visual Neuroscience* **13**, 173–186.

Eisner, A., Shapiro, A.G. & Middleton, J.A. (1998). Equivalence between temporal frequency and modulation depth for flicker response suppression: Analysis of a three-process model of visual adaptation. *Journal of the Optical Society of America A* **15**, 1987–2002.

Foerster, M.H., van de Grind, W.A. & Grüsser, O.-J. (1977). Frequency transfer properties of three distinct types of cat horizontal cells. *Experimental Brain Research* **29**, 347–366.

Foley, J.M. & Boynton, G.B. (1993). Forward pattern masking and adaptation: Effects of duration, interstimulus interval, contrast, and spatial and temporal frequency. *Vision Research* **33**, 959–980.

Foley, J.M. (1994). Human luminance pattern-vision mechanisms: Masking experiments require a new model. *Journal of the Optical Society of America A* **11**, 1710–1719.

Frishman, L.J., Freeman, A.W., Troy, J.B., Schweitzer-Tong, D.E. & Enroth-Cugell, C. (1987). Spatiotemporal frequency responses of cat retinal ganglion cells. *Journal of General Physiology* **89**, 599–628.

Gaudiano, P. (1994). Simulations of X and Y retinal ganglion cell behavior with a nonlinear push–pull model of spatiotemporal retinal processing. *Vision Research* **34**, 1767–1784.

Gazères, N., Borg-Graham, L.J. & Frégnac, Y. (1998). A phenomenological model of visually evoked spike trains in cat geniculate nonlagged X-cells. *Visual Neuroscience* **15**, 1157–1174.

Graham, N. & Hood, D.C. (1992a). Quantal noise and decision rules in dynamic models of light adaptation. *Vision Research* **32**, 779–787.

Graham, N. & Hood, D.C. (1992b). Modeling the dynamics of light adaptation: The merging of two traditions. *Vision Research* **32**, 1373–1393.

Hayhoe, M.M., Levin, M.E. & Koshel, R.J. (1992). Subtractive processes in light adaptation. *Vision Research* **32**, 323–333.

He, S. & MacLeod, D.I.A. (1998). Contrast-modulation flicker: Dynamics and spatial resolution of the light adaptation process. *Vision Research* **38**, 985–1000.

Hood, D.C., Graham, N., von Wiegand, T.E. & Chase, V.M. (1997). Probed-sinewave paradigm: A test of models of light-adaptation dynamics. *Vision Research* **37**, 1177–1191.

Hood, D.C. (1998). Lower-level visual processing and models of light adaptation. *Annual Review of Psychology* **49**, 503–535.

Hood, D.C. & Graham, N. (1998). Threshold fluctuations on temporally modulated backgrounds: A possible physiological explanation based upon a recent computational model. *Visual Neuroscience* **15**, 957–967.

Kasdin, N.J. (1995). Discrete simulation of colored noise and stochastic processes and $1/f^{\alpha}$ power law noise generation. *Proceedings of the IEEE* **83**, 802–827.

Kelly, D.H. (1969). Diffusion model of linear flicker responses. *Journal of the Optical Society of America* **59**, 1665–1670.

Koenderink, J.J. & van Doorn, A.J. (1987). Representation of local geometry in the visual system. *Biological Cybernetics* **55**, 367–375.

Kortum, P.T. & Geisler, W.S. (1995). Adaptation mechanisms in spatial vision—II. Flash thresholds and background adaptation. *Vision Research* **35**, 1595–1609.

Koutalos, Y. & Yau, K.-W. (1996). Regulation of sensitivity in vertebrate rod photoreceptors by calcium. *Trends in Neurosciences* **19**, 73–81.

Lankheet, M.J.M., van Wezel, R.J.A., Prickaerts, J.H.H.J. & van de Grind, W.A. (1993). The dynamics of light adaptation in cat horizontal cell responses. *Vision Research* **33**, 1153–1171.

Laughlin, S.B. & Hardie, R.C. (1978). Common strategies for light adaptation in the peripheral visual systems of fly and dragonfly. *Journal of Comparative Physiology* **128**, 319–340.

Lee, B.B., Pokorny, J., Smith, V.C. & Kremers, J. (1994). Responses to pulses and sinusoids in macaque ganglion cells. *Vision Research* **34**, 3081–3096.

Lee, B.B. (1996). Receptive field structure in the primate retina. *Vision Research* **36**, 631–644.

Lee, B.B., Dacey, D.M., Smith, V.C. & Pokorny, J. (1997). Time course and cone specificity of adaptation in primate outer retina. *Investigative Ophthalmology and Visual Science* **38**, S1163.

Lu, Z.-L. & Sperling, G. (1996). Contrast gain control in first- and second-order motion perception. *Journal of the Optical Society of America A* **13**, 2305–2318.

MacLeod, D.I.A., Williams, D.R. & Makous, W. (1992). A visual nonlinearity fed by single cones. *Vision Research* **32**, 347–363.

Magnussen, S. & Greenlee, M.W. (1985). Marathon adaptation to spatial contrast: Saturation in sight. *Vision Research* **25**, 1409–1411.

Maruyama, K. & Takahashi, M. (1977). Wave form of flickering stimulus and visual masking function. *Tohoku Psychologica Folia* **36**, 120–133.

Morrone, M.C. & Owens, R.A. (1987). Feature detection from local energy. *Pattern Recognition Letters* **6**, 303–313.

Normann, R.A. & Perlman, I. (1979). The effects of background illumination on the photoresponses of red and green cones. *Journal of Physiology* (London) **286**, 491–507.

Ohlson, J.E. (1974). Exact dynamics of automatic gain control. *IEEE Transactions on Communications* **COM-22**, 72–75.

Oldham, K.B. & Spanier, J. (1974). *The Fractional Calculus.* New York and London: Academic Press.

Pluvinage, V. & Green, D.G. (1990). Evidence for a power law intensity code in the coupled cones of the turtle. *Vision Research* **30**, 673–682.

Poot, L., Snippe, H.P. & van Hateren, J.H. (1997). Dynamics of adaptation at high luminances: Adaptation is faster after luminance decrements than after luminance increments. *Journal of the Optical Society of America A* **14**, 2499–2508.

Reeves, A., Wu, S. & Schirillo, J. (1998). The effect of photon noise on the detection of white flashes. *Vision Research* **38**, 691–703.

Schiller, P.H. (1992). The ON and OFF channels of the visual system. *Trends in Neurosciences* **15**, 86–92.

Schnapf, J.L., Nunn, B.J., Meister, M. & Baylor, D.A. (1990). Visual transduction in cones of the monkey *Macaca Fascicularis*. *Journal of Physiology* **427**, 681–713.

Sengpiel, F., Baddeley, R.J., Freeman, T.C.B., Harrad, R. & Blakemore, C. (1998). Different mechanisms underlie three inhibitory phenomena in cat area 17. *Vision Research* **38**, 2067–2080.

Shah, S. & Levine, M.D. (1996). Visual information processing in primate cone pathways—Part I: A model. *IEEE Transactions on Systems, Man, and Cybernetics* **B**, 259–274.

SHAPLEY, R.M. & VICTOR, J.D. (1978). The effect of contrast on the transfer properties of cat retinal ganglion cells. *Journal of Physiology* **285**, 275–298.
SHAPLEY, R.M. & VICTOR, J.D. (1979). Non-linear spatial summation and the contrast gain control of cat retinal ganglion cells. *Journal of Physiology* **290**, 141–161.
SHAPLEY, R.M. (1997). Retinal physiology: Adapting to the changing scene. *Current Biology* **7**, R421–R423.
SHICKMAN, G.M. (1970). Visual masking by low-frequency sinusoidally modulated light. *Journal of the Optical Society of America* **60**, 107–117.
SPERLING, G. & SONDHI, M.M. (1968). Model for visual luminance discrimination and flicker detection. *Journal of the Optical Society of America* **58**, 1133–1145.
SPERLING, G. (1989). Three stages and two systems of visual processing. *Spatial Vision* **4**, 183–207.
SRINIVASAN, M.V., LAUGHLIN, S.B. & DUBS, A. (1982). Predictive coding: A fresh view of inhibition in the retina. *Proceedings of the Royal Society B* **216**, 427–459.
STOCKMAN, A. & PLUMMER, D.J. (1998). Color from invisible flicker: A failure of the Talbot-Plateau law caused by an early 'hard' saturating nonlinearity used to partition the human short-wave cone pathway. *Vision Research* **38**, 3703–3728.
THORSON, J. & BIEDERMAN-THORSON, M. (1974). Distributed relaxation processes in sensory adaptation. *Science* **183**, 161–172.
TRANCHINA, D. & PESKIN, C.S. (1988). Light adaptation in the turtle retina: Embedding a parametric family of linear models in a single nonlinear model. *Visual Neuroscience* **1**, 339–348.
TYLER, C.W. & LIU, L. (1996). Saturation revealed by clamping the gain of the retinal light response. *Vision Research* **36**, 2553–2562.
VALETON, J.M. & VAN NORREN, D. (1983). Light adaptation of primate cones: An analysis based on extracellular data. *Vision Research* **23**, 1539–1547.
VICTOR, J.D. (1985). The nonlinear dynamics of the center mechanism of cat retinal ganglion cells. *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, pp. 867–873. Tucson, Arizona.
VICTOR, J.D. (1987). The dynamics of the cat retinal X cell centre. *Journal of Physiology* (London) **386**, 219–246.
VAN DE GRIND, W.A., LANKHEET, M.J.M., VAN WEZEL, R.J.A., ROWE, M.H. & HULLEMAN, J. (1996). Gain control and hyperpolarization level in cat horizontal cells as a function of light and dark adaptation. *Vision Research* **36**, 3969–3985.
VAN HATEREN, J.H. (1993). Spatiotemporal contrast sensitivity of early vision. *Vision Research* **33**, 257–267.
VAN HATEREN, J.H. (1997). Processing of natural time series of intensities by the visual system of the blowfly. *Vision Research* **37**, 3407–3416.
VON WIEGAND, T.E., HOOD, D.C. & GRAHAM, N. (1995). Testing a computational model of light-adaptation dynamics. *Vision Research* **35**, 3037–3051.
WATANABE, T., MORI, N. & NAKAMURA, F. (1992). A new superbright LED stimulator: Photodiode-feedback design for linearizing and stabilizing emitted light. *Vision Research* **32**, 953–961.
WATSON, A.B. & SOLOMON, J.A. (1997). Model of visual contrast gain control and pattern masking. *Journal of the Optical Society of America A* **14**, 2379–2391.
WERBLIN, F.S. & COPENHAGEN, D.R. (1974). Control of retinal sensitivity—III. Lateral interactions at the inner plexiform layer. *Journal of General Physiology* **63**, 88–110.
WILSON, H.R. & HUMANSKI, R. (1993). Spatial frequency adaptation and contrast gain control. *Vision Research* **33**, 1133–1149.
WILSON, H.R. (1997). A neural model of foveal light adaptation and afterimage formation. *Visual Neuroscience* **14**, 403–423.
WILSON, H.R. & KIM, J. (1998). Dynamics of a divisive gain control in human vision. *Vision Research* **38**, 2735–2741.
WU, S., BURNS, S.A. & ELSNER, A.E. (1995). Effects of flicker adaptation and temporal gain control on the flicker ERG. *Vision Research* **35**, 2943–2953.
WU, S., BURNS, S.A., ELSNER, A.E., ESKEW, R.T., JR., & HE, J. (1997). Rapid sensitivity changes on flickering backgrounds: Tests of models of light adaptation. *Journal of the Optical Society of America A* **14**, 2367–2378.
YEH, T., LEE, B.B. & KREMERS, J. (1996). The time course of adaptation in macaque retinal ganglion cells. *Vision Research* **36**, 913–931.

Appendix

Fractional differentiation

Subtractive light adaptation in the model is implemented such that it is equivalent to a fractional differentiation. Since the concept of fractional differentiation may be not very well known within the visual community, a brief description is provided here. For more extensive discussion and implementation algorithms, see Oldham and Spanier (1974) and Kasdin (1995).

Fractional differentiation is most easily explained in the Fourier domain. First, consider a simple (first-order) differentiation of a harmonic function:

$$\frac{\mathrm{d}}{\mathrm{d}t}\sin\omega t = \omega\cos\omega t = \omega\sin\left(\omega t + \frac{\pi}{2}\right). \tag{A1}$$

Relative to the input function, the differentiation multiplies the amplitude by ω and produces a phase advance of $\pi/2$ radians. Likewise, a second-order differentiation $\mathrm{d}^2/\mathrm{d}t^2$ leads to an amplitude factor ω^2 and a phase advance $2\pi/2$. Fractional differentiation is an extension of these results to noninteger order q:

$$\frac{\mathrm{d}^q}{\mathrm{d}t^q}\sin\omega t \equiv \omega^q\sin\left(\omega t + q\,\frac{\pi}{2}\right). \tag{A2}$$

Fractional differentiation of order q can thus be described by a linear operator with a complex transmission function:

$$A_q(\omega) = |\omega|^q e^{iq(\pi/2)\mathrm{sgn}(\omega)} \tag{A3}$$

in which $\mathrm{sgn}(\omega)$ represents the sign of ω.

The pulse-response function of this operator is the inverse Fourier transform of this transmission function. This pulse response consists of a brief excitatory impulse at $t = 0$, immediately followed by an inhibition that is sharp at first and then has a tail which, for noninteger q, behaves as a power function t^{-1-q}. The power-law tail of the pulse response of a fractional differentiation differs from the functions usually considered in linear systems theory which have an exponential tail $e^{-t/\tau}$ that can be described by a single time scale τ. Fractional differentiation is time-scale invariant, although for a given frequency range in practice it can be described using a finite number of time scales (Thorson & Biederman-Thorson, 1974). This is in fact how it was calculated in the model: a low-pass filter consisting of the superposition of a range of first-order low-pass filters (LP_4 in Fig. 2C) forms the inhibitory tail of the impulse response of the fractional differentiation.

Implementation of the contrast gain control

Contrast gain control in the model consists of a divisive feedforward loop (see Fig. 2C). The calculation that yields the contrast gain control signal is shown in Fig. 7. Although we do not claim that this implementation nec-

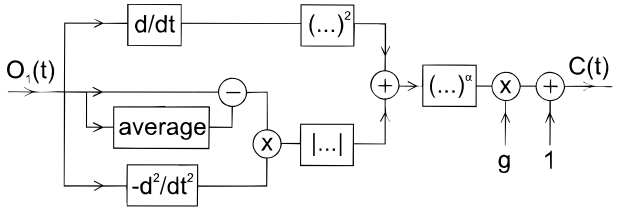

Fig. 7. Calculation procedure for the contrast gain control signal $C(t)$. See the text for further explanation.

essarily reflects the underlying physiology, the predictions of the model are consistent with our measurements.

To understand the logic of this method to calculate a contrast signal, imagine that the input $i(t)$ of the contrast calculation consists of a harmonic signal with frequency ω and contrast c, possibly superimposed on a time-averaged signal i_{av}:

$$i(t) = i_{av} + c \sin \omega t. \tag{A4}$$

In practice, the time-averaging operation in the contrast calculation of Fig. 7 would be performed using a low-pass filter in the visual system. Here we do not specify the precise value of the time constant of this filter, but we assume that it is sufficiently slow such that it does not transmit the modulation frequencies ω used in the present experiments. Then the product of $i - i_{av}$ and $-\mathrm{d}^2 i/\mathrm{d}t^2$ (see Fig. 7) yields $\omega^2 c^2 \sin^2 \omega t$, whereas the square of $\mathrm{d}i/\mathrm{d}t$ yields $\omega^2 c^2 \cos^2 \omega t$. Hence the sum of these signals, $\omega^2 c^2$, is independent of time t (since $\sin^2 + \cos^2 = 1$), and increases both with contrast and with temporal frequency, as is necessary to explain our data. This method of obtaining a time-invariant contrast signal is similar to methods that construct quadrature filters by performing a Hilbert transformation of the input signal (e.g. Adelson & Bergen, 1985; Morrone & Owens, 1987). However, contrary to a Hilbert transformation, the present calculation is causal, and fast. The result of this calculation is compressed with a power $\alpha = 0.35$, multiplied by a constant $g = 0.4$, and added to 1 (the default setting when no contrast is present), which yields the divisive contrast gain control signal $C(t)$ in Fig. 2C.