INSTITUTE OF PHYSICS PUBLISHING NETWORK: COMPUTATION IN NEURAL SYSTEMS

Network: Comput. Neural Syst. **14** (2003) 119–135 PII: S0954-898X(03)57926-X

Adaptive encoding in the visual pathway

Nicholas A Lesica, Alireza S Boloori and Garrett B Stanley[1]

Division of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA

E-mail: gstanley@deas.harvard.edu

Received 17 March 2002, in final form 6 December 2002
Published 17 January 2003
Online at stacks.iop.org/Network/14/119

Abstract

In a natural setting, the mean luminance and contrast of the light within a visual neuron's receptive field are constantly changing as the eyes saccade across complex scenes. Adaptive mechanisms modulate filtering properties of the early visual pathway in response to these variations, allowing the system to maintain differential sensitivity to nonstationary stimuli. An adaptive variant of the reverse correlation technique is used to characterize these changes during single trials. Properties of the adaptive reverse correlation algorithm were investigated via simulation. Analysis of data collected from the mammalian visual system demonstrates the ability to continuously track adaptive changes in the encoding scheme. The adaptive estimation approach provides a framework for characterizing the role of adaptation in natural scene viewing.

1. Introduction

The early visual pathway is a spatiotemporal filter that maps visual stimuli into neuronal response. Classically, techniques that assume time invariance, such as reverse correlation, have been used to analyse the behaviour of the pathway in steady state. While these studies characterize the system's response to a specific stimulus ensemble, they fail to capture important changes in the encoding process that occur during natural vision. As an eye saccades across a complex natural scene, the mean and variance of the light falling in any given cell's receptive field (RF) are constantly changing. Adaptive mechanisms modulate the filtering properties of the pathway in order to maintain sensitivity and signal to noise ratio in the face of these changes.

Two distinct types of adaptation have been observed in the early visual pathway, termed *light adaptation* and *contrast adaptation/gain control*. These mechanisms alter encoding properties such as gain and bandwidth on two primary timescales. A rapid adaptation operating over milliseconds shifts the operating range of the system in response to a change in mean luminance or contrast [15, 29]. A second slower adaptation responds to prolonged exposure to fixed luminance or contrast over the course of seconds [1, 11, 34]. These phenomena were originally observed in the retina and have recently been reported in the thalamus [22, 23, 27] and

[1] Author to whom any correspondence should be addressed.

0954-898X/03/010119+17$30.00

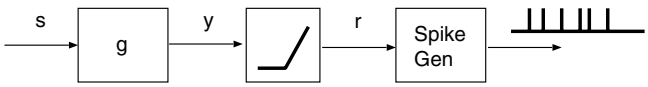

Figure 1. Wiener system model of neural encoding.

cortex [22, 28]. Although adaptive phenomena in the early visual pathway have been studied for many years, the role of adaptation in natural vision is not well understood. One theory suggests that these mechanisms serve to decorrelate redundant natural scenes and maximize signal energy in the frequency band where the signal to noise ratio is highest [2] and recent studies have shown that some adaptive mechanisms may aid in maintaining a constant rate of information flow in the face of changing stimulus statistics [16].

Reverse correlation techniques have been used to analyse adaptation either after the adaptation reaches a steady state [11], or across multiple stimulus trials [34]. However, in a natural setting steady state may never be reached and multiple trials may not be available. The use of time invariant analysis techniques in a nonstationary setting results in a system estimate that describes the average behaviour over the trial, neglecting adaptive changes in the encoding scheme. In this paper, an adaptive reverse correlation (ARC) approach is developed to capture time varying characteristics of the spatiotemporal encoding mechanisms in the early visual pathway during single trials. Adaptive filter estimation is a well-established area of engineering and statistics, but has only recently been applied to capture nonstationary features of neuronal spike trains and the time varying nature of stimulus/response relationships [10, 35]. We apply the ARC technique to both simulated and experimentally observed neuronal response data to demonstrate the properties of the algorithm and its ability to track dynamic properties on a variety of timescales at experimental noise levels. Spatiotemporal binary noise stimuli with varying mean luminance and contrast were used as input to retinal adaptation models and for single electrode extracellular recordings in cat lateral geniculate nucleus (LGN) and primate LGN and visual cortex. The results demonstrate the ability of adaptive estimation to track changes in dynamic filtering properties that may be important during natural vision.

2. Methods

2.1. Encoding model and experiments

For many neurons in the early visual pathway, the firing rate can be described as a quasi-linear function of the spatiotemporal stimulus. More specifically, this encoding process can be modelled as a cascade of a linear spatiotemporal filter (kernel, RF) and a static nonlinearity [12], often referred to as a Wiener system, as illustrated in figure 1. The visual stimulus s is convolved with the linear spatiotemporal kernel g to produce intermediate output y, which is then passed through the rectifying static nonlinearity to yield the firing rate r. This firing rate can be envisioned as the rate parameter driving a probabilistic spike generating mechanism. The fundamental construct of the adaptive estimation described in this paper involves temporally local estimates of stimulus/response relationships based on the above model.

Data from experiments using anesthetized cats were analysed in this framework. Computer controlled spatiotemporal binary noise with fixed luminance and contrast was presented and the resulting neuronal response was recorded extracellularly using a multi-electrode array from X cells in the LGN. The spatial extent of the stimulus grid was significantly larger than the classical RF. The temporal frequency was 128 Hz and the width of each square in the grid was either 0.2° or 0.4°, depending on the optimal spatial frequency of the cell. We also

conducted one experiment on an anesthetized macaque monkey. Spatiotemporal binary noise stimulus was again used and neuronal responses were recorded extracellularly with a single electrode. In this experiment, the mean luminance and contrast of the stimulus were varied. The stimulus was confined to the classical RF while the rest of the monitor was uniform at the mean luminance of the noise. The spatial frequency of the stimulus was set to four times that of the optimal grating for the cell. The monitor refresh rate for the primate was 100 Hz. Details of the surgical and experimental preparations for the cat and primate are given in [36] and [23], respectively. Neuronal recordings of spike times were binned to produce a rate representative of r, and from this continuous process the underlying kernel g was estimated as it evolved over time, as described in the next section.

2.2. Adaptive reverse-correlation (ARC)

To track adaptive changes in the linear kernel of the Wiener system, we have developed a recursive variation on frequency domain reverse correlation [19]. At each time step, the current estimate is updated using new data obtained since the last step, while results from previous estimates are downweighted. Given zero mean white noise stimulus (s, light intensity) and response (r, neuronal firing rate) data of length D, reverse correlation in the frequency domain is performed as follows. For a desired kernel estimate that is L samples long, the data are partitioned into $M = D/L$ segments. For each segment, the fast Fourier transform (FFT) is applied to the stimulus and response signals to obtain S and R. Next, the auto- and cross-spectral densities $\Phi_{SS} = SS^*$ and $\Phi_{RS} = RS^*$ are computed (* represents complex conjugation). Since the stimulus is uncorrelated, the relevant quantity is its power level, P, which is a scalar equal to the mean of the auto-spectral density. The transfer function estimate for each segment of data is estimated by $G = \Phi_{RS}/P$. When the above algorithm has been applied to the M segments, the transfer function estimates are averaged to give the overall transfer function estimate for the data set. The total transfer function estimate for the M segments is given by

$$G = \frac{1}{M}\sum_{m=1}^{M}\frac{\Phi_{R_m S_m}}{P_m}$$

where $\Phi_{R_m S_m}$ and P_m are the cross-spectral density and input power level of the mth segment, respectively. The inverse FFT of G is taken to obtain the kernel estimate g. It has been shown that the static nonlinearity has the effect of scaling the correlation structure [35]. Under the conditions of these experiments, this scaling can be approximated by multiplying the estimate by a factor of two. The estimate is also improved by zero padding to eliminate wraparound errors [19].

We have developed a novel approach in which the frequency domain algorithm was extended to provide an adaptive estimate of the kernel. At any point in time, the kernel can be computed as a weighted sum of transfer function estimates from past segments. When new data are available, new cross-spectral and power level estimates are calculated. The resulting new transfer function estimate is added to the current sum while older terms are downweighted, so that the kernel estimate reflects the current state of the system. This process can be written as a simple recursion. The total transfer function estimate at time n is denoted by G^n. Assume the estimate is updated every δ samples and the desired kernel length is L samples. At time $n+\delta$ new data segments are formed with the data from time $n+\delta$ back to $n+\delta-(L-1)$ (note that choice of δ and L determine the separation/overlap of successive segments). Using these

new segments the transfer function estimate is updated according the following formula:

$$G^{n+\delta} = \frac{1}{1+\lambda}(G_{n+\delta} + \lambda G^{n})$$

where $G^{n+\delta}$ is the new transfer function estimate using all the data up to and including the time $n + \delta$, λ is the 'forgetting factor' used to downweight past data ($\lambda \in [0, 1]$) and $G_{n+\delta}$ is the transfer function estimate from the data in the segment formed at time $n + \delta$. The inverse FFT of $G^{n+\delta}$ is taken to obtain the kernel estimate $g^{n+\delta}$. For spatiotemporal stimuli, the above quantities correspond to matrices and the kernels mapping each pixel of the stimulus to the response are estimated simultaneously.

2.2.1. Choice of parameters. The performance of the ARC algorithm is greatly affected by the choice of δ and λ. An analysis of the system in question must be made to determine the optimal set of values for these parameters. For offline analysis, the estimate should be updated every sample ($\delta = 1$). For online analysis, it may be more computationally efficient to update less frequently, especially if the system in question is adapting slowly. Ideally, the combination of δ and λ would produce an estimate with memory comparable to the fastest time constant of adaptation in the system. However, in experimental situations, the selection of estimate memory involves a tradeoff between adaptability and robustness. For an estimate with a short memory, a small amount of data is used. This allows the estimate to track a quickly adapting system (because the behaviour is averaged over a small part of the data set at any given time), but the estimate may suffer from noise problems due to the limited data used in the computation. Conversely, an estimate with a long memory may average over short adaptations but the additional data will make the estimate more robust with respect to experimental noise.

Empirical testing determined that the minimum memory which produced reliable estimates at our experimental signal to noise ratio was roughly twice the integration time of the system ($2 \times L$). Some of the adaptive mechanisms that have been observed in the visual system operate on a faster timescale than this. For example, contrast gain control in the retina may have a time constant of 50 ms, while the cells may have an integration time of 150 ms. However, adaptive changes can still be followed during saccades around a natural scene, due to the stationarity of the stimulus between saccades. If the average intersaccade interval is 350 ms, then 50 ms of this interval will be spent adapting from the previous state to the new state, while 300 ms will be spent in the new state. If an estimate with a memory of 300 ms ($2 \times L$) is used, the dynamics of the state transitions will be missed, but the state changes in the system can still be tracked. In fact, no functional difference was observed in contrast gain control models where the state transitions were taken to be instantaneous [20].

If we define memory as the duration for which a given data segment retains more than 37% of its initial weight in the overall estimate, then we can quantify the memory in terms of a number of estimate updates u with the formula $0.37 = \lambda^{u}$. Multiplying u by the update size δ yields the memory of the system. For the analyses here, we set δ to one sample, and matched λ to the time constants of adaptation in the systems we tested (as reported in previous studies). For cases where this time constant was less than twice the integration time of the system, the memory was set to $2 \times L$. For calculations in adaptation model simulations, a forgetting factor (λ) of 0.97 and kernel update (δ) of 10 ms (one sample) were used. These values yield an estimate with 320 ms of memory. For analysis of the cat data, $\lambda = 0.995$ and $\delta = 8$ ms were utilized to yield an estimate with 1.6 s of memory. For the primate studies, $\lambda = 0.98$ and $\delta = 10$ ms produced an estimate with 500 ms of memory.

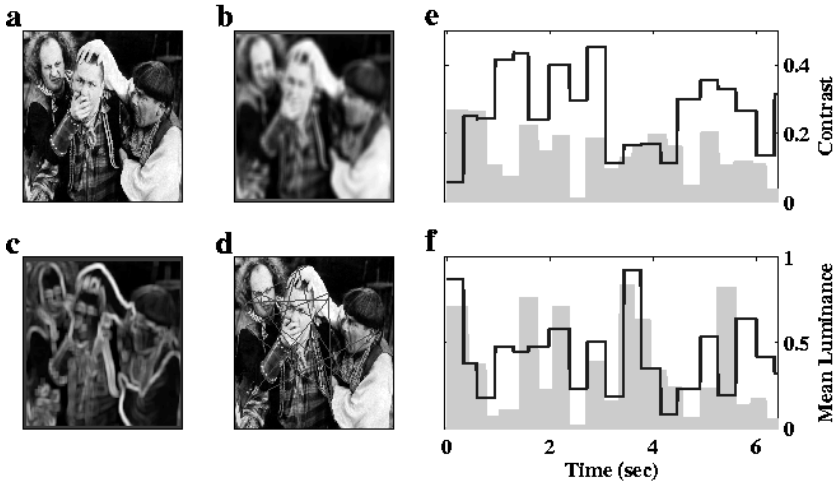

Figure 2. Simulated saccade path. (a) An example of a natural scene image. (b) The image shows the corresponding local mean luminance. (c) The image shows the corresponding local contrast measure. (d) Simulated saccade path across the image. (e) Local contrast measure over time for the saccade model based on contrast (thick curve) and for a uniformly distributed saccade path (shaded). (f) Same for the mean luminance. For this simulation, $m = 0.2$.

2.3. Natural scene simulation

The ultimate goal of investigation in the visual pathway is to understand its function in the natural environment, where sudden mean luminance and contrast changes accompany saccades and the movement of objects in and out of the RF. To simulate dynamic viewing, binary noise was generated, with mean luminance and contrast modulated according to a saccade path across a natural scene. A set of grey-scale images was used with intensities ranging from 0 to 1, with 256 grey levels. Spatial scale was taken to be 25 pixels/degree. Local measures of contrast were taken over $1^\circ \times 1^\circ$ square patches, and were computed as follows [24]:

$$C_{ij} = \frac{\left[\sum_{k,l \in D_{ij}} (I_{kl} - \bar{I}_{D_{ij}})^2\right]^{\frac{1}{2}}}{(2n+1)^2} \qquad D_{ij} = \{I_{kl} : |k-i| < n, |l-j| < n\}$$

where C_{ij} is the local contrast of the patch centred at pixel (i, j), I_{kl} is the intensity of the pixel at position (k, l), D_{ij} is the patch centred at pixel (i, j), $\bar{I}_{D_{ij}}$ is the mean intensity of the patch centred at (i, j) and n is the half-width of the patch in pixels ($n = 12$ for this case). Reinagel and Zador [24] found that these local measures of contrast were highly correlated with saccade target during 'free' viewing of natural scenes. Simulated saccade paths over natural scenes can thus be realized through probability models based on local contrast. Specifically, as a simple model, suppose that at each saccade, the contrast level is drawn from an exponential distribution ranging from high to low contrast. Once generated, the target location of the saccade is chosen from a random distribution of patches with contrast falling in a bin defined by the exponential draw. Specifically, $C_k \sim \exp(m)$, where C_k is the randomly chosen contrast of the kth saccade, and $\exp(m)$ is the exponential distribution with parameter m. A sample path of such a construct is shown in figure 2.

The time interval between saccades was taken to follow a Gaussian distribution, with the contrast and mean luminance described in the following manner:

$$C_s(t) = C_k \qquad \text{and} \qquad \bar{I}_s(t) = \bar{I}_k \qquad \text{for } \tau_k \leqslant t < \tau_{k+1}$$

Figure 3. Binary noise sequence simulated from natural scene statistics. The stimulus consists of a 25 × 25 pixel binary noise sequence at 100 Hz, where the mean luminance and contrast are dictated by the saccade simulation described in the text. Frames are shown only at transitions to new mean luminance and contrast.

where the intervals between saccades were defined as $T_k = \tau_{k+1} - \tau_k \sim \mathcal{N}(\mu_T, \sigma_T^2)$, where $\mathcal{N}$ represents a Gaussian distribution, with mean $\mu_T = 350$ ms and standard deviation $\sigma_T = 50$ ms [38]. Since the standard deviation is small relative to the mean, the chances of an interval being negative are very small. However, to avoid this problem, the Gaussian was truncated at five standard deviations. This Gaussian model has successfully captured the saccade dynamics of a macaque monkey freely viewing a natural scene [40]. For this simplistic model, the saccades were considered instantaneous. A realization of local contrast over time for this model is shown in figure 2(e). For reference, a time series of contrast for a saccade model in which the target location is drawn from a uniform distribution on the image is shown with the shaded region, illustrating the non-uniformity of saccade target locations. Similar plots are shown for the local mean luminance in figure 2(f).

The simulated contrast and mean luminance levels were utilized to generate binary noise restricted to an assumed $1^\circ \times 1^\circ$ square patch representative of the classical RF of a cell, frames of which are shown in figure 3.

3. Results

The early visual pathway adapts to changes in the mean luminance and contrast of the visual world. In a natural setting, the statistics of the light falling within a given cell's RF are changing with each saccade, prompting corresponding changes in the filtering characteristics of the system. These changes prevent the system from reaching a steady state and thus an adaptive analysis technique must be employed. The ARC approach presented here allows for observation of changes in system dynamics over short trials. The tracking ability of the ARC algorithm was tested in a number of settings. Simulations of retinal adaptation models were used to probe the properties of the algorithm. The results were compared to those obtained by standard reverse correlation. Data from the cat and primate visual systems were analysed to demonstrate the adaptive estimation of filtering properties such as kernel peak, latency and bandwidth over single experimental trials.

3.1. Simulations

Adaptive mechanisms operating on the millisecond timescale serve to rapidly shift the operating range of the visual system in response to a changing stimulus. Neurons in the early visual pathway exhibit adaptation to mean luminance levels (*light adaptation*) and contrast levels (*contrast adaptation/gain control*). Initial work in adaptation centred on functional mechanisms in the retina, but further work has shown its presence in the thalamus [22, 23, 27] and cortex [22, 28] as well, so it is likely that the general adaptive properties are ubiquitous in the visual pathway. These properties are illustrated through the behaviour of the light adaptation and contrast gain control models (see the appendix). One effect of the adaptive mechanisms is to attenuate the response at high mean luminance or contrast, which is illustrated in figure 4(a) for the light adaptation model and figure 4(d) for the contrast gain model. For comparison,

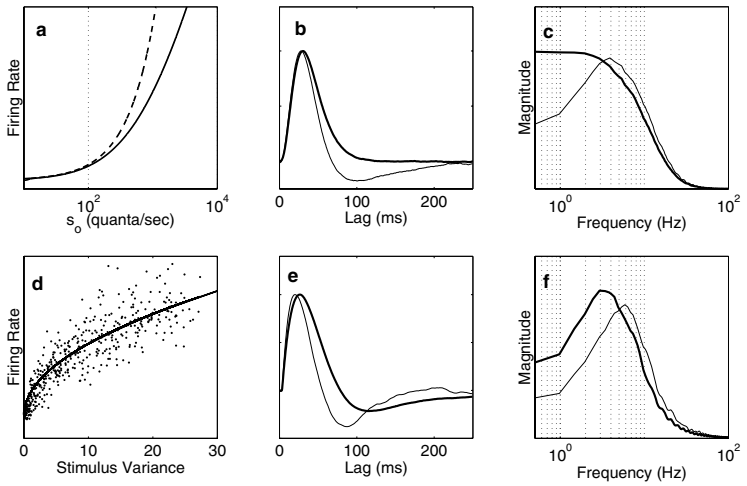

Figure 4. Adaptation model characteristics. Subplots (a)–(c) describe the light adaptation model. (a) Steady state firing rate as a function of luminance for adaptive (full) and non-adaptive (broken) models. (b) First-order kernel estimates for pre- (thick) and post-adapted (thin) states; kernels scaled to identical peak values, post-adapted state attenuated by a factor of 10. (c) Transfer function magnitudes for pre- (thick) and post-adapted (thin) states, where each kernel was first normalized: $\bar{g} = g/\|g\|$. Subplots (d)–(f) describe the contrast gain control model. (d) Mean firing rate as a function of stimulus variance. (e) Kernel estimates at low (thick) and high (thin) contrast, normalized to illustrate the temporal compression. (f) Corresponding transfer function magnitudes, where each kernel was first normalized: $\bar{g} = g/\|g\|$.

the linear relationship between the background luminance and the firing rate is shown with the broken curve in 4(a). This suggests, at minimum, a static nonlinearity in the relationship between stimulus and response. Beyond the simple static nonlinearity, comparison of first-order kernels at different luminances/contrasts also suggest a fundamental difference in the temporal dynamics of the system as seen in figures 4(b) and (e) for the light adaptation and contrast gain models, respectively. The thick curve represents the kernel at low background luminance/contrast, while the thin curve is the kernel at higher background luminance/contrast. Both have been scaled to eliminate the obvious decrease in magnitude with increased mean luminance/contrast, revealing the temporal 'compression' of the kernel and enhanced biphasic response at increased background luminance. Figures 4(c) and (f) present the magnitude of the transfer function of the pre- and post-adapted states (thick and thin curves, respectively). To obtain magnitude characteristics on the same scale, the kernels were first normalized by their overall power: $\bar{g} = g/\|g\|$. The pre-adapted state (low luminance/contrast) is low pass in nature, while the adaptation to increased luminance/contrast effectively produces a band pass characteristic, also slightly increasing the high-frequency cutoff. While these models describe adaptive function in the retina, thalamic and cortical phenomena are very similar.

The fast adaptation described by the above models serves to shift the system to its new operating range quickly. A second slower adaptation modulates the system properties over the course of seconds. A slow form of luminance adaptation has been noted in rabbit [21] and primate retinas [39]. This mechanism decreases the magnitude and integration time of the neural response after a step change in luminance with a time constant of the order of 10 s. Studies of slow contrast adaptation in the primary visual cortex of the cat [1, 22] and the macaque monkey [28] showed a similar decrease in response amplitude and increase in differential sensitivity to sustained contrast on the timescale of seconds. Initial investigation

revealed limited adaptation effects in the geniculate, but recent studies have shown significant adaptation [32], with a similar time course to that observed in the cortex [26]. In the retina, slow contrast adaptation has been observed in the salamander, rabbit and primate [11, 18, 34]. All forms of slow adaptation exhibit the same decrease in response magnitude, accompanied by a compression of kernel dynamics, in response to sustained mean luminance/contrast.

With adaptive mechanisms operating all along the visual pathway, experiments may be designed to maintain constant stimulus statistics during the course of a trial (to eliminate fast adaptation) and the first segment of a data set may be discarded (to eliminate slow adaptation). While such measures are valid under artificial experimental conditions, the stimulus falling onto particular photoreceptors during natural vision is constantly varying. To understand the visual function in such a nonstationary environment, where multiple trials may not be available, an adaptive system identification technique is required. Using the ARC technique, stimulus induced modulations in the encoding dynamics can be tracked. The models described above were used to simulate an experiment where the mean luminance or contrast of a binary noise stimulus were stepped up at 5 s increments. The firing rate output of the model was used to drive an inhomogeneous Poisson spike generator, introducing noise into the simulation that is comparable to that encountered experimentally. The results of these simulations are summarized in figure 5. Figures 5(a)–(h) show the results for the light adaptation model and figures 5(i)–(p) show the results for the contrast gain model. The visual stimulus was binary noise at 100 Hz. The output of the model describes the response of the cell to the appearance of this stimulus in the centre of its RF. Figures 5(b) and (j) show the decrease in magnitude and latency that result from increases in luminance/contrast. These changes were tracked throughout the trial and can be observed in the top–down view of the kernel estimates (figures 5(c) and (k)), the plot of the kernel peak throughout the trial (figures 5(d) and (l)) and the plot of the latency throughout the trial (figures 5(e) and (m)). Kernel estimates from standard and ARC were used to predict the response to a subsequent trial (different realization of the binary noise, same changes in mean luminance and contrast). Although the adaptive estimate uses more parameters, noise fitting effects are eliminated by computing and testing the estimates on independent data sets. These predictions are compared to the output of the model in figures 5(f)–(h) and (n)–(p). Using the ARC estimate results in a much lower prediction error as expected. The adaptation described by the models operates on a timescale that is close to the integration time of the system. In order to track the adaptation on this timescale, only a small amount of data is used in the estimate. This results in the expression of experimental noise in the estimate (as discussed in the methods) as evidenced in the early part of figure 5(h). While the adaptive estimate is noisy, it is able to track the system as it changes state, resulting in better overall performance. For instance, the standard reverse correlation overestimates the kernel at high luminance/contrast, evidenced by the relatively large response magnitude at the end of the trial (figures 5(g) and (o)). This is due to the fact that standard reverse correlation weights data from the whole trial equally when forming an estimate, including when the luminance/contrast is low and the kernel is large. The adaptive estimate is able to more accurately reflect the current state of the system at any point during the trial by focusing on data in the recent past.

The above simulations demonstrate the ability of the adaptive approach to capture the time-varying nature of the encoding mechanisms in response to step changes in mean luminance and contrast. In a natural setting, the local mean luminance and contrast are constantly varying as the eyes move across the visual field and objects move in and out of view. Natural scenes have been utilized as visual stimuli in a number of recent studies, but are not ideal for spatiotemporal kernel estimation due to the spatial and temporal correlation structure inherent to the natural environment. While white noise stimuli excite all modes of a system, a band-limited natural

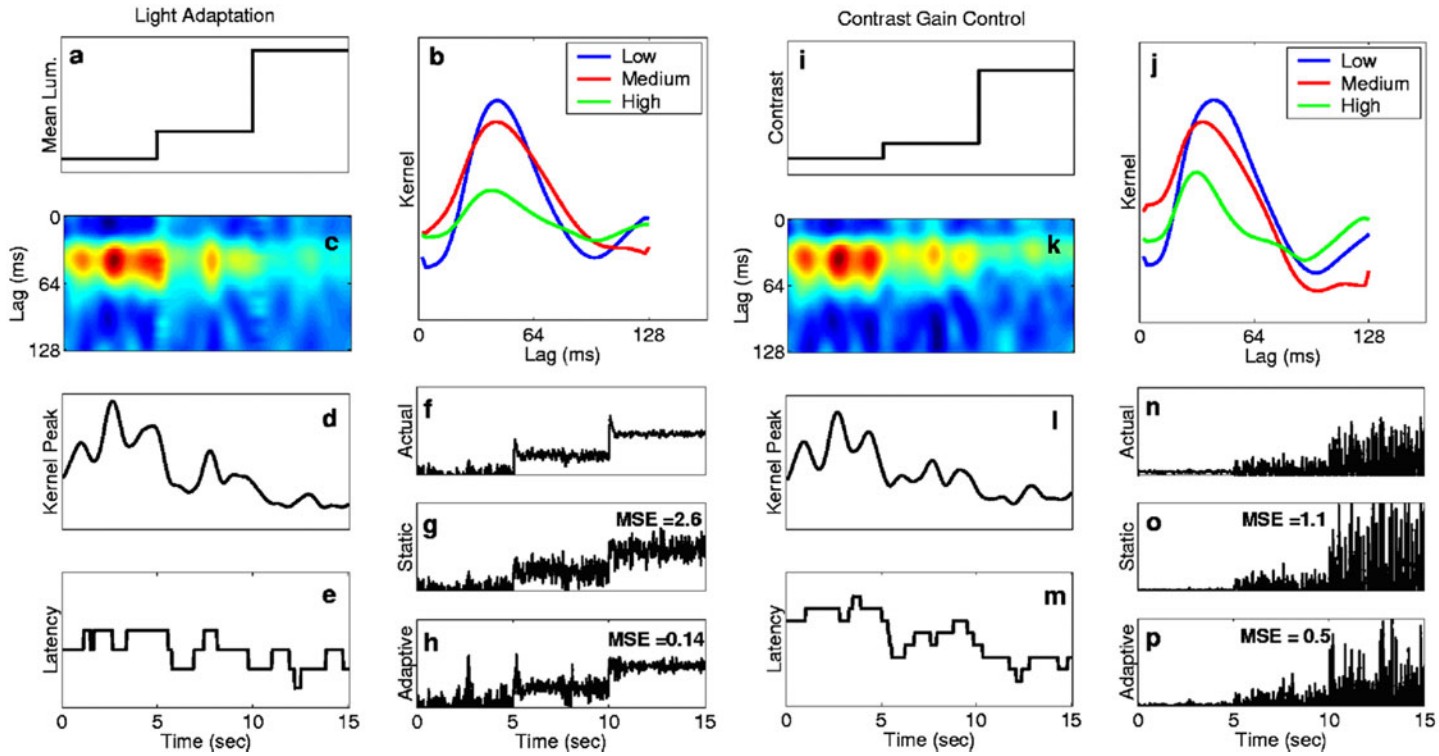

Figure 5. Adaptive estimation of model response to step changes. Results for the light adaptation model are given in (a)–(h). Results for the contrast gain model are given in (i)–(p). (a), (i) The values of the relevant features of the binary noise stimulus for the two simulations, (a) shows the value of the mean luminance throughout the trial, (i) shows the value of the contrast throughout the trial. (b), (j) The kernel at the three time instants during the trial corresponding to the three values of luminance/contrast. (c), (k) A top–down view of the kernel estimate throughout the trial. Maximum values are indicated by red, minimum values by blue. (d), (l) Attenuation of the kernel peak (spikes s^{-1}) in response to step changes in mean luminance/contrast over the stimulus trial. (e), (m) Decrease in latency (ms) corresponding to changes in luminance/contrast. (f), (n) The firing rate (spikes s^{-1}) produced by the model. Kernel estimates were used to predict the response of the system for comparison. The predicted rates for standard reverse correlation (g), (o) and ARC (h), (p) are shown with the mean square error (MSE) of their predictions. MSE was calculated as $\text{mean}((r-\hat{r})^2/\text{var}(r))$, where r is the response of the model and $\hat{r}$ is the predicted response.

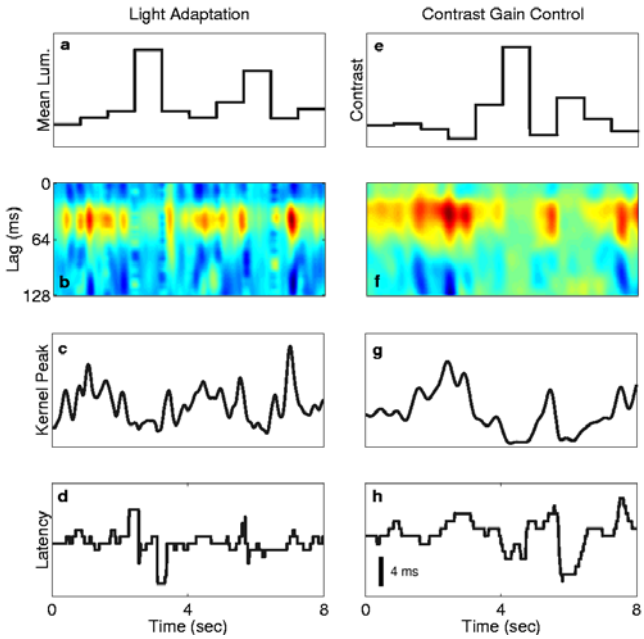

Figure 6. ARC tracking of model responses to natural scene stimulus. (a), (e) Mean luminance and contrast of the binary noise. (b), (f) A top–down image of the kernel estimate throughout the trial; red represents maximum values, blue represents minimum values. (c), (g) Tracking of the kernel peak over the stimulus trial. (d), (h) Tracking of the latency throughout the trial.

scene may not. To emulate the natural environment for the reverse correlation studies, spatiotemporal binary noise was generated in which the mean luminance and contrast were varied according to a saccade model (see methods). The resulting stimulus captures the luminance and contrast dynamics that are characteristic of natural scenes, while providing a rich stimulus for kernel estimation. Applying ARC in this emulated natural setting demonstrates the ability of the technique to track the encoding properties of the system as they are continuously modulated. The results of the adaptive estimation for a simulation involving the light adaptation and contrast gain control models are shown in figure 6. Note that the kernel adapts to a new state with each saccade to a new area in the natural scene. The behaviour described by the earlier simulations, namely attenuation and compression at higher luminance/contrast values, is observed. Each saccade induced change in the stimulus leads to a corresponding change in the system, as shown in the top–down view of the kernel estimates (figures 5(b) and (f)). The adaptation in kernel peak and latency throughout the trial are shown explicitly in figures 5(c), (g) and (d), (h), respectively. Again the estimate for the light adaptation model suffers from noise problems. For both models, the kernel peaks are tracked well, as increases in luminance/contrast correspond to decreases in peaks. The latency estimate for the contrast gain model is clean, with higher contrasts corresponding to lower latencies. The latency estimate for the light adaptation model, however, is noisy (figure 5(d)), making it difficult to observe distinct state transitions.

3.2. Experiments

To test the performance of the ARC algorithm under experimental conditions, we analysed data from the cat and macaque visual systems. All trials involved single electrode extracellular recordings of the response to binary spatiotemporal noise as described in the methods.

3.2.1. Cat LGN. The response of cat LGN X cells to a 16×16 grid of spatiotemporal binary noise was recorded. At the start of the trial, the stimulus appeared at 100% contrast. The firing rate of a cell was calculated by binning spikes in 7.8 ms bins. The ARC algorithm was applied to the data to adaptively estimate the linear kernel over the first minute of the trial. The results for typical on and off cells are shown in figure 7. The adaptive kernel estimates are shown in figures 7(a) and (b). Specifically, the kernel exhibits a clear decrease in response magnitude (peak for the on cell, trough for the off cell) over the first 10–30 s of the trial, shown in figures 7(e) and (g). Additionally the kernel showed a sharpening in bandwidth over the same time course. The frequency response at the start and end of the trial is shown in 7(c) and (d) and a continuous plot of the bandwidth is shown in 7(f) and (h). The bandwidth was calculated as the range of frequencies for which the transfer function magnitude was greater than 50% of the peak value.

Thus far, the discussion has focused on adaptation in the temporal aspects of the encoding process, in terms of the centre of the RF. The ARC technique is well suited for tracking changes in the spatial RF as well. Plotting the spatial RF at the peak of the temporal kernel at different time instants throughout the experimental trial reveals the modulation of spatial RF properties, illustrated in figure 8. For this example, the spatial distribution flattened throughout the trial. Figures 8(a)–(d) illustrate the spatial RF map at the peak in temporal response at 8 s intervals over the stimulus trial. The RF map is normalized by the peak of the centre pixel at each time slice. Figures 8(e) and (f) show the pre- (thick) and post-adapted (thin) kernels for the centre pixel and a typical pixel outside the RF centre, respectively. The extent of the spatial RF in figure 8(a) appears to increase due to the fact that the magnitude of the centre pixel shows a greater decrease than that of the off centre pixel over the course of the trial, which is more clearly shown in figures 8(e) and (f).

3.2.2. Macaque LGN and V1. A similar experiment was carried out in the early visual pathway of the macaque monkey. The experiments in the cat LGN allowed us to track adaptation to fixed luminance and contrast. For the macaque, the mean luminance and contrast of the binary noise stimulus were changed at fixed intervals so that adaptation to a nonstationary stimulus could be observed. A 10×10 grid of binary noise was presented in the classical RF. The ARC algorithm was used to estimate the kernel while the luminance and contrast were varied. Results for a cortical simple cell and magnocellular LGN cell are shown in figure 9. Typical kernel estimates for the RF centre of each cell are shown in figures 9(a) and (f). Over the 20 min trial the contrast was stepped up and down over a range of five values at 30 s intervals, while the mean luminance was stepped over a range of five values at 4 min intervals (after an initial period of full contrast stimulation). This allowed us to observe adaptation due to changes in both mean luminance and contrast. Figures 9(b) and (g) show the stimulus profile and 9(c) and (h) show the corresponding kernel magnitude. The oscillations in kernel magnitude are due to contrast adaptation, while light adaptation causes a slow rise in the mean value. A closer look at an interval of the trial where the mean luminance was fixed shows more details of contrast adaptation. Figures 9(d) and (i) show stimulus contrast stepped over four values. Figures 9(e) and (j) show the changes in kernel peak over this interval. Note the exponential decay that accompanies each change in contrast. These plots illustrate the ability of the ARC algorithm to capture adaptive phenomena over a single trial in the absence of steady state and are in agreement with previous studies that relied on multiple trials.

4. Discussion

The ARC approach presented here represents a general framework that is useful in understanding the complex mechanisms that control the sensory encoding process in response

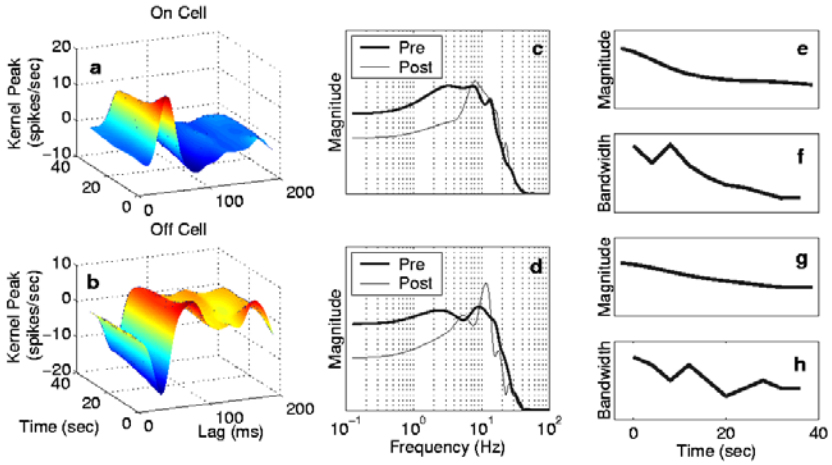

Figure 7. Continuous tracking of adaptation in LGN X cells. The temporal kernel between stimulus and firing rate for the centre of the RF is plotted as a function of the time since stimulus onset (seconds) for an on cell (a) and an off cell (b); red indicates positive values, blue negative. The magnitude of the transfer function of the system at the beginning (thick) and end (thin) of the trial are shown in (c) and (d) for the on and off cells, respectively. Note that the magnitudes are normalized to emphasize the change in bandwidth. Continuous tracking of kernel magnitude (e), (g), along with plots of the bandwidth throughout the trial are shown for each cell (f), (h). The bandwidth was calculated as the range of frequencies for which the transfer function magnitude was greater than 50% of the peak value.

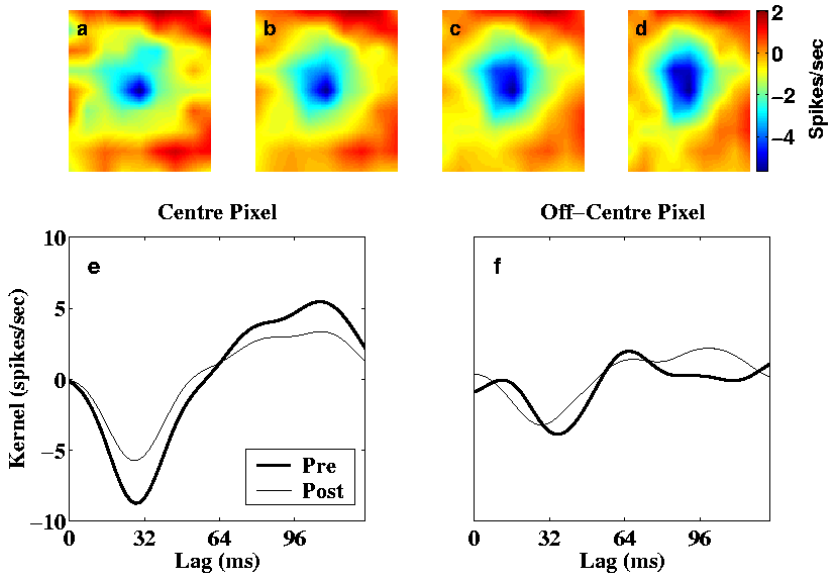

Figure 8. Adaptation of spatial properties of LGN X cell RF. (a)–(d) Spatial RF at peak in temporal kernel at 8, 16, 24 and 32 s after stimulus onset (the region is approximately 1.8° of visual space). Blue represents minimum values, red maximum values. The RF map is normalized by the peak of the centre pixel at each time slice. (e) Pre- (thick) and post-adapted (thin) kernels at RF centre. (f) Similar plot for a typical pixel outside the RF centre.

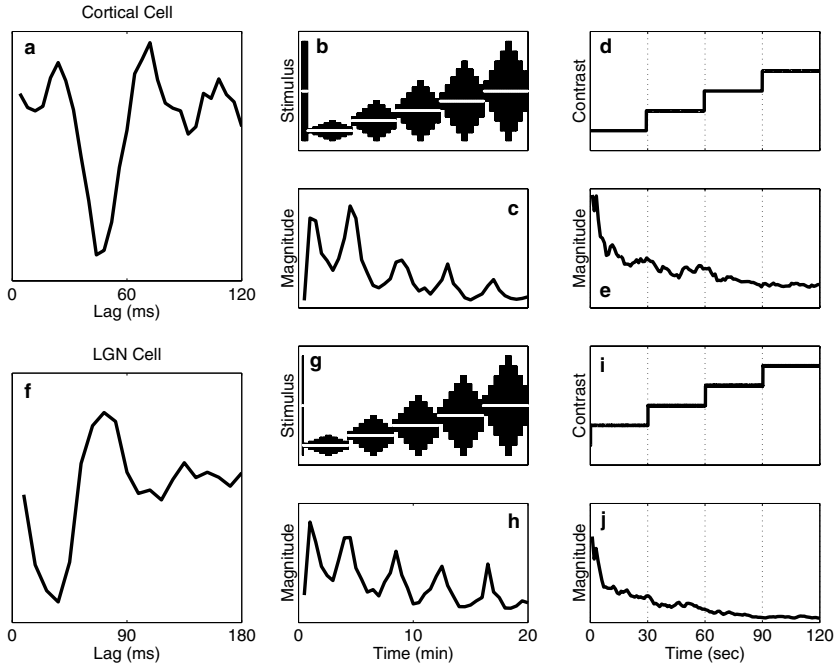

Figure 9. Adaptation to changing luminance and contrast in the primate visual system. Typical first-order kernel estimate (RF centre) for cortical (a) and LGN (f) cells. (b), (g) The stimulus profile throughout the 20 min trial. The mean luminance is superimposed in white. For each mean luminance value, the contrast was stepped up and down a series of five values. (c), (h) Tracking changes in the kernel trough over the trial. (d), (i) One interval of the trial in which contrast was stepped up over four values. (e), (j) Tracking changes in kernel peak over the short interval shown in (d), (i).

to natural stimuli. The formulation used here was based on a spatiotemporal kernel estimation framework in the frequency domain as an illustration, but other formulations exist. We have previously formulated the adaptive estimation in the time domain, based on recursive least squares techniques [35]. Frank *et al* [10] have developed adaptive approaches to neural RF estimation within a point process framework for capturing encoding properties in the hippocampal place cells. This is likely to be important in pathways exhibiting coding mechanisms that cannot be well described through rate coding. More recently, the recursive approach to RF estimation was also implemented in capturing encoding properties in cortical response to natural scenes, but was implemented for efficiency of estimation and not presented within an adaptive framework, although tracking time-varying properties is mentioned as a possible benefit of the approach [25]. These studies represent an important direction in understanding the role of adaptive mechanisms in natural sensory function.

The ARC approach successfully captured changes in encoding dynamics in both simulated and experimental data. Modulation of properties such as kernel peak, latency and bandwidth were observed. Simulations using light adaptation and gain control models demonstrated the ability of the algorithm to robustly track the characteristic changes observed in the early visual pathway that are neglected by time invariant techniques. Results were noisy when the adaptive time constant of the system approached the integration time of the cell, especially for the light adaptation model (figures 5(h) and (d)). Our results indicate that the minimum time constant of adaptation that can be faithfully tracked with this approach at experimental signal to noise

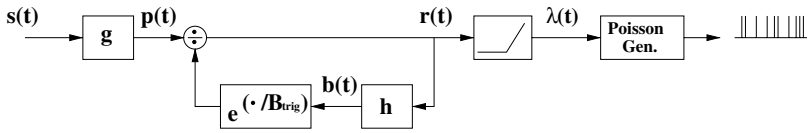

Figure 10. Light adaptation model adapted from Enroth-Cugell and Shapley [15].

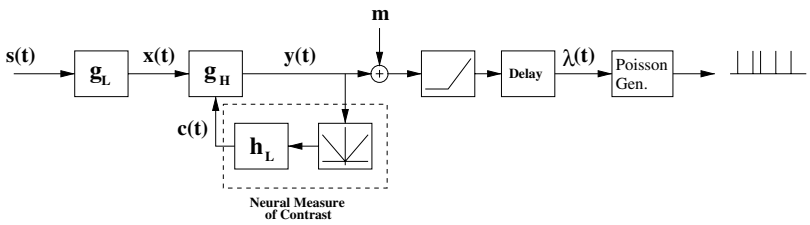

Figure 11. Contrast gain control model after Victor [37].

ratios is twice the integration time of the system. However, faster systems can be tracked if the stimulus is relatively stationary over an interval that is twice the integration time of the system, as is often the case in natural vision. Analysis of data from single cell responses in cat and primate yielded results that agree with those previously observed. Studies conducted at various points in the visual pathway [1, 15, 23, 34] have demonstrated similar adaptive phenomena such as attenuation and compression of temporal kernels at high luminance/contrast. These studies have relied on experimental paradigms involving multiple trials and steady state analysis. Through ARC, we have succeeded in capturing the same adaptive phenomena during single trials using a nonstationary stimulus.

While the different forms of visual adaptation may be isolated under artificial experimental conditions, it is clear that their interaction is crucial in a natural setting. The visual world exhibits variations in contrast and mean luminance over a wide range of spatial scales; the local variance is often very well correlated with local mean luminance, implying that adaptive mechanisms may be difficult to decouple in the natural environment. Recent studies have characterized the nature of spontaneous saccades in the natural environment [17, 38], and have found that saccade targets are indeed highly correlated with local contrast and spatial correlation structure [24]. Our studies in the visual system of the primate demonstrate the dependence of system characteristics on both luminance and contrast for a stimulus where both are varied. In addition to interaction between different forms of adaptation, both light and contrast adaptation have been shown to operate on multiple timescales. If the memory of the estimate is matched to the fastest adaptation in the system, then adaptation on all timescales can be tracked. If the different adaptive time constants are close to one another, their effects may be coupled. In the early visual pathway, the two major adaptive time constants that have been observed are separated by two orders of magnitude. This may allow them to be tracked independently while operating simultaneously, as the faster adaptation will be relatively instantaneous. Further investigation in this direction may lead to an understanding of how different adaptive mechanisms combine to accomplish natural vision. This approach could also be used to answer questions that have recently arisen regarding spatial interaction in natural scene responses [38].

The early visual pathway is a spatiotemporal filter mapping light varying over several orders of magnitude into spikes for transmission to higher areas. Clearly, only an adaptive

system could maintain differential sensitivity over such a range of inputs. Since the visual input only occupies a small portion of this range at any time, we would expect efficient visual encoding mechanisms would adapt their strategies to the full shape of the luminance distribution [34]. In addition to maintaining sensitivity in the system a number of other roles for adaptation have been proposed. One theory that has been tested experimentally is that the early visual pathway decorrelates redundant natural scenes [3]. It has also been suggested that the role of adaptation is to provide efficient coding [2, 4, 5, 33] of the visual world by maximizing information transmission, or at least maintaining information transmission in the face of perturbations to a system with limited dynamic range [9, 16]. To test such theories, the system must be observed in the environment in which it naturally operates. Steady state analysis techniques applied under such conditions would describe the average behaviour of the system over some range, neglecting important adaptive changes. Adaptive estimation provides a tool for observing these changes and understanding the role of adaptation in natural vision.

Acknowledgments

We thank Yang Dan for the use of the cat LGN data and Dan Pollen and Andrzej Przybyszewski for helping us to collect the primate data in their laboratory at the Department of Neurology at the UMass Medical Center.

Appendix

The early visual pathway modifies its encoding strategy based on the mean luminance of the visual stimulus. The most obvious effect of light adaptation is the modulation of retinal sensitivity. By varying its gain inversely with stimulus strength [7, 15], the retina is able to maintain differential sensitivity over many orders of magnitude. In addition to sensitivity, the integration time of the response decreases with luminance level [13, 15], while the relative strength of the surround region of the spatial RF increases [6, 14]. This allows the retina to produce a significant response at low luminance levels, while maintaining sensitivity to temporal and spatial contrast at high luminance levels. The fast component of the adaptation occurs a few hundred milliseconds after a change in luminance [8, 15]. The work of Enroth-Cugell and Shapley [15] experimentally characterized the nature of this phenomenon and developed a light adaptation model. In the first stage of the light adaptation model, the stimulus $s(t)$ is transduced into a photocurrent $p(t)$ by the photoreceptors in a linear fashion: $p(t) = \int g(\tau)s(t-\tau)\,\mathrm{d}\tau$, where the linear filter has the form $g(t) = \frac{\bar{g}}{3!}(t/\tau_p)^3\mathrm{e}^{-t/\tau_p}$ for $t > 0$, and $g(t) = 0$ otherwise. The resulting receptor activity $r(t)$ is transmitted to bipolar and horizontal cells. However, the sensitivity of the transmission is controlled by a divisive term involving bipolar cell activity: $r(t) = p(t)/\exp(b(t)/B_{trig})$ and $b(t) = \int h(\tau)r(t-\tau)\,\mathrm{d}\tau$, where $b(t)$ is the bipolar activity and B_{trig} is the threshold value of bipolar cell activity necessary to induce the gain control. The bipolar activity is, in turn, a linear function of the receptor output. The linear filter $h(t)$ has the form: $h(t) = \frac{1}{\tau_h}\mathrm{e}^{-t/\tau_h}$ for $t > 0$, and $h(t) = 0$ otherwise. Simulations were conducted with $\tau_h = 0.2$ s, $\tau_p = 0.01$ s, $\bar{g} = 3\ \mu\mathrm{V}/(\text{quanta s}^{-1})$ and $B_{trig} = 30$ mV. The ratio of ganglion cell firing rate to potential was assumed to be $(30\ \text{impulse s}^{-1})/(1\ \text{mV})$. The firing rate output of the model was used to drive a Poisson spike generator for our simulations.

As with background luminance, the early visual pathway also exhibits adaptation in response to changes in contrast. Many of the phenomena observed in light adaptation are also found in *contrast gain control*. When the contrast of the stimulus is increased, there is a

decrease in response magnitude accompanied by an increase in response speed [29, 30]. These changes have been shown to take place in a few tens of milliseconds. Early work by Shapley and Victor [31] explored this mechanism empirically in the retinal ganglion cells of the cat, and the subsequent modelling work of Victor resulted in the contrast gain control model [37]. In the contrast gain control model, the stimulus first passes through a low pass filter g_L, of the following form: $x(t) = \int g_L(\tau)s(t-\tau)\,\mathrm{d}\tau = \int(\tau/T_L)\mathrm{e}^{-\tau/T_L}s(t-\tau)\,\mathrm{d}\tau/(N_L-1)!$. The adaptation of the high pass stage of the model can be captured in the following three equations: $T_s(t)\dot{y}(t) = -y(t) + T_s(t)\dot{x}(t) + (1-H_s)x(t)$, $T_s(t) = T_0/(1+c(t)/c_{50})$ and $T_c\dot{c}(t) = |y(t)| - c(t)$. Finally, the output rate is a shifted and rectified version of the signal y: $\lambda(t) = \lfloor y(t) + m \rfloor$. Simulations were conducted with $H_s = 0.806$, $N_L = 16$, $M_0 = 31$ impulses s^{-1}, $T_L = 1.94$ ms, $T_0 = 0.193$ s, $c_{50} = 0.054$, $D = 3$ ms, $A_0 = 440$ impulses s^{-1} and $T_C = 15$ ms. For simulations, the firing rate was used to drive a Poisson spike generator.

References

[1] Albrecht D G, Farrar S B and Hamilton D B 1984 Spatial contrast adaptation characteristics of neurones recorded in the cat's visual cortex *J. Physiol.* **347** 713–39

[2] Atick J J 1992 Could information theory provide an ecological theory of sensory processing? *Network: Comput. Neural Syst.* **3** 213–51

[3] Atick J J and Redlich A N 1992 What does the retina know about natural scenes? *Neural Comput.* **4** 196–210

[4] Attneave F 1954 Some informational aspects of visual perception *Psychol. Rev.* **61** 183–93

[5] Barlow H B 1961 Possible principles underlying the transformations of sensory messages *Sensory Communication* ed W A Rosenblith (Cambridge, MA: MIT Press) pp 217–34

[6] Barlow H B, FitzHugh R and Kuffler S W 1957 Changes of organization of the receptive fields of the cat's retina during dark adaptation *J. Physiol.* **137** 338–54

[7] Barlow H B and Levick W R 1969 Three factors limiting the reliable detection of light by retinal ganglion cells of the cat *J. Physiol.* **200** 1–24

[8] Baylor D A and Hodgkin A L 1974 Changes in time-scale and sensitivity in turtle photoreceptors *J. Physiol.* **242** 729–58

[9] Brenner N, Bialek W and de Ruyter van Steveninck R 2000 Adaptive rescaling maximizes information transmission *Neuron* **26** 695–702

[10] Brown E N, Nguyen D P, Frank L M, Wilson M A and Solo V 2001 An analysis of neural receptive field plasticity by point process adaptive filtering *Proc. Natl Acad. Sci.* **98** 12261–6

[11] Chander D and Chichilnisky E J 2001 Adaptation to temporal contrast in primate and salamander retina *J. Neurosci.* **21** 9904–16

[12] Chichilnisky E J 2001 A simple white noise analysis of neuronal light responses *Network: Comput. Neural Syst.* **12** 199–213

[13] Donner K, Koskelainen A, Djupsund K and Hemila S 1995 Changes in retinal time scale under background light: observations on rods and ganglion cells in the frog retina *Vis. Res.* **35** 2255–66

[14] Donner K and Reuter T 1965 The dark-adaptation of single units in the frog's retina and its relation to the regeneration of rhodopsin *Vis. Res.* **5** 615–32

[15] Enroth-Cugell C and Shapley R M 1973 Adaptation and dynamics of cat retinal ganglion cells *J. Physiol.* **233** 271–309

[16] Fairhall A L, Lewen G D, Bialek W and de Ruyter van Stevenick R R 2001 Efficiency and ambiguity in an adaptive neural code *Nature* **412** 787–90

[17] Gallant J L, Connor C E and Van Essen D C 1998 Neural activity in areas v1, v2 and v4 during free viewing of natural scenes compared to controlled viewing *Neuroreport* **9** 1673–8

[18] Kim K J and Rieke F 2001 Temporal contrast adaptation in the input and output signals of salamander retinal ganglion cells *J. Neurosci.* **21** 287–99

[19] Marmarelis P Z and Marmarelis V Z 1978 *Analysis of Physiological Systems* (New York: Plenum)

[20] Meister M and Berry M J 1999 The neural code of the retina *Neuron* **22** 435–50

[21] Nakatini K, Tamura T and Yau K W 1991 Light adaptation in retinal rods of the rabbit and two other primate animals *J. Gen. Physiol.* **97** 413–35

[22] Ohzawa I, Sclar G and Freeman R D 1985 Contrast gain control in the cat's visual system *J. Neurophysiol.* **54** 651–67
[23] Przybyszewski A W, Gaska J P, Foote W and Pollen D A 2000 Striate cortex increases contrast gain of macaque lgn neurons *Vis. Neurosci.* **17** 485–94
[24] Reinagel P and Zador A M 1999 Natural scene statistics at the centre of gaze *Network: Comput. Neural Syst.* **10** 341–50
[25] Ringach D L, Hawken M J and Shapley R 2002 Receptive field structure in neurons in monkey primary visual cortex revealed by stimulation with natural image sequences *J. Vis.* **2** 12–24
[26] Sanchez-Vives M V, Nowak L G and McCormick D A 2000 Membrane mechanisms underlying contrast adaptation in cat area 17 in vivo *J. Neurosci.* **20** 4267–85
[27] Sclar G 1987 Expression of retinal contrast gain control by neurons of the cat's lateral geniculate nucleus *Exp. Brain Res.* **66** 589–96
[28] Sclar G, Lennie P and DePriest D D 1989 Contrast adaptation in striate cortex of macaque *Vis. Res.* **29** 747–55
[29] Shapley R M and Victor J D 1979 The contrast gain control of the cat retina *Vis. Res.* **19** 431–4
[30] Shapley R M and Victor J D 1981 How the contrast gain control modifies the frequency responses of cat retinal ganglion cells *J. Physiol.* **318** 161–79
[31] Shapley R M and Victor J D 1978 The effect of contrast on the transfer properties of cat retinal ganglion cells *J. Physiol.* **285** 275–98
[32] Shou T, Li X, Zhou Y and Hu B 1996 Adaptation of visually evoked responses of relay cells in the dorsal lateral geniculate nucleus of the cat following prolonged exposure to drifting gratings *Vis. Neurosci.* **13** 605–13
[33] Simoncelli E P and Olshausen B A 2001 Natural image statistics and neural representation *Annu. Rev. Neurosci.* **24** 1193–216
[34] Smirnakis S M, Berry M J, Warland D K, Bialek W and Meister M 1997 Adaptation of retinal processing to image contrast and spatial scale *Nature* **386** 69–73
[35] Stanley G B 2002 Adaptive spatiotemporal receptive field estimation in the visual pathway *Neural Comput.* **14** 2925–46
[36] Stanley G B, Li F F and Dan Y 1999 Reconstruction of natural scenes from ensemble responses in the lateral geniculate nucleus *J. Neurosci.* **19** 8036–42
[37] Victor J D 1987 The dynamics of the cat retinal x cell centre *J. Physiol.* **386** 219–46
[38] Vinje W E and Gallant J L 2000 Spare coding and decorrelation in primary visual cortex during natural vision *Science* **287** 1273–6
[39] Yeh T, Lee B B and Kremers J 1996 The time course of adaptation in macaque retinal ganglion cells *Vis. Res.* **36** 913–31
[40] Gallant J L 2002 Private communication