A model of the dynamics of retinal activity during natural visual fixation

GAËLLE DESBORDES[1] AND MICHELE RUCCI[2]

[1]Division of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts
[2]Department of Cognitive and Neural Systems, Boston University, Boston, Massachusetts

Abstract

During visual fixation, small eye movements keep the retinal image continuously in motion. It is known that neurons in the visual system are sensitive to the spatiotemporal modulations of luminance resulting from this motion. In this study, we examined the influence of fixational eye movements on the statistics of neural activity in the macaque's retina during the brief intersaccadic periods of natural visual fixation. The responses of parvocellular (P) and magnocellular (M) ganglion cells in different regions of the visual field were modeled while their receptive fields scanned natural images following recorded traces of eye movements. Immediately after the onset of fixation, wide ensembles of coactive ganglion cells extended over several degrees of visual angle, both in the central and peripheral regions of the visual field. Following this initial pattern of activity, the covariance between the responses of pairs of P and M cells and the correlation between the responses of pairs of M cells dropped drastically during the course of fixation. Cell responses were completely uncorrelated by the end of a typical 300-ms fixation. This dynamic spatial decorrelation of retinal activity is a robust phenomenon independent of the specifics of the model. We show that it originates from the interaction of three factors: the statistics of natural scenes, the small amplitudes of fixational eye movements, and the temporal sensitivities of ganglion cells. These results support the hypothesis that fixational eye movements, by shaping the statistics of retinal activity, are an integral component of early visual representations.

Keywords: Retinal ganglion cells, Fixational eye movements, Natural images, Pairwise correlation, Redundancy reduction, Microsaccade

1. Introduction

The structure of correlated activity appears to play a critical role in the neural encoding of visual information. Correlated cell responses, both in the form of coactivity of instantaneous firing rates (Roelfsema et al., 2004) and as synchronous spikes (Singer & Gray, 1995), might signal the presence of important features in the visual scene, such as an edge or an object. Furthermore, in the early stages of the visual system, simultaneously active neurons are more likely than isolated cell responses to influence neural activity at later stages (Alonso et al., 1996; Usrey & Reid, 1999).

In the presence of natural visual stimulation, however, a scheme of neuronal encoding based on correlated activity faces a serious challenge. On average, the luminance of natural images tend to vary smoothly over space (Field, 1987). Therefore, in order to exploit input correlations caused by relevant visual features, the visual system needs to distinguish them from the "uninteresting" correlations generally present in natural scenes. Elimination of these predictable correlations has also been proposed as a means of

coping with the information bottleneck created by the optic nerve (Atick, 1992). In principle, center-surround spatial filters, qualitatively similar to the receptive fields of retinal ganglion cells (RGCs), could eliminate the average correlation of natural scenes (Srinivasan et al., 1982; Atick & Redlich, 1992; van Hateren, 1992). However, a closer look at these theoretically-derived filters reveals that they do not agree quantitatively with neurophysiological data (see Discussion).

Under natural viewing conditions, the input signal to the retina depends not only on the visual scene, but also on the oculomotor activity performed by the observer during the acquisition of visual information. Eye movements are always present, even during visual fixation when microscopic saccades alternate with periods of drift, and contribute to shaping the spatiotemporal input to the retina (Ratliff & Riggs, 1950; Ditchburn & Ginsborg, 1953; Steinman et al., 1973). Neurophysiological recordings have shown that fixational saccades modulate the responses of neurons in the lateral geniculate nucleus (LGN) and primary visual cortex (V1) (Gur et al., 1997; Leopold & Logothetis, 1998; Martinez-Conde et al., 2000; Snodderly et al., 2001). Furthermore, previous modeling work has suggested that fixational eye movements attenuate the impact of the broad spatial correlations of natural images on neural activity (Rucci et al., 2000).

Address correspondence and reprint requests to: Michele Rucci, Boston University, Department of Cognitive and Neural Systems, 677 Beacon Street, Boston, MA 02215. E-mail: rucci@cns.bu.edu

The goal of this study is to examine the impact of eye movements that occur under natural viewing conditions on the representation of visual information in the retina. During natural viewing, saccades continually alternate with brief periods of fixational eye movements. We use neural modeling to analyze the way the natural alternation between macroscopic and microscopic eye movements influences the second-order statistics of neural activity in populations of RGCs. This study differs from our previous modeling work in two significant ways. First, whereas in previous studies, we modeled neural responses in the cat's LGN during artificially-generated eye movements (Rucci & Casile, 2004, 2005; Casile & Rucci, 2006), here we analyze neural activity in the macaque's retina when natural images are scanned by real, previously recorded patterns of eye movements. This difference is important, as the structure of correlated activity depends on the properties of cell responses, and the characteristics of eye movements. Since macaques are similar to humans in visual performance and eye movements, results from this analysis are likely to be relevant for human vision, and can be examined in light of the vast literature in human psychophysics. A second important difference with respect to our previous work concerns the periods of interest in the analysis of correlated activity. In previous models, in order to isolate the contributions of fixational eye movements, we focused on steady-state levels of neural activity, which were measured either during sustained fixation or during late periods of fixation. In this study, instead, we examine the full temporal evolution of correlated retinal activity during the time-course of natural visual fixation. We show that eye movements profoundly influence the statistics of neural activity in models of ganglion cells. The alternation of macroscopic and microscopic eye movements that continually occurs during natural viewing gives rise to a well-defined dynamics of retinal activity: wide pools of coactive ganglion cells emerge after saccades and shrink during the course of fixation in a way that compensates for the broad correlations of luminance present in natural scenes. We show that this stereotypical dynamics of neural activity occurs only during viewing of scenes with a power spectrum similar to that of natural images. These results support the argument that fixational eye movements are a critical component of the acquisition and neural encoding of visual information.

2. Materials and methods

2.1. Visual stimulation

To replicate the spatiotemporal input to the retina that occurs under natural viewing conditions, the responses of macaque RGCs were simulated as their receptive fields followed the trajectories of recorded eye movements. In the main simulations of this study, visual input was provided by 20 gray-scale photographs of natural scenes chosen from a public domain database (van Hateren & van der Schaaf, 1998). The average power spectrum of these images was best fit by $1/f^{2.02}$ with spatial frequency f, which is consistent with other measurements of the power spectrum of natural images (Field, 1987; Ruderman, 1994). In these images, one pixel corresponded to $1'$ of visual angle, and each image spanned a visual field of approximately 25×17 square degrees. In control simulations, we used three sets of 20 synthetic images consisting of noise with predefined power spectra varying as $1/f^2$, $1/f^3$, and $1/f^4$. These images were generated by filtering white noise patterns with appropriate spatial kernels.

In each trial of our simulations, the spatiotemporal input to modeled retinal neurons was a movie, $I(\mathbf{x}, t)$, produced by scanning an image following a prerecorded trajectory of eye movements. Eye movements were recorded from four human observers with normal vision during free viewing of the same 20 natural images of our simulations. Informed consent was obtained from all subjects following the procedures approved by the Boston University Charles River Campus Institutional Review Board. The use of human eye movements to simulate neural responses in the macaque retina is justified by the similarity between the oculomotor behaviors of humans and macaques (Snodderly & Kurtz, 1985; Snodderly, 1987). Eye movements were measured by means of a Generation 6 Dual-Purkinje-Image eyetracker (Fourward Technologies Inc., Buena Vista, VA). This device has a nominal delay of less than 1 ms and a nominal spatial resolution of approximately $1'$ (Crane & Steele, 1985; Stevenson & Roorda, 2005), therefore enabling the recording of fine fixational eye movements. A dental imprint bite bar and a forehead rest prevented movements of the head. Vertical and horizontal eye position signals were sampled at 1 kHz.

In control simulations, to study the influence of the amplitude of eye movements, we used synthetic fixational eye movements. In these simulations, each fixation started from a randomly selected location on the image and lasted for 300 ms. Two separate random walks modeled horizontal and vertical eye position. For each random walk, a new step was drawn every millisecond from a normal distribution with zero mean and standard deviation σ. The spatial extent of the region covered by these synthetic fixations varied with σ. We report results for three values of σ corresponding to average angular spans of $6'$, $1°$, and $2°$, where the angular span of a fixation is defined as the maximum distance between the scanpath's centroid and each of its points.

2.2. Modeling neuronal responses

The responses of RGCs were simulated using standard firing-rate models composed of the cascade of linear and nonlinear stages (Carandini et al., 2005). In the context of this paper, since the instantaneous firing rate already provides statistical averaging, it is a more appropriate signal to simulate than the actual train of spikes. Following convention, we term magnocellular-projecting parasol ganglion cells, "M cells," and parvocellular-projecting midget ganglion cells, "P cells." We modeled populations of ON-center M and P cells in two regions of the visual field: central ($0–5°$ eccentricity) and peripheral ($20–30°$ eccentricity). Parameter values for each of the four types of cells ("Central M," "Central P," "Peripheral M," "Peripheral P") were taken from corresponding neurophysiological measurements of macaque RGCs as explained below. For each population, we simulated 305 cells arranged on a pattern composed of 19 concentric circles with 16 cells on each circle, and one cell at the center of the population. The radii of the circles varied from $5'$ to $95'$ in steps of $5'$.

The response of a model cell to visual stimulation simulated the deviation in the neuron's instantaneous firing rate with respect to the level of spontaneous activity. The visual response at time t of a cell at location $\mathbf{z} = (x, y)$ in the retina, $\alpha_{\mathbf{z}}(t)$, was computed as the rectified temporal convolution between the spatiotemporal stimulus $I(\mathbf{x}, t)$—the movie obtained by sampling the image following eye movements—and the cell spatiotemporal kernel $h_\alpha(c, \mathbf{x}, t)$, such that

$$\alpha_{\mathbf{z}}(t) = \left[\int_{S_z} \int_{-\infty}^{t} h_\alpha(c, \mathbf{x}', t') I(\mathbf{x}', t - t') \, \mathbf{dx}' \, dt' \right]_\beta, \qquad (1)$$

where c parametrizes the local level of contrast, S_z is the area of the retina spanned by the cell's receptive field, and the operator $[\cdot]_\beta$ indicates rectification with threshold β, *i.e.,* a threshold-linear static nonlinearity such that $[z]_\beta = z - \beta$ if $z > \beta$, and $[z]_\beta = 0$ if $z \leq \beta$. Time was discretized in 1-ms steps.

RGC kernels measured physiologically are not space-time separable, because the response elicited by stimulation of the surround is delayed relative to the response elicited by stimulation of the center (Enroth-Cugell et al., 1983; Benardete & Kaplan, 1999*b*). In order to account for this center-surround delay, the spatiotemporal kernel $h_\alpha(c, \mathbf{x}, t)$ was modeled by the sum of two space-time separable terms, which represented the contributions from the center and surround of the receptive field, as in Cai et al. (1997):

$$h_\alpha(c, \mathbf{x}, t) = h_{\mathrm{cnt}}(c, \mathbf{x}, t) + h_{\mathrm{srn}}(c, \mathbf{x}, t)$$
$$= g_{\mathrm{cnt}}(c, t) f_{\mathrm{cnt}}(\mathbf{x}) + g_{\mathrm{srn}}(c, t) f_{\mathrm{srn}}(\mathbf{x}), \qquad (2)$$

where g_{cnt} and f_{cnt} are the temporal and spatial components of the center, and g_{srn} and f_{srn} are those of the surround. The spatial kernels $f_{\mathrm{cnt}}(\mathbf{x})$ and $f_{\mathrm{srn}}(\mathbf{x})$ were modeled as two-dimensional Gaussian functions, with amplitudes A_{cnt} and A_{srn} and standard deviations σ_{cnt} and σ_{srn}. Spatial parameters were set to the median values measured by Croner and Kaplan (1995) in macaque RGCs with receptive fields in the central and peripheral regions of the visual field.

Following the model introduced by Victor (1987) for cat RGCs, the temporal profile of the receptive field center, $g_{\mathrm{cnt}}(c, t)$, was computed as the inverse Fourier transform of the temporal-frequency response $K(c, \omega)$ to a sine-wave grating with contrast c modulated at temporal frequency ω. The temporal-frequency response $K(c, \omega)$ was modeled by the series of a set of low-pass filters, and a high-pass stage in the form of

$$K(c, \omega) = A e^{-i\omega D} \left(1 - \frac{H_S}{1 + i\omega\tau_S} \right) \left(\frac{1}{1 + i\omega\tau_L} \right)^{N_L}, \qquad (3)$$

where A represents the overall gain; H_S, the strength of the subtractive stage; τ_S and τ_L, the time constants of the high- and low-pass stages; N_L, the number of low-pass stages; and D, the brief transmission delay along the RGC axon, which was included to model the output of the retina. In this model, the cascade of filters is a mathematical commodity with the purpose of data fitting; these filters are not meant to describe individual biological components. The filter parameters define the range and shape of the contrast sensitivity function. In particular, τ_S regulates the response transience by controlling sensitivity to low temporal frequencies. This parameter varied with contrast, as explained below. This model has been successfully applied to fit data from macaque RGCs (Benardete et al., 1992; Benardete & Kaplan, 1997, 1999*a*, 1999*b*). As with spatial parameters, temporal parameters of the model were set to the median values measured in macaque RGCs (P cells: Benardete & Kaplan, 1997; M cells: Benardete & Kaplan, 1999*a*). The temporal profile of the surround, $g_{\mathrm{srn}}(c, t)$, was similar to that of the center, except for a center-surround delay τ_{CS}, such that

$$g_{\mathrm{srn}}(c, t) = g_{\mathrm{cnt}}(c, t - \tau_{\mathrm{CS}}).$$

In the main simulations of this study, the value of τ_{CS} was set to 3 ms. Control simulations examined the effect of varying τ_{CS} between 0 and 10 ms, which is the range of center-surround delays

measured in RGCs, and cells in the lateral geniculate nucleus of the macaque (Benardete & Kaplan, 1997, 1999*b*; Reid & Shapley, 2002).

In the case of P cells, the stimulus contrast c is known to have little influence on receptive-field properties (Lee, 1996; Benardete & Kaplan, 1997, 1999*b*). M cells however, exhibit a mechanism of contrast gain control by which the temporal profile changes with stimulus contrast (Shapley & Victor, 1979; Benardete et al., 1992; Lee et al., 1994; Benardete & Kaplan, 1999*a*). In the model of M cells, the value of τ_S in Eq. (3) varied as a function of contrast c such that

$$\tau_S(c) = \frac{T_0}{1 + \left(\dfrac{c}{C_{1/2}} \right)^2}, \qquad (4)$$

where T_0 is the time constant at zero contrast, and $C_{1/2}$ represents the contrast at which τ_S is half its initial value (Victor, 1987; Benardete & Kaplan, 1999*a*). In this study, the local contrast of the input image was on average 12%, which is consistent with measurements of the contrast distribution of natural scenes (Ruderman, 1994; Brady & Field, 2000).

Since activity in the model was defined as the deviation from the spontaneous firing rate, positive (negative) values in the output of simulated cells correspond to firing rates higher (lower) than spontaneous activity. Rectification was present in the model to account for the asymmetry between the possible ranges of negative and positive responses: real firing rates can increase largely above the spontaneous firing rate, but they cannot go below zero. In the simulations, the rectification level was defined as the percentage of the range of negative activity in the linear (unrectified) response that was eliminated. Unless otherwise specified, we used 50% rectification, that is, we eliminated half of the range of negative activity present in the unrectified model.

2.3. Statistical analysis of simulated neural activity

We examined the temporal evolution of patterns of coactive units during visual fixation. During the course of each fixation, the correlation between the responses of two cells with receptive fields centered at $\mathbf{x}_i$ and $\mathbf{x}_j$, $\alpha_{\mathbf{x}_i}(t)$ and $\alpha_{\mathbf{x}_j}(t)$, was evaluated at several times. At each time t_d after fixation onset, levels of correlation were computed over an interval of duration T centered at t_d, as in

$$r_{\mathbf{x}_i \mathbf{x}_j}(t_d) = \langle \alpha_{\mathbf{x}_i} \alpha_{\mathbf{x}_j} \rangle_{[t_d - (T/2),\, t_d + (T/2)]}$$
$$= \frac{1}{T} \int_{t_d - (T/2)}^{t_d + (T/2)} \alpha_{\mathbf{x}_i}(\tau) \alpha_{\mathbf{x}_j}(\tau)\, d\tau. \qquad (5)$$

Levels of correlation were averaged over all pairs of simulated cells with receptive fields at the same separation d, and averaged over a large set of fixations $\mathcal{F}$ on the database of natural images, giving the spatial correlation function

$$r(d, t) = \langle r_{\mathbf{x}_i \mathbf{x}_j}(t) \rangle_{\|\mathbf{x}_j - \mathbf{x}_i\| = d,\, \mathcal{F}}.$$

In the main simulations, levels of correlation were averaged over a total of $N = 522$ fixations, each with its unique pattern of eye movements. In control experiments with synthetic fixational eye movements, we simulated 30 fixations on each image, for a total of $N = 600$ fixations. In some simulations we also computed the

spatial covariance function during fixation, $c(d, t)$. The procedure followed to evaluate $c(d, t)$ was similar to the one described above for $r(d, t)$, except that in each fixation, before multiplying the responses of two cells, we first subtracted their mean activities over the fixation period, $\overline{\alpha_{\mathbf{x}_i}}$ and $\overline{\alpha_{\mathbf{x}_j}}$. That is,
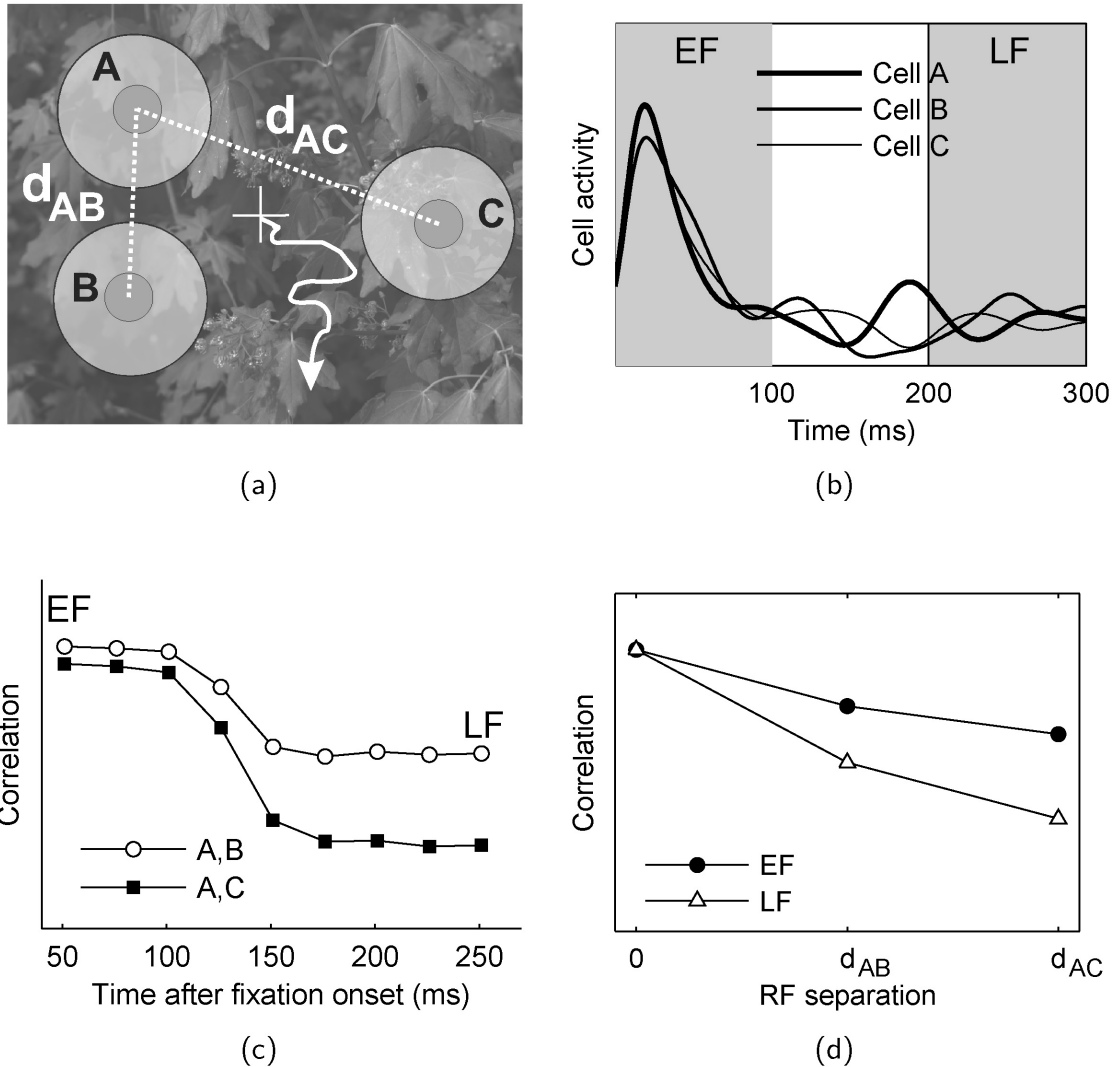
$$c_{\mathbf{x}_i \mathbf{x}_j}(t_d) = \langle (\alpha_{\mathbf{x}_i}(t) - \overline{\alpha_{\mathbf{x}_i}})(\alpha_{\mathbf{x}_j}(t) - \overline{\alpha_{\mathbf{x}_j}}) \rangle_{[t_d - (T/2), t_d + (T/2)]}. \quad (6)$$

In the rest of this article, we use the terms "correlation" and "covariance" to refer to the normalized versions of the spatial correlation and covariance functions, such that the autocorrelation and autocovariance (i.e., the values for a receptive-field separation $d = 0$) have unit amplitude. We focused on two time intervals: the

first 100 ms after fixation onset, which we indicate as Early Fixation, and the 200–300 ms interval after fixation onset, referred to as Late Fixation. The procedure for measuring correlation is illustrated by the example in Fig. 1.

3. Results

In this study, we examined the influence of eye movements on the spatial structure of correlated activity in a model of the macaque's retina. Fig. 1 demonstrates the procedure followed to measure levels of correlation between the simulated responses of pairs of ganglion cells. As illustrated in Fig. 1a, to replicate the spatiotemporal input to the retina which occurs during natural viewing, the
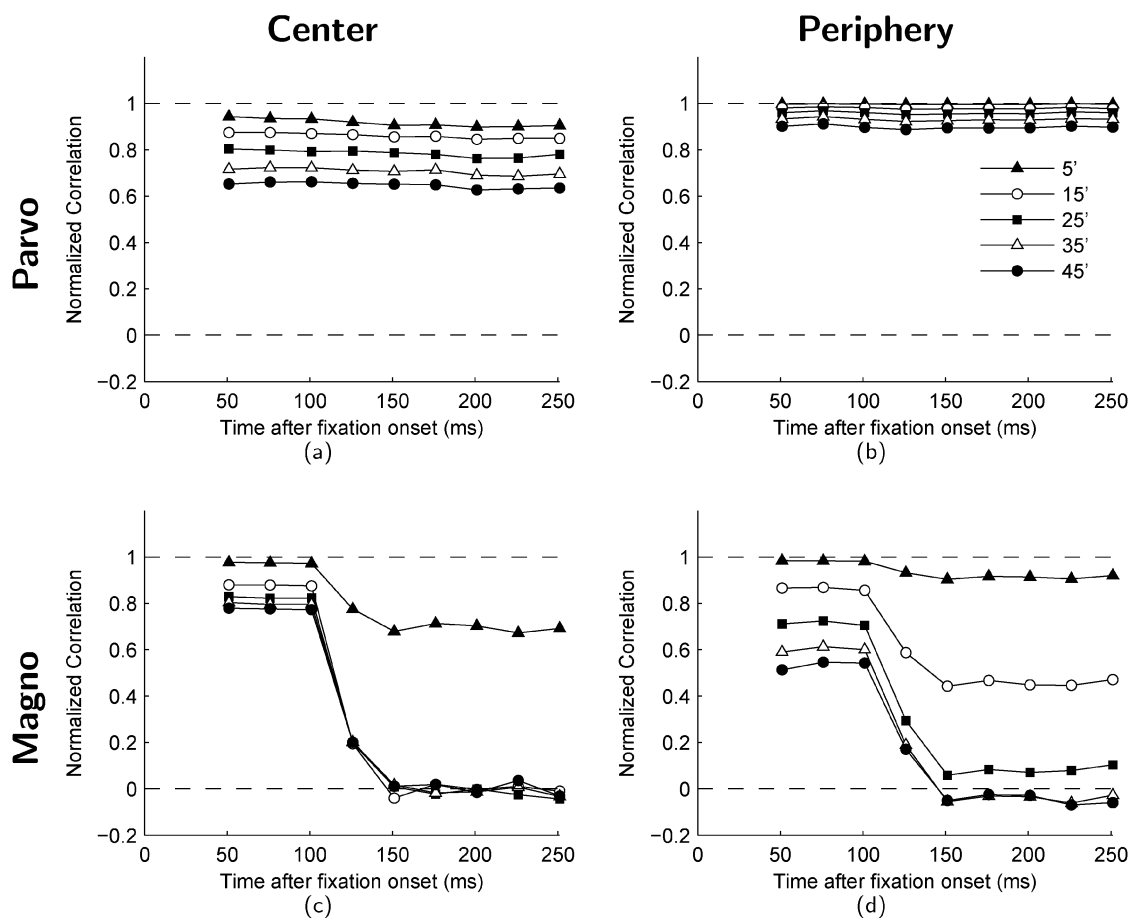

Fig. 1. Procedure for measuring the dynamics of patterns of correlated activity in populations of modeled RGCs. (a) Cell responses were simulated while their receptive fields scanned images of natural scenes following sequences of recorded eye movements. The white arrow represents the motion of the eye during a period of fixation. For clarity, only three cells, A, B, and C, are shown here (not to scale). d_{AB} and d_{AC} represent the distances between the receptive field centers of the two pairs of cells (A, B) and (A, C), respectively. (b) Example of simulated neural responses during a 300-ms period of visual fixation. Two intervals of interest, Early Fixation (EF) and Late Fixation (LF), are highlighted in gray. (c) Correlation as a function of time after fixation onset for the two pairs of cells (A, B) and (A, C). A data point at time t represents pairwise correlation in the $[t - 50, t + 50]$ interval after fixation onset. Correlation levels over the Early and Late Fixation periods are given by the first and last data point, respectively. (d) Correlation as a function of receptive field separation, plotted separately for Early Fixation and Late Fixation.

receptive fields of populations of RGCs scanned natural images following sequences of recorded eye movements. Fig. 1b shows the responses of three cells during a 300-ms period of visual fixation, the typical duration of intersaccadic intervals (Harris et al., 1988; Andrews & Coppola, 1999). After an initial burst of activity at fixation onset, cell responses were strongly modulated by the occurrence of fixational eye movements, as can be seen in the 100–300 ms interval.

As described in the Materials and methods, the degree of coactivity of pairs of RGCs was quantified by the correlation between their responses at zero time delay. Levels of correlation were evaluated at several times during the course of fixation. They were averaged over a large number of fixations on different images and also averaged over all pairs of simulated cells with receptive fields at a given separation *d*. The resulting pattern of correlation was a function of two variables: *t*, the time relative to fixation onset at which correlation was evaluated, and *d*, the distance between receptive field centers. Fig. 1c shows the time course of correlation for the two pairs of cells in Fig. 1a. In both curves, each data point represents the mean level of correlation between the responses of the two cells evaluated over the 100-ms interval, centered at a different time *t* following the onset of visual fixation. In Fig. 1d, the same data as in Fig. 1c are replotted as a function

of the separation between the receptive fields of the two cells. The two curves in this graph represent correlation during two different time intervals: the first 100 ms of fixation (Early Fixation) and the period from 200 to 300 ms after fixation onset (Late Fixation).

The main results of our simulations are reported in Fig. 2. Each panel in Fig. 2 shows the typical evolution of correlated activity over the time course of fixation for one of the four simulated neuronal populations: Central M cells, Central P cells, Peripheral M cells, and Peripheral P cells. The curves in each panel represent mean levels of correlation between pairs of cells with receptive fields at different separations. At the onset of fixation, the sudden change in visual input caused by the saccade gave rise to wide pools of synchronously active cells in all simulated neuronal populations. As explained later in this section, this pattern of activity was a consequence of the broad spatial correlations present in natural images. In populations of P cells, both in the central (Fig. 2a), and in the peripheral region of the visual field (Fig. 2b), correlation changed little during the course of fixation. In contrast, fixational eye movements strongly influenced the correlation between the responses of M cells. Approximately 100 ms after the onset of fixation, levels of correlation dropped in pairs of M cells with receptive fields separated by more than 5′ in the central visual field (Fig. 2c) and 25′ in the periphery (Fig. 2d). That is, during the

Fig. 2. Dynamics of correlated retinal activity during visual fixation on natural images. Each panel shows the temporal evolution of the patterns of coactive units in a given neuronal population: (a) central P cells, (b) peripheral P cells, (c) central M cells, (d) peripheral M cells. The curves in each panel show levels of correlation between pairs of cells with receptive fields separated by different distances in the range 5′–45′. Data points represent pairwise correlation averaged over *N* = 522 fixations on 20 natural images.

normal instability of visual fixation, the spatial structure of correlated activity in populations of M cells quickly became unaffected by the broad correlations of natural scenes.
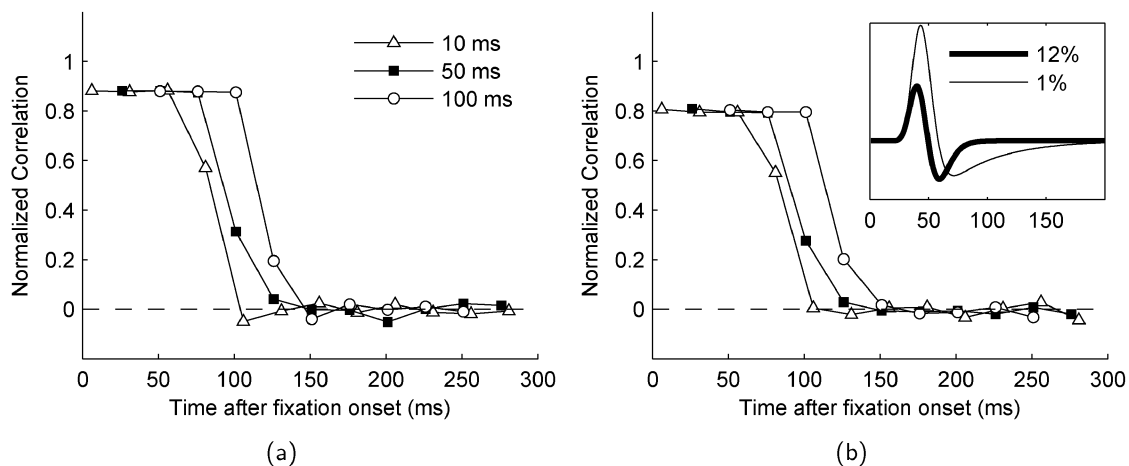
Levels of correlation in Fig. 2 were evaluated over a temporal window of duration $T = 100$ ms. To achieve a finer temporal resolution, in Fig. 3 we used windows of shorter durations. As shown by these data, levels of correlation started to decrease approximately 50 ms after the onset of fixation, and decorrelation was virtually complete after 100 ms. As shown in the insert in Fig. 3b, 50 ms corresponded to the duration of the positive lobe of the temporal impulse response of M cells, while 100 ms were its entire duration. That is, responses became progressively more decorrelated as neurons lost memory of the visual input from the previous fixation.

To determine the spatial extent of correlated activity, Fig. 4 shows the same data as Fig. 2 plotted as a function of the separation between receptive field centers. Only the correlation levels computed during the two time intervals of Early Fixation (the first 100 ms after fixation onset) and Late Fixation (the period from 200 to 300 ms after fixation onset) are shown in this figure. During Early Fixation, correlation extended over several degrees of visual field. That is, even pairs of cells with no overlap in their receptive fields were strongly correlated at the onset of visual fixation. In the case of P cells, fixational eye movements had little impact on levels of correlation. As shown in Figs. 4a, 4b, the spatial structure of correlation during the period of Late Fixation was very similar to that present during Early Fixation both in the center and the periphery of the visual field. The situation differed in the case of M cells, for which levels of correlation changed considerably during the course of visual fixation. For these cells, the spatial extent of correlated activity was substantially narrower during the period of Late Fixation than during Early Fixation. As shown in Fig. 4c, in the central region, the responses of M cells with receptive fields as close as 15′ became virtually uncorrelated in the presence of fixational eye movements. This distance was entirely determined by the diameter of the receptive field center, which was 12′ for central M cells in our model. Indeed, as shown in Fig. 4d, the spatial extent of correlation was slightly broader between peripheral cells, which had larger receptive field centers
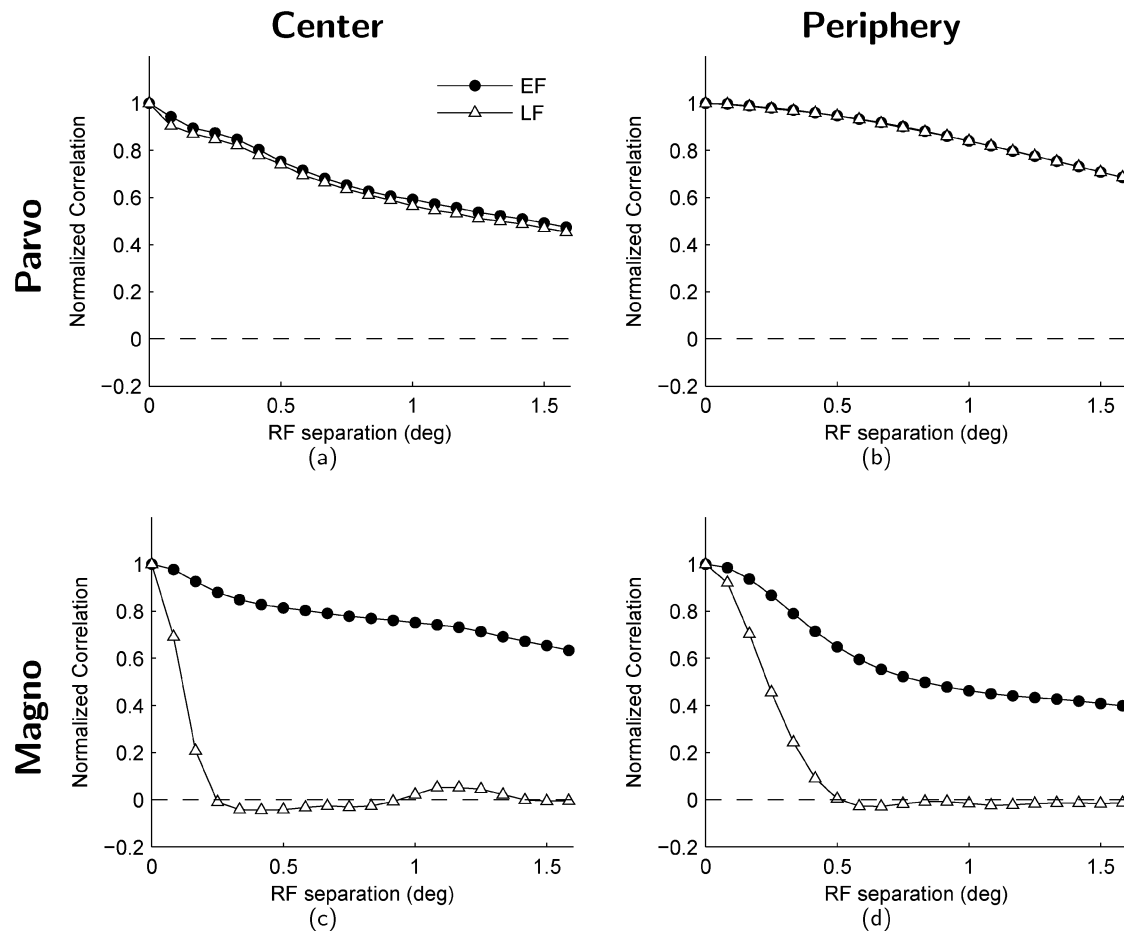
(28′). Hence, when viewing natural images in the presence of fixational eye movements, the responses of pairs of M cells with non-overlapping receptive field centers became uncorrelated during the time course of a typical fixation.

Cortical neurons may be sensitive to other features of their input than the mean levels of activity of pre-synaptic afferents. For instance, they may respond to changes in the incoming firing patterns (Abbott et al., 1997) or to the fine temporal structure of incoming spike trains (Usrey et al., 1998; Greschner et al., 2006). While our model, based on firing rates, cannot address the latter, we can investigate the former by examining the way retinal cell responses covaried around their average firing rates. That is, in addition to analyzing pairwise correlation in the responses of simulated RGCs, we also computed levels of covariance during visual fixation. In this case, the mean responses of modeled cells during a period of fixation were subtracted from the instantaneous firing rates before multiplying them. The results of this analysis are shown in Fig. 5 for cells in the center of the visual field. Both the dynamics and the spatial structure of covarying activity in populations of P and M cells are shown. In populations of M cells, the dynamics of correlation and the dynamics of covariance were almost indistinguishable (compare Fig. 2c to Fig. 5b). As explained later in this section, this result was a consequence of the temporal sensitivity of M cells. A different scenario occurred in populations of simulated P cells. Even though correlation between the responses of P cells remained high during fixation, their covariance decreased in a way which was similar to that of M cells (compare Fig. 2a to Fig. 5a). Therefore, levels of covariance dropped for both P and M cells during fixation on a natural image. In both neuronal populations, the spatial structure of covariance at the end of a 300-ms fixation period was only determined by the overlap between cell receptive fields. Since the same mechanisms were responsible for the dynamics of covariance and the dynamics of correlation in the two cell types, in the remaining of this section we focus on the pairwise correlation between the responses of M cells.

The decorrelation of cell responses shown in Figs. 2 and 5 critically depended on three main factors: the sensitivity of simulated neurons to the spatiotemporal modulations of luminance resulting from fixational eye movements, the small spatial region

(a)   (b)

Fig. 3. Dynamics of correlated activity for different durations of the correlation window (parameter T in Eq. (5)). The two panels show the temporal evolution during the course of visual fixation of pairwise correlation in M cells with receptive fields centers separated by 15′ (a) and 35′ (b). Model parameters replicated receptive fields in the central region of the visual field. The insert in (b) shows the temporal impulse responses of simulated M cells for two different levels of stimulus contrast: 1% and 12%.
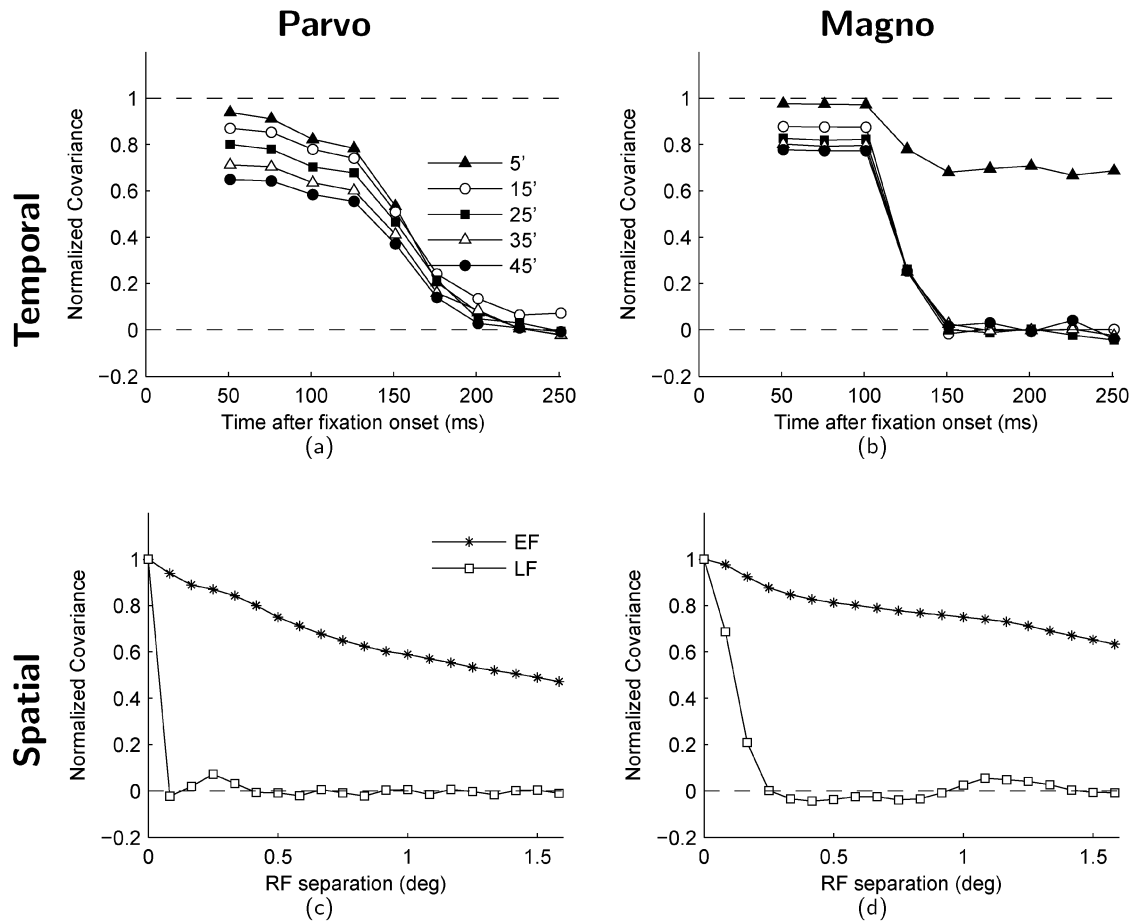
Fig. 4. Spatial structure of correlated retinal activity during visual fixation on natural images. Correlation data are the same as in Fig. 2, but plotted as a function of the distance between receptive field centers. Only two time intervals are shown: Early Fixation (0–100 ms, black symbols) and Late Fixation (200–300 ms, white symbols).

covered by fixational eye movements, and the second-order statistics of natural images. Fig. 6 analyzes the impact of these factors by separately manipulating each of them.

To examine the relative contributions of the spatial and temporal characteristics of cell responses, in control simulations, we exchanged the spatial and temporal receptive fields of simulated P and M cells. In a first simulation, we made the responses of M cells less transient, by substituting their impulse response with that of P cells. The resulting modeled neuron was a hybrid between a P cell and an M cell, which combined the spatial profile of a central M cell with the temporal sensitivity of a P cell. As shown in Fig. 6a, levels of correlation between these artificial units remained high during the course of visual fixation, like in the case of P cells. Fig. 6b shows the results of a second simulation, in which hybrid cells combined the spatial profile of a central P cell with the temporal sensitivity of an M cell. In this case, decorrelation occurred as in normal M cells. These results show that it was the temporal sensitivity of simulated cells, and not their spatial sensitivity, that was responsible for the dynamic decorrelation observed when natural images were viewed in the presence of fixational eye movements. Independently of the spatial characteristics of simulated neurons, levels of correlation dropped during fixational eye movements in pairs of neurons which possessed the temporal profile of M cells, but not in pairs of neurons with the temporal profile of P cells.

To evaluate the influence of the spatial scale of fixational eye movements, we examined the structure of correlated activity obtained with synthetic fixational eye movements of various amplitudes. As explained in the Materials and methods, in these simulations eye movements were modeled as random walks with various step sizes. We used three different step sizes, which, over a period of 300 ms, yielded average angular spans of approximately 6′, 1°, and 2°. For each amplitude, correlations were evaluated over 600 randomly selected fixations on the 20 natural images of our database. As shown in Fig. 6c, decorrelation occurred only for eye movements smaller than 1°. Broad spatial correlations were instead present in the responses of M cells for a fixational span of 2°. These results are consistent with the dynamics of correlated activity observed in our simulations of free-viewing: the wide ensembles of coactive units that emerged after macroscopic saccades shrunk during visual fixation, when small eye movements occurred.

To study the influence of the second-order statistics of natural images, we examined the structure of correlated activity during presentation of images with various power spectra. As explained in the Materials and methods, synthetic images with different spectral densities were created by filtering white-noise images with appropriate spatial kernels. The power spectrum of natural images is known to decline as f^{-2} with spatial frequency (Field, 1987). Images with power spectra that decline faster than f^{-2}, for exam-

Fig. 5. Covariance of retinal activity during visual fixation on natural images. *Top Row:* Dynamics of covarying activity. The temporal evolution of levels of covariance between cells with receptive fields at various distances is plotted as in Fig. 2. *Bottom Row:* Spatial structure of covarying activity. Levels of covariance are plotted as a function of the distance between receptive field centers for the two periods of Early Fixation and Late Fixation as in Fig. 4. Data refer to P and M cells in the central visual field.
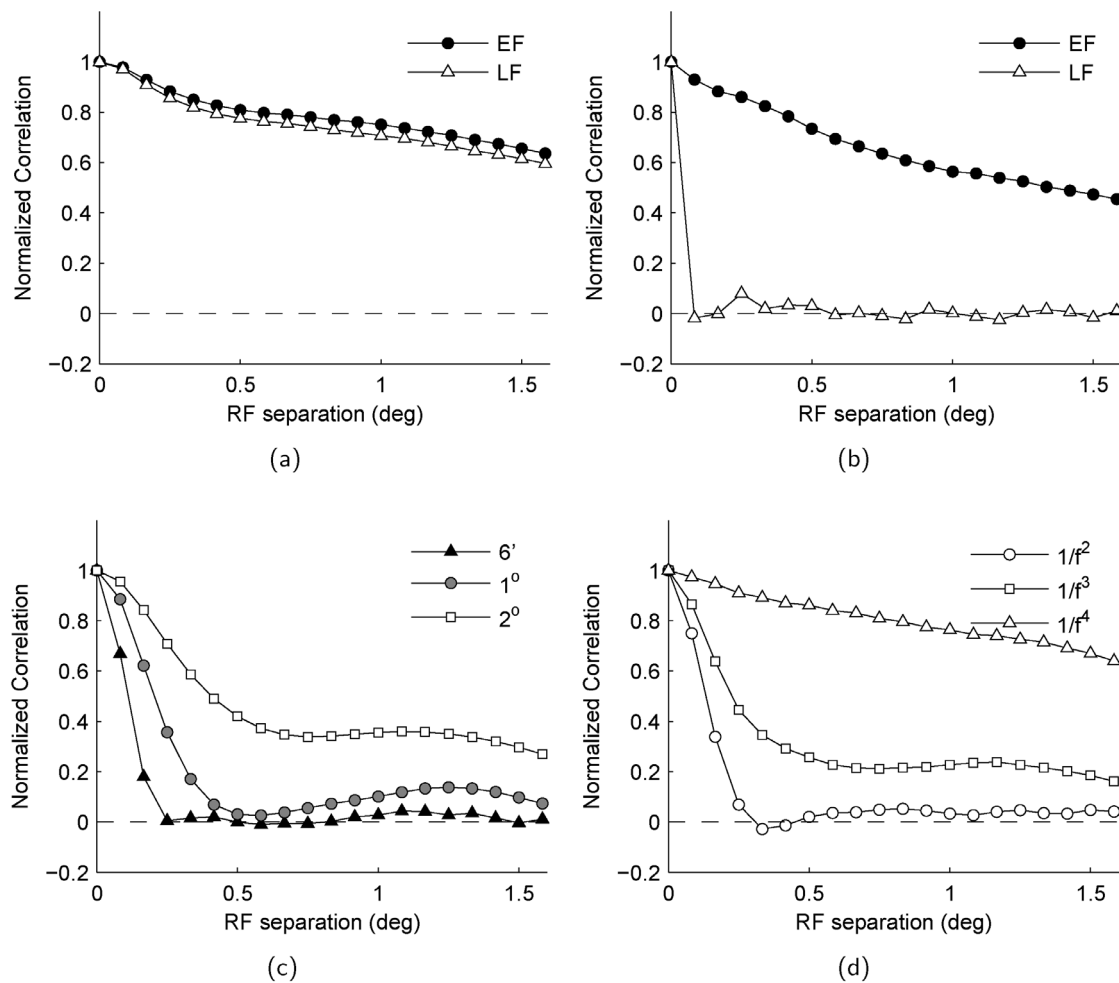
ple, f^{-3} and f^{-4}, have broader spatial correlations than natural images. For each pre-defined power spectrum, levels of correlations were estimated over 20 artificial images. The receptive fields of modeled neurons scanned these images following the same sequences of eye movements used in the simulations of Fig. 2. The data in Fig. 6c show that fixational eye movements gave rise to decorrelated cell responses only in the presence of visual input with a scale-invariant power spectrum, like natural images. Stimulation with an f^{-3} or an f^{-4} spectrum did not result in decorrelated responses by the end of a 300-ms period of fixation. This specificity for images with a f^{-2} spectral density is explained later in this section.

Whereas the factors described in Fig. 6 were critical for the dynamic decorrelation of cell responses, our results were only marginally affected by the precise choice of model parameters. To test the robustness of decorrelation, we systematically varied the values of several parameters in a series of control simulations. These simulations examined the impact of the two sources of nonlinearity present in our model, rectification and contrast gain control, as well as the influence of the degree of space-time inseparability of modeled receptive fields. The results of these simulations are shown in Fig. 7.

Figs. 7a and 7b shows that patterns of correlated activities, as well as their dynamics, were virtually unaffected by the degree of

rectification used in the simulations. Even with 100% rectification, i.e., when we eliminated the full range of negative activity, which would have been present in a linear model, the structure of correlated activity was highly similar to that observed in the absence of rectification. Figs. 7c and 7d shows the results of simulations in which we investigated the effect of varying the average contrast of the image. Whereas P cells are well approximated by linear models for a wide range of contrast levels, M cells can be described by a linear model only for small-amplitude signals at a fixed average contrast. As explained in Materials and methods, our model accounted for contrast gain control by means of the parameter c in Eq. (4). This parameter modified the temporal impulse response of M cells according to the contrast of the stimulus within the receptive field. Figs. 7c and 7d shows the spatial structure of correlated activity in M cells during stimulation with three different levels of average contrast: 1%, 5%, and 12%. The responses of M cells tend to saturate beyond a contrast level of 12% (Lee et al., 1990). As shown by these data, results obtained at different contrasts were almost indistinguishable. Finally, the data in Figs. 7e and 7f, shows the results of control simulations in which we analyzed the influence of the space-time inseparability of cell kernels. Again, varying the center-surround delay τ_{CS} from 0 to 10 ms, the range of values found in macaque RGCs, had negligible influence on the observed patterns of correlation. The

Fig. 6. Main factors contributing to the dynamic decorrelation of retinal activity during fixational eye movements. *Top Row:* Importance of the temporal characteristics of modeled neurons. (a) Spatial structure of correlated activity in a hybrid cell model, which combined the spatial receptive field of a central M cell with the temporal sensitivity of a P cell. Correlation between these cells was similar to correlation between P cells (see Fig. 4a). (b) Spatial structure of correlated activity in a hybrid cell model, which combined the spatial receptive field of a central P cell with the temporal sensitivity of an M cell. Correlation between these cells was similar to correlation between M cells (see Fig. 4c). (c) Importance of the small amplitude of fixational eye movements. Spatial structure of correlated activity between pairs of central M cells, when fixational eye movements of various amplitudes were simulated as random walks. (d) Importance of the input power spectrum. Spatial structure of correlated activity in populations of central M cells, during presentation of images with various power spectra. Data in (c) and (d) refer to the period of Late Fixation.
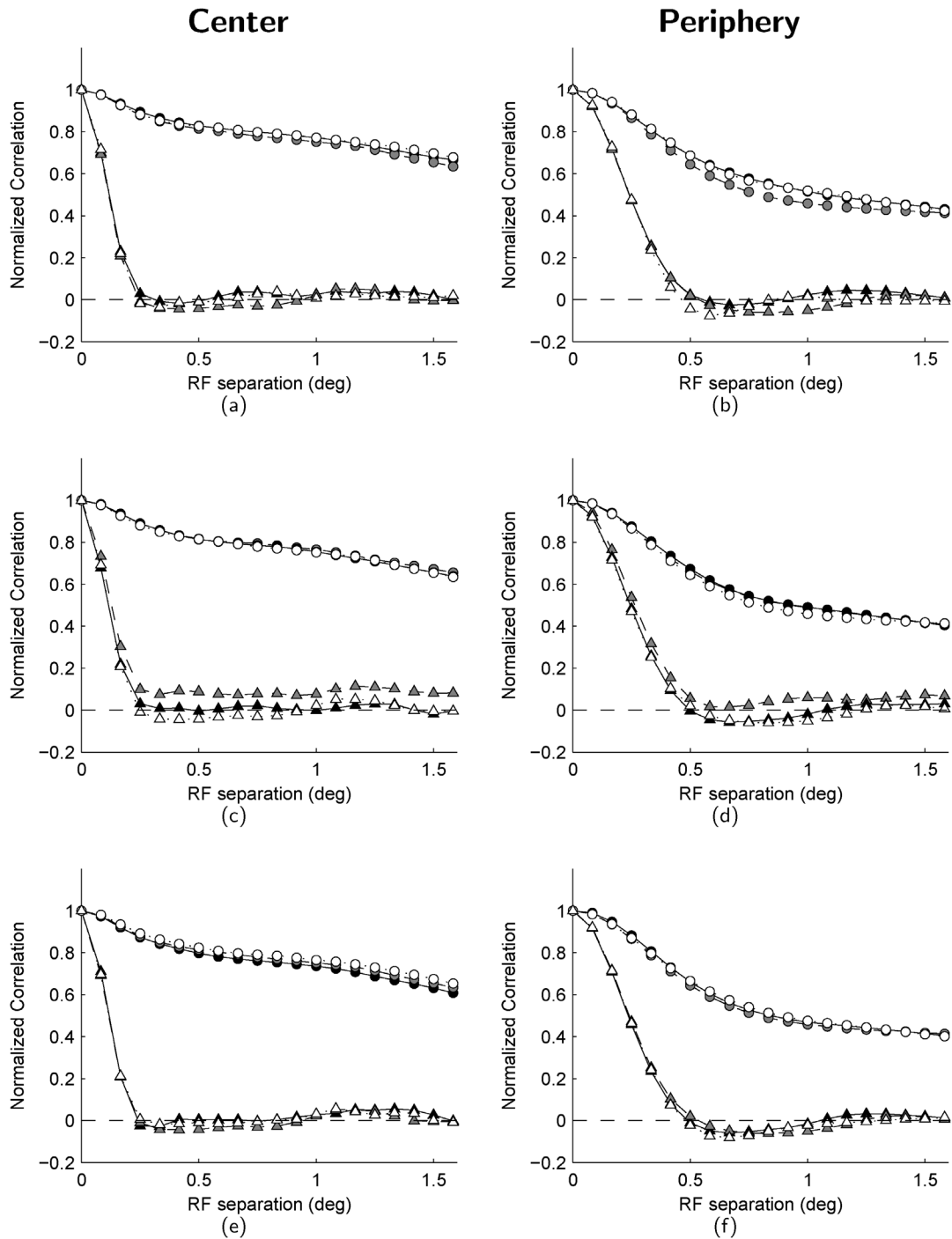
results of these control simulations show that the dynamic decorrelation of the responses of M cells was a robust phenomenon during viewing of natural images. Decorrelation was unaffected by the visual eccentricity of receptive fields, the percentage of rectification, the transmission gain, and the degree of space-time separability of cell kernels. A similar level of robustness was also present in the analysis of patterns of covariance (data not shown).

What are the origins of the progressive decorrelation of simulated neural activity during fixation on a natural image? To provide an intuitive explanation, Fig. 8 examines the responses of two ganglion cells during a single period of fixation, the fixation β shown in the enlarged window in Fig. 8a. Both the cases of a pair of P cells and a pair of M cells are considered. The luminance within the area covered by receptive fields varied as the two cells scanned the image following eye movements (see bottom panel of Fig. 8b). Input signals increased after the saccade, as receptive fields moved from a dark region of the scene to a region with

higher intensity. After the saccade, fixational eye movements modulated the visual input by introducing temporal fluctuations of luminance (ΔI_A and ΔI_B), which reflected the spatial changes of luminance within the area covered by the receptive fields during fixation.

The responses of a pair of P cells are shown in the top panel of Fig. 8b. Both cells fired vigorously following the saccade and continued to respond strongly during fixation. Hence, their responses were highly correlated throughout the fixation period. This high level of correlation was a consequence of the similar input received by the two cells, which is typical of natural scenes. Since the intensity values of natural images tend to be correlated over large distances, pairs of cells with nearby receptive fields often receive similar input, as in the example of Fig. 8.
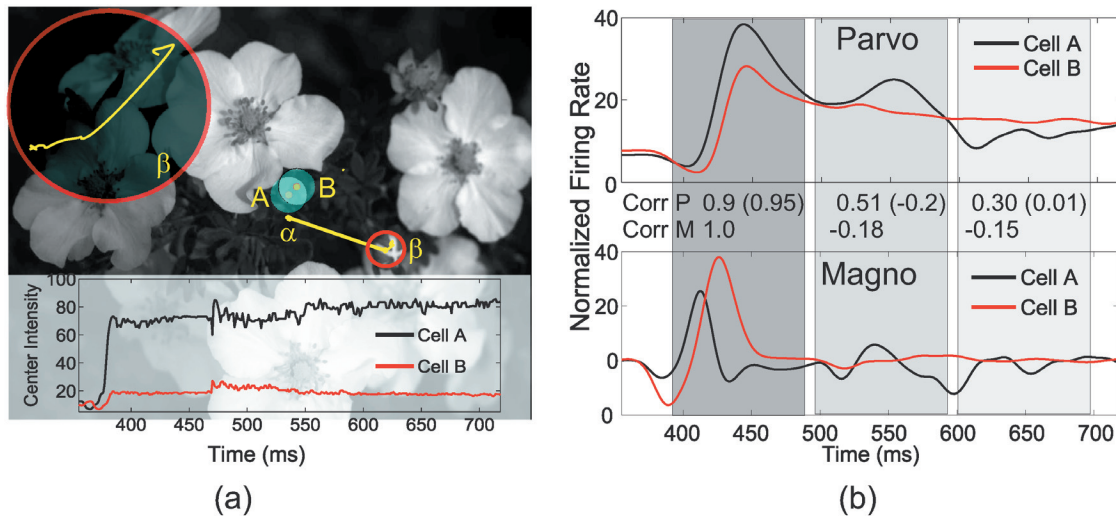
The statistics of natural images were also responsible for the dynamics of correlation between the responses of M cells (bottom panel of Fig. 8b). Unlike P cells, however, M cells are character-

Center   Periphery

Fig. 7. Influence of model parameters on the dynamics of correlated activity in M cells. The left column shows results for central cells (0–5% eccentricity). The right column shows results for peripheral cells (20–30% eccentricity). Circles denote Early Fixation (0–100 ms), triangles represent Late Fixation (200–300 ms). (a, b) Effect of rectification. The rectification level varied from 0% (continuous lines, black symbols) to 50% (dashed lines, gray symbols) to 100% (dotted lines, white symbols). (c, d) Effect of contrast. The contrast level varied from 1% (dashed lines, gray symbols) to 5% (continuous lines, black symbols) to 12% (dotted lines, white symbols). (e, f) Effect of center-surround delay. The center-surround delay varied from 0 ms (continuous lines, black symbols) to 3 ms (dashed lines, gray symbols) to 10 ms (dotted lines, white symbols).

ized by transient temporal responses, i.e., they are highly sensitive to input temporal changes and respond poorly to stationary stimuli. This transient nature contributed to establishing two different regimes of neural sensitivity. Immediately after the saccade, M

cells were strongly influenced by the new stimulus and by the way the visual input changed with respect to the previous fixation. Later during the course of visual fixation, M cells responded primarily to the luminance modulations caused by fixational eye

Fig. 8. Example of dynamics of correlated activity during a single fixation. (a) A natural scene is scanned by a sequence of eye movements composed of two fixations, α and β, separated by a saccade. The fixational eye movements occurring during the period β, shown in the small red circle, are magnified in the window at the top left of the figure. The receptive fields of two cells, A and B, are shown to scale. Center and surround regions within receptive fields are marked by red and blue circles, respectively. The graph on the bottom shows the spatiotemporal input to the two cells as the gaze location moved from α to β. Each curve represents the average pixel intensity within the area covered by the center of a cell receptive field. During β, fixational eye movements modulated the input signal to each cell by introducing fluctuations ($\Delta I_{\text{Cell}}(t)$) around the average luminance experienced by the cell over the fixation period ($\bar{I}_{\text{Cell}}$), *i.e.*, $I_{\text{Cell}}(t) = \bar{I}_{\text{Cell}} + \Delta I_{\text{Cell}}(t)$. (**b**) Simulated responses of a pair of P cells (top panel) and a pair of M cells (bottom panel) during the fixation β. The gray-shaded areas mark three intervals over which levels of correlation were evaluated. Correlation values are given between the two panels. P-cell covariance values are shown in parentheses.

movements. In natural images, these two input regimes differ significantly in their spatial correlations. Whereas luminance values tend to be correlated over large distances, local changes of luminance are spatially uncorrelated. This property of natural images is a direct consequence of their scale-invariant power spectrum and is explained in the Appendix.

It is this change in the component of the visual input effective in driving cell responses which is responsible for the dynamic decorrelation observed in our model. In the example of Fig. 8, the input luminance to both cells increased with the saccade. Thus, during the initial period of fixation, the two cells produced similar, highly correlated responses. Correlation dropped during fixation, as M cells were driven by the spatially uncorrelated fluctuations introduced by fixational eye movements. This transition was less evident in the case of P cells because of their sensitivity to stationary stimuli, which, during fixation, made them continue to respond to the broadly correlated component of their visual input ($\bar{I}_A$ and $\bar{I}_B$). However, elimination of this steady component from cell responses—as in the calculation of levels of covariance instead of correlation—revealed the influence of the spatially uncorrelated fluctuations of luminance introduced by fixational eye movements. Thus, the covariance between the responses of P cells decreased during fixation even if their level of correlation remained high.

4. Discussion

Simultaneously active ganglion cells signal important information about the visual scene. In our model, the alternation between macroscopic and microscopic eye movements that occurs during natural viewing profoundly influenced the structure of correlated activity. Immediately following a large saccade, wide ensembles of coactive ganglion cells extended over several degrees of visual field. These neuronal ensembles shrunk in size as fixational eye movements elicited uncorrelated changes in the firing rates of modeled neurons. This dynamic decorrelation of retinal activity was extremely robust, as it originated from the spatiotemporal stimulus on the retina during free-viewing of natural images. It was little affected by the spatial characteristics of simulated receptive fields, their eccentricity, their degree of space-time inseparability, or the presence of nonlinear mechanisms in the generation of neural responses.

What could be gained from a dynamic spatial decorrelation of retinal activity during natural visual fixation? According to the redundancy reduction hypothesis, a first advantage is a more efficient encoding of visual information (Attneave, 1954; Barlow, 1961). Natural images tend to vary smoothly in space (Field, 1987). Removal of these spatial correlations from the responses of retinal neurons has been proposed as a way of enabling a more efficient transmission of visual information from the retina to the cortex (Srinivasan et al., 1982; Atick & Redlich, 1992; Atick, 1992; van Hateren, 1992). These proposals argue that spatial filters qualitatively similar to the receptive fields of RGCs are capable of eliminating the broad correlations of natural images. However, neurophysiological data are not compatible with these theories, as the sensitivity of RGCs to low spatial frequencies tends to be significantly higher than the frequency response of an ideal decorrelating spatial filter (e.g., Croner & Kaplan, 1995). Indeed, broad correlations between ganglion cells have been measured *in vitro* with presentation of natural images (Puchalla et al., 2005). Consistent with these results, our simulations, which rigorously modeled neurons on the basis of neurophysiological data, produced

broadly correlated cell responses at the onset of fixation, before fixational eye movements could have any effect. Furthermore, correlation remained high throughout fixation in the case of P cells.

The above theories of spatial decorrelation implicitly assume a direct projection of the visual scene onto the retina without considering that, during natural viewing, fixational eye movements continually modulate the visual input to the eye. In contrast to these proposals, the decorrelation of neural activity observed in our simulations originated from the interaction between these input modulations and the temporal sensitivity of RGCs, not from their spatial characteristics. It is a property of natural images that pixel intensities tend to be correlated over large distances whereas changes of intensity are uncorrelated. As explained in the Appendix, this property can be understood by considering that the operation of differentiation corresponds to multiplication by frequency in the Fourier domain. Thus, the first spatial derivative of an image with an f^{-2} spectral density has a flat power spectrum; that is, it is spatially uncorrelated. Fixational eye movements transform local spatial variations into temporal fluctuations in the visual input to the retina. In the model, neurons with transient response characteristics, such as M cells, were highly sensitive to these input modulations, and their responses quickly became uncorrelated during fixation. On the contrary, cells with more sustained responses, such as P cells, remained sensitive to the broad correlations of natural images. Therefore, fixational eye movements had the effect of decreasing the covariance, but not the correlation, between the responses of P cells.

Even though the analysis of the temporal structure of correlated activity is beyond the scope of this paper, it is worth mentioning that in our model fixational eye movements led not only to a spatial decorrelation of cell responses, but also to a reduction in their correlation in time. This result can be explained on the basis of Eq. (8) in the Appendix, which shows that the temporal spectral density of static images scanned by fixational eye movements is approximately equal to the power spectrum of fixational eye movements. This spectrum is known to decline with temporal frequency as ω^{-2} (Eizenman et al., 1985), an input regime that appears to be well counterbalanced by the temporal sensitivity of M cells. A similar interaction has already been proposed as the reason for the temporal decorrelation of geniculate responses observed during stimulation with natural time-varying scenes (Dan et al., 1996).

A second advantage of a dynamic decorrelation of neural responses is the possibility of a "temporal multiplexing" of visual information during the course of fixation. According to this proposal, the normal alternation between large saccades and periods of minute eye movements would enable sequential encoding of image features at different spatial scales. In the model, in the period immediately following the onset of visual fixation, the recent occurrence of a macroscopic saccade enabled the responses of RGCs to integrate input from different regions of the visual scene. Thus, the broad pools of coactive units present during this phase conveyed wide-scale information about the scene. Conversely, responses at a later interval were determined by the spatiotemporal input received from the single region of the scene scanned during the period of fixation. At this time, the uncorrelated modulations of neural activity caused by fixational eye movements signaled spatial changes within a neighborhood of each cell receptive field. That is, by bringing information from different spatial scales within the cell's temporal window of integration, eye movements enabled a hierarchical representation in which the global

structure of the scene was encoded before its fine details. Consistent with this multiplexing proposal, psychophysical evidence exists in support of both (a) a tendency to process global structures before local ones, *i.e.,* the forest before the trees—a phenomenon known as the "global precedence effect" (Navon, 1977); and (b) a coarse-to-fine integration of visual information, in which low spatial frequencies are processed before high frequencies (Parker et al., 1992; Schyns & Oliva, 1994). The results of our simulations suggest a possible contribution of fixational eye movements to these phenomena, which can be tested experimentally. Psychophysical studies of the time course of visual perception have not examined the fine eye movements performed by subjects.

An interesting prediction of a temporal multiplexing of visual information is an involvement of fixational eye movements in fine-scale spatial vision, a hypothesis shared with other recent studies (Ahissar & Arieli, 2001; Greschner et al., 2002). A contribution of fixational eye movements to the perception of spatial detail was originally speculated by Hering (1899) and later refined into the dynamic theories of visual acuity (Averill & Weymouth, 1925; Marshall & Talbot, 1942; Arend, 1973). Interest in these proposals declined after classical experiments on retinal image stabilization—a procedure in which retinal image motion is eliminated—reported little or no change in visual acuity (Riggs et al., 1953; Tulunay-Keesey & Jones, 1976). These pioneering experiments, however, compared results obtained in the absence of retinal image motion to levels of performance measured while subjects maintained steady and accurate fixation. Sustained fixation is an unnatural viewing condition, which is known to reduce the amount of fixational instability (Steinman et al., 1967; Skavenski et al., 1979; Kapoula et al., 1986). Therefore, it is possible that the reduction in retinal image motion caused by sustained fixation contributed to the lack of distinction reported by classical experiments between performances measured under retinal stabilization and in the presence of retinal motion. In contrast, recent psychophysical experiments which avoided sustained fixation found that discrimination of the orientation of a briefly-flashed small bar was impaired under conditions of retinal stabilization (Rucci & Desbordes, 2003). These results support a role for fixational eye movements in the perception of fine spatial detail.

Despite over a century of research, the visual functions of fixational eye movements have remained controversial (see Steinman & Levinson, 1990, for a comprehensive review). It is often claimed that these small eye movements are necessary in order to refresh neural activity to prevent the visual fading experienced under retinal stabilization (e.g., Martinez-Conde et al., 2006). Yet fading of exoptic images actually occurs with prolonged absence of retinal-image motion (Ditchburn & Ginsborg, 1952; Riggs & Ratliff, 1952), lasting from several seconds to several minutes, intervals much longer than the duration of natural visual fixation (Harris et al., 1988; Andrews & Coppola, 1999). Furthermore, under natural viewing conditions, fixational eye movements seem unnecessary to prevent fading as other body movements produce sufficient retinal image motion (Skavenski et al., 1979). Several other possible functions of fixational eye movements have been proposed besides the prevention of image fading (Averill & Weymouth, 1925; Marshall & Talbot, 1942; Arend, 1973; Ahissar & Arieli, 2001; Greschner et al., 2002; Ölveczky et al., 2003; Rucci & Casile, 2005). Our study joins these previous proposals in arguing for a role of fixational eye movements in the structuring, rather than the refreshing, of neural activity. The results of our simulations suggest that, in the early stages of the visual system, the incessant fixational motion of the retinal image contributes

to the establishment of efficient neural representations of natural stimuli.

References

ABBOTT, L.F., VARELA, J.A., SEN, K. & NELSON, S.B. (1997). Synaptic depression and cortical gain control. *Science* **275**, 220–224.

AHISSAR, E. & ARIELI, A. (2001). Figuring space by time. *Neuron* **32**, 185–201.

ALONSO, J.M., USREY, W.M. & REID, R.C. (1996). Precisely correlated firing in cells of the lateral geniculate nucleus. *Nature* **383**, 815–819.

ANDREWS, T.J. & COPPOLA, D.M. (1999). Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision Research* **39**, 2947–2953.

AREND, L.E. (1973). Spatial differential and integral operations in human vision: Implications of stabilized retinal image fading. *Psychological Review* **80**, 374–395.

ATICK, J.J. (1992). Could information theory provide an ecological theory of sensory processing? *Network: Computation in Neural Systems* **3**, 213–251.

ATICK, J.J. & REDLICH, A. (1992). What does the retina know about natural scenes? *Neural Computation* **4**, 196–210.

ATTNEAVE, F. (1954). Some informational aspects of visual perception. *Psychological Review* **61**, 183–193.

AVERILL, H.I. & WEYMOUTH, F.W. (1925). Visual perception and the retinal mosaic, II. The influence of eye movements on the displacement threshold. *Journal of Comparative Psychology* **5**, 147–176.

BARLOW, H.B. (1961). Possible principles underlying the transformations of sensory messages. In *Sensory Communication*, ed. ROSENBLITH, W.A., pp. 217–234. Cambridge, MA: MIT Press.

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*a*). The dynamics of primate M retinal ganglion cells. *Visual Neuroscience* **16**, 355–368.

BENARDETE, E.A. & KAPLAN, E. (1999*b*). Dynamics of primate P retinal ganglion cells: Responses to chromatic and achromatic stimuli. *Journal of Physiology (London)* **519.3**, 775–790.

BENARDETE, E.A., KAPLAN, E. & KNIGHT, B.W. (1992). Contrast gain control in the primate retina: P cells are not X-like, some M cells are. *Visual Neuroscience* **8**, 483–486.

BRADY, N. & FIELD, D.J. (2000). Local contrast in natural images: Normalisation and coding efficiency. *Perception* **29**, 1041–55.

CAI, D., DEANGELIS, G.C. & FREEMAN, R.D. (1997). Spatiotemporal receptive field organization in the lateral geniculate nucleus of cats and kitten. *Journal of Neurophysiology* **78**, 1045–1061.

CARANDINI, M., DEMB, J.B., MANTE, V., TOLHURST, D.J., DAN, Y., OLSHAUSEN, B.A., GALLANT, J.L. & RUST, N.C. (2005). Do we know what the early visual system does? *Journal of Neuroscience* **25**, 10577–97.

CASILE, A. & RUCCI, M. (2006). A theoretical analysis of the influence of fixational instability on the development of thalamo-cortical connectivity. *Neural Computation* **18**, 569–590.

CRANE, H.D. & STEELE, C.M. (1985). Generation V dual Purkinje-image eyetracker. *Applied Optics* **24**, 527–537.

CRONER, L.J. & KAPLAN, E. (1995). Receptive fields of P and M ganglion cells across the primate retina. *Vision Research* **35**, 7–24.

DAN, Y., ATICK, J.J. & REID, R.C. (1996). Efficient coding of natural scenes in the lateral geniculate nucleus: Experimental test of a computational theory. *Journal of Neuroscience* **16**, 3351–3362.

DITCHBURN, R.W. & GINSBORG, B.L. (1952). Vision with a stabilized retinal image. *Nature* **170**, 36–37.

DITCHBURN, R.W. & GINSBORG, B.L. (1953). Involuntary eye movements during fixation. *Journal of Physiology (London)* **119**, 1–17.

EIZENMAN, M., HALLETT, P. & FRECKER, R.C. (1985). Power spectra for ocular drift and tremor. *Vision Research* **25**, 1635–1640.

ENROTH-CUGELL, C., ROBSON, J.G., SCHWEITZER-TONG, D.E. & WATSON, A.B. (1983). Spatio-temporal interactions in cat retinal ganglion cells showing linear spatial summation. *Journal of Physiology (London)* **341**, 279–307.

FIELD, D.J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A, Optics and Image Science* **4**, 2379–94.

GRESCHNER, M., BONGARD, M., RUJAN, P. & AMMERMÜLLER, J. (2002). Retinal ganglion cell synchronization by fixational eye movements improves feature estimation. *Nature Neuroscience* **5**, 341–7.

GRESCHNER, M., THIEL, A., KRETZBERG, J. & AMMERMÜLLER, J. (2006). Complex spike-event pattern of transient on-off retinal ganglion cells. *Journal of Neurophysiology* **96**, 2845–56.

GUR, M., BEYLIN, A. & SNODDERLY, D.M. (1997). Response variability of neurons in primary visual cortex (V1) of alert monkeys. *Journal of Neuroscience* **17**, 2914–20.

HARRIS, C.M., HAINLINE, L., ABRAMOV, I., LEMERISE, E. & CAMENZULI, C. (1988). The distribution of fixation durations in infants and naive adults. *Vision Research* **28**, 419–432.

HERING, E. (1899). über die Grenzen der Sehschärfe. berichte der Königlichen Sächsischen Gesellshaft der Wissenschaften. *Mathematisch-Physische Klasse* **20**, 16–24.

KAPOULA, Z.A., ROBINSON, D.A. & HAIN, T.C. (1986). Motion of the eye immediately after a saccade. *Experimental Brain Research* **61**, 386–394.

LEE, B.B. (1996). Receptive field structure in the primate retina. *Vision Research* **36**, 631–44.

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., POKORNY, J., SMITH, V.C., MARTIN, P.R. & VALBERG, A. (1990). Luminance and chromatic modulation sensitivity of macaque ganglion cells and human observers. *Journal of the Optical Society of America A, Optics and Image Science* **7**, 2223–36.

LEOPOLD, D.A. & LOGOTHETIS, N.K. (1998). Microsaccades differentially modulate neural activity in the striate and extrastriate visual cortex. *Experimental Brain Research* **123**, 341–345.

MARSHALL, W.H. & TALBOT, S.A. (1942). Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity. In *Biological Symposia—Visual Mechanisms*, ed. KLUVER, H., volume 7, pp. 117–164. Lancaster, PA: Cattel.

MARTINEZ-CONDE, S., MACKNIK, S.L. & HUBEL, D.H. (2000). Microsaccadic eye movements and firing of single cells in the striate cortex of macaque monkeys. *Nature Neuroscience* **3**, 251–258.

MARTINEZ-CONDE, S., MACKNIK, S.L. & HUBEL, D.H. (2002). The function of bursts of spikes during visual fixation in the awake primate lateral geniculate nucleus and primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 13920–13925.

MARTINEZ-CONDE, S., MACKNIK, S.L., TRONCOSO, X.G. & DYAR, T.A. (2006). Microsaccades counteract fading during fixation. *Neuron* **49**, 297–305.

NAVON, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology* **9**, 353–383.

ÖLVECZKY, B.P., BACCUS, S.A. & MEISTER, M. (2003). Segregation of object and background motion in the retina. *Nature* **423**, 401–408.

PARKER, D.M., LISHMAN, J.R. & HUGHES, J. (1992). Temporal integration of spatially filtered visual images. *Perception* **21**, 147–60.

PUCHALLA, J.L., SCHNEIDMAN, E., HARRIS, R.A. & BERRY, M.J. (2005). Redundancy in the population code of the retina. *Neuron* **46**, 493–504.

RATLIFF, F. & RIGGS, L.A. (1950). Involuntary motions of the eye during monocular fixation. *Journal of Experimental Psychology* **40**, 687–701.

REID, R.C. & SHAPLEY, R.M. (2002). Space and time maps of cone photoreceptor signals in macaque lateral geniculate nucleus. *Journal of Neuroscience* **22**, 6158–6175.

RIGGS, L.A. & RATLIFF, F. (1952). The effects of counteracting the normal movements of the eye. *Journal of the Optical Society of America* **42**, 872–873.

RIGGS, L.A., RATLIFF, F., CORNSWEET, J.C. & CORNSWEET, T.N. (1953). The disappearance of steadily fixated visual test objects. *Journal of the Optical Society of America* **43**, 495–501.

ROELFSEMA, P.R., LAMME, V.A. & SPEKREIJSE, H. (2004). Synchrony and covariation of firing rates in the primary visual cortex during contour grouping. *Nature Neuroscience* **7**, 982–91.

RUCCI, M. & CASILE, A. (2004). Decorrelation of neural activity during

fixational instability: Possible implications for the refinement of V1 receptive fields. *Visual Neuroscience* **21**, 725–738.

RUCCI, M. & CASILE, A. (2005). Fixational instability and natural image statistics: Implications for early visual representations. *Network: Computation in Neural Systems* **16**, 121–138.

RUCCI, M. & DESBORDES, G. (2003). Contributions of fixational eye movements to the discrimination of briefly presented stimuli. *Journal of Vision* **3**, 852–64.

RUCCI, M., EDELMAN, G.M. & WRAY, J. (2000). Modeling LGN responses during free-viewing: A possible role of microscopic eye movements in the refinement of cortical orientation selectivity. *Journal of Neuroscience* **20**, 4708–4720.

RUDERMAN, D.L. (1994). Statistics of natural images. *Network: Computation in Neural Systems* **5**, 517–548.

SCHYNS, P.G. & OLIVA, A. (1994). From blobs to boundary edges: Evidence for time-and spatial-scale-dependent scene recognition. *Psychological Science* **5**, 195–200.

SHAPLEY, R.M. & VICTOR, J.D. (1979). Nonlinear spatial summation and the contrast gain control of cat retinal ganglion cells. *Journal of Physiology (London)* **290**, 141–61.

SINGER, W. & GRAY, C.M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience* **18**, 555–586.

SKAVENSKI, A.A., HANSEN, R.M., STEINMAN, R.M. & WINTERSON, B.J. (1979). Quality of retinal image stabilization during small natural and artificial body rotations in man. *Vision Research* **19**, 675–683.

SNODDERLY, D.M. (1987). Effect of light and dark environments on macaque and human fixational eye movements. *Vision Research* **27**, 401–415.

SNODDERLY, D.M., KAGAN, I. & GUR, M. (2001). Selective activation of visual cortex neurons by fixational eye movements: Implications for neural coding. *Visual Neuroscience* **18**, 259–77.

SNODDERLY, D.M. & KURTZ, D. (1985). Eye position during fixation tasks: Comparison of macaque and human. *Vision Research* **25**, 83–98.

SRINIVASAN, M.V., LAUGHLIN, S.B. & DUBS, A. (1982). Predictive coding: A fresh view of inhibition in the retina. *Proceedings of the Royal Society, Series B: Biological Sciences* **216**, 427–459.

STEINMAN, R.M., CUNITZ, R.J., TIMBERLAKE, G.T. & HERMAN, M. (1967). Voluntary control of microsaccades during maintained monocular fixation. *Science* **155**, 1577–1579.

STEINMAN, R.M., HADDAD, G.M., SKAVENSKI, A.A. & WYMAN, D. (1973). Miniature eye movement. *Science* **181**, 810–819.

STEINMAN, R.M. & LEVINSON, J.Z. (1990). The role of eye movements in the detection of contrast and spatial detail. In *Eye Movements and Their Role in Visual and Cognitive Processes*, ed. KOWLER, E., chapter 3, pp. 115–212. Elsevier Science Publishers BV.

STEVENSON, S.B. & ROORDA, A. (2005). Correcting for miniature eye movements in high resolution scanning laser ophthalmoscopy. In *Ophthalmic Technologies XV*, ed. MANNS, F., SODERBERG, P.G., HO, A., STUCK, B.E. & BELKIN, M., volume 5688 of *Proceedings of SPIE*, pp. 145–151. SPIE–The International Society for Optical Engineering, Bellingham, WA.

TULUNAY-KEESEY, Ü. & JONES, R.M. (1976). The effect of micromovements of the eye and exposure duration on contrast sensitivity. *Vision Research* **16**, 481–488.

USREY, W.M. & REID, R.C. (1999). Synchronous activity in the visual system. *Annual Review of Physiology* **61**, 435–456.

USREY, W.M., REPPAS, J.B. & REID, R.C. (1998). Paired-spike interactions and synaptic efficacy of retinal inputs to the thalamus. *Nature* **395**, 384–387.

VAN HATEREN, J.H. (1992). A theory of maximizing sensory information. *Biological Cybernetics* **68**, 23–29.

VAN HATEREN, J.H. & VAN DER SCHAAF, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society, Series B: Biological Sciences* **265**, 359–366.

VICTOR, J.D. (1987). The dynamics of the cat retinal X cell centre. *Journal of Physiology (London)* **386**, 219–246.

Appendix

The spatial decorrelation of neural activity observed in the model originated from the interaction between the temporal response characteristics of simulated cells and the second-order statistics of the modulations of luminance resulting from fixational eye movements. This Appendix explains why, in the case of natural images, these input fluctuations are spatially uncorrelated. For simplicity, we focus on the case of a one-dimensional spatial stimulus. The derivation below can be directly extended to the case of a two-dimensional stimulus.

Let us first consider the visual input to the retina in the absence of fixational eye movements. Given a static stimulus $L(x)$, the input signal entering the eye can be expressed as $I_S(x,t) = L(x)v(t)$, where $v(t) = 1$ at all times. Since $I_S(x,t)$ does not change with time, its power is given by $P_{I_S}(u,\omega) = |L(u)|^2\delta(\omega)$, where u and ω represent spatial and temporal frequencies, $\delta(\omega)$ is the unit impulse function, and $L(u)$ is the spatial power spectrum of $L(x)$. All power is on the spatial frequency axis, i.e., there is no power for temporal frequencies different from zero.

Fixational eye movements modulate the visual input to the retina. The input signal generated by viewing the stimulus $L(x)$ when the eye moves following a trajectory $\xi(t)$ can be expressed as $I(x,t) = L[x + \xi(t)]$. By the conservation of energy, when the same set of images are observed with and without eye movements, the total power of $I(x,t)$ is the same as that of $I_S(x,t)$. However, I and I_S have a different distribution of spectral density. As a consequence of the motion of the eye, some of the power of $I_S(x,t)$ spreads to temporal frequencies different from zero.

To study the impact of fixational eye movements on the visual input, it is useful to approximate $I(x,t)$ with a Taylor expansion. For small eye movements, we can consider a second-order approximation

$$I(x,t) \approx I(x,0) + \frac{\partial I(x,t)}{\partial x}\,\xi(t) + \frac{1}{2}\,\frac{\partial^2 I(x,t)}{\partial x^2}\,\xi^2(t),$$

which, in the Fourier domain, becomes

$$I_T(u,\omega) \approx L(u)\left[\delta(\omega) + ju\xi_T(\omega) - \frac{1}{2}\,u^2\Gamma_T(\omega)\right],$$

where $I_T(u,\omega) = \mathcal{F}_T\{I(x,t)\}$ is the Fourier Transform of the signal $I(x,t)$ over a finite period T, $\xi_T(\omega) = \mathcal{F}_T\{\xi(t)\}$, and $\Gamma_T(\omega) = \mathcal{F}_T\{\xi^2(t)\}$. This equation allows an approximation of the power spectrum of the input signal $I(x,t)$:

$$P_I(u,\omega) \approx |L(u)|^2\{\delta(\omega) + u^2[|\xi(\omega)|^2 - \delta(\omega)\sigma_\xi^2]\}, \tag{7}$$

where σ_ξ^2 is the variance of the motion of the eye, and elimination of the subscript T indicates the calculation of the limit, *i.e.* $\xi(\omega) = \lim_{T\to\infty}(1/T)\xi_T(\omega)$. Eq. (7) shows that the way the power spreads to the temporal domain because of oculomotor activity can be approximated by $u^2|\xi(\omega)|^2$.

An M cell responds poorly to stationary stimuli but is sensitive to temporal changes within the region covered by its receptive field. To examine the visual input effective in driving this neuron, we can estimate the dynamic input power $P^D(u)$, that is, the total amount of power at each spatial frequency which, as a consequence of the motion of the eye, becomes available at nonzero temporal frequency. Since P_I and P_{I_S} possess the same amount of total power at each spatial frequency, $P^D(u)$ can be calculated by subtracting $P_I(u,0)$ from $P_{I_S}(u,0)$. From Eq. (7), after some simplifications, we obtain:

$$P^D(u) = P_{I_S}(u,0) - P_I(u,0) \approx u^2|L(u)|^2\sigma_\xi^2. \tag{8}$$

Eq. (8) shows that, at low spatial frequencies, where the Taylor approximation holds and within the range of sensitivity of M cells, the amount of power that spreads into the temporal domain increases with spatial frequency as u^2. With presentation of natural images, this term counterbalances the power spectrum of the stimulus, $|L(u)|^2 \approx u^{-2}$, to yield a constant dynamic power at all spatial frequencies: $P^D(u) \approx \sigma_\xi^2$. Therefore, with presentation of natural images, an ideal neuron equally sensitive to all nonzero temporal frequencies is driven by a visual input signal with flat spatial spectrum, *i.e.* a signal that is spatially uncorrelated. Consistent with this approximation, in our model, the responses of pairs of M cells with non-overlapping receptive fields became uncorrelated as soon as neurons stopped responding to the DC level of the current fixation.