Network: Computation in Neural Systems
June/September 2005; 16(2/3): 121–138

Taylor & Francis
Taylor & Francis Group

Fixational instability and natural image statistics: Implications for early visual representations

MICHELE RUCCI[1] & ANTONINO CASILE[2]

[1]*Department of Cognitive and Neural Systems, Boston University, Boston, MA, USA, and*
[2]*Department of Cognitive Neurology, University Clinic, Tübingen, Germany*

(Received 1 December 2004; revised 25 June 2005; accepted 30 June 2005)

Abstract
Under natural viewing conditions, small movements of the eye, head and body prevent the maintenance of a steady direction of gaze. It is known that stimuli tend to fade when they are stabilized on the retina for several seconds. However, it is unclear whether the physiological motion of the retinal image serves a visual purpose during the brief periods of natural visual fixation. This study examines the impact of fixational instability on the statistics of the visual input to the retina and on the structure of neural activity in the early visual system. We show that fixational instability introduces a component in the retinal input signals that, in the presence of natural images, lacks spatial correlations. This component strongly influences neural activity in a model of the LGN. It decorrelates cell responses even if the contrast sensitivity functions of simulated cells are not perfectly tuned to counter-balance the power-law spectrum of natural images. A decorrelation of neural activity at the early stages of the visual system has been proposed to be beneficial for discarding statistical redundancies in the input signals. The results of this study suggest that fixational instability might contribute to the establishment of efficient representations of natural stimuli.

Keywords: *Lateral geniculate nucleus, retina, eye movement, visual fixation, drift, saccade, redundancy reduction*

Introduction

During natural viewing, the projection of the visual scene on the retina is never stationary. Saccades relocate the direction of gaze every few hundred milliseconds. Even during the brief periods in between saccades, the physiological instability of visual fixation keeps the retinal image in permanent motion. Several sources contribute to this constant jittering of the eye. Fixational eye movements, of which we are not aware, alternate small saccades with periods of drifts even when subjects are instructed to maintain steady fixation (Ratliff & Riggs 1950; Ditchburn 1955; Steinman et al. 1973). Following macroscopic redirection of gaze, other small eye movements, such as corrective saccades and post-saccadic drifts, are likely to occur. Furthermore, outside of the controlled conditions of a laboratory when the head is not constrained by a bite bar, movements of the head and body, as well as imperfections in the vestibulo-ocular reflex, significantly amplify the motion of the retinal image (Skavenski et al. 1979). Fixational instability moves the stimulus by an amount that should be clearly

Correspondence: M. Rucci, Cognitive and Neural Systems Department, Boston University, 677 Beacon Street, Boston, MA 02215, USA. Tel: 617 353 7671. Fax: 617 353 7755. E-mail: rucci@cns.bu.edu

ISSN: 0954-898X print / ISSN 1361-6536 online
DOI: 10.1080/09548980500300507

visible (see Murakami & Cavanagh 1998 for a striking demonstration of the motion of the retinal image during fixation). It is remarkable that the brain is capable of constructing a stable percept despite this constant jitter.

Little is known about the purposes of the physiological self-motion of the retinal image. While it is possible that fixational instability is simply the result of an imperfect motor system, several theories have argued for possible visual functions (Marshall & Talbot 1942; Ditchburn 1955; Ahissar & Arieli 2001). It is often claimed that small saccades are needed to refresh neuronal responses and prevent the disappearance of a stationary scene (reviewed in Martinez-Conde et al. 2004). This claim, which follows the observation that images tend to fade when they are stabilized on the retina, has, however, remained controversial. Image fading is typically observed when retinal motion is eliminated for several seconds or minutes, a period much longer than the typical duration of natural visual fixation (Ditchburn & Ginsborg 1952; Riggs & Ratliff 1952; see Steinman & Levinson 1990 for a comprehensive review of early stabilization experiments). While it has been observed that entoptic images, images that are intrinsically perfectly stabilized, can disappear in less than 80 ms in the absence of retinal image motion (Coppola & Purves 1996), disagreement exists over whether perfectly stabilized external images disappear completely even with long stimulus presentations (Arend & Timberlake 1986; Ditchburn 1987; Steinman & Levinson 1990). No fading of the stimulus was observed in recent experiments that stabilized external stimuli for brief periods comparable to the duration of natural visual fixation (Rucci & Desbordes 2003).

Yet, evidence emerging from neurophysiological and modeling studies suggests that fixational instability is an important component of the way information about a visual scene is acquired and encoded in the brain. Recent neurophysiological investigations have shown that fixational eye movements strongly influence the activity of neurons in several areas of the monkey's brain (Gur et al. 1997; Leopold & Logothetis 1998; Martinez-Conde et al. 2000; Snodderly et al. 2001). Analysis of the impact of fixational instability in the retina indicates that it may contribute to the improvement of feature estimation (Greschner et al. 2002) and figure/ground segregation (Olveczky et al. 2003). Furthermore, modeling studies that simulated neural responses during free-viewing suggest that fixational instability profoundly affects the statistics of thalamic (Rucci et al. 2000) and thalamocortical activity (Rucci & Casile 2004; Parsons & Rucci 2005).

Building on the results of our previous modeling work, this study proposes an alternative theory for the existence of fixational instability. Instead of regarding the jitter of visual fixation as necessary for *refreshing* neuronal responses, it is argued that fixational instability is essential for properly *structuring* neural activity in the early visual system into a format that is suitable for processing at later stages. This theory can be cast within the framework of the redundancy reduction hypothesis (Barlow 1961). It is proposed that fixational instability is part of a strategy of acquisition of visual information that enables compact representations in the presence of natural visual input. We provide a mathematical formulation of this theory by describing the impact of fixational instability on the second-order statistics of neural activity in a model of the LGN.

Neural decorrelation and early visual processing

Images of natural scenes tend to vary smoothly over space and time. As a consequence, it is often possible to predict the intensity of a pixel at a certain location on the basis of the intensity values of adjacent pixels. In the visual system, a moderate degree of redundancy

in the input signals can be beneficial for increasing the signal-to-noise ratio and establishing robust representations in the presence of noise. However, if not eliminated, an elevated level of redundancy implies that many neurons will represent the same information over relatively long periods of time.

It is a long-standing proposal that an important function of early visual processing is the removal of part of the redundancy that characterizes natural visual input (Attneave 1954; Barlow 1961). Less redundant signals enable more compact representations, in which the same amount of information can be represented by smaller neuronal ensembles. Although in the cortex the huge number of neurons appears to be able to afford redundant sensory representations (Barlow 2002), a compact neural code is desirable in the early stages of the visual system. The neural pathway from the retina to the cortex creates a bottleneck in the processing of visual information. A compression of information at this level would be beneficial not only for increasing the efficiency of transmission, but also for emphasizing the most interesting elements of the visual stimulus, those that cannot be predicted from the mere knowledge of the statistical properties of the environment. That is, compact early representations may facilitate the computational tasks of later processing stages where these elements are extracted and used.

While several methods for eliminating input redundancies can be followed (Field 1994), a possible approach is the removal of the pairwise correlations between the intensity values of nearby pixels. Elimination of these spatial correlations allows efficient representations in which neuronal responses tend to be less statistically dependent. This idea is explained in Figure 1. In the frequency domain, the broad spatial correlations of images of natural scenes are reflected by the scaling invariance of their power spectrum. Several studies have shown that the $1/u^2$ approximation of this power spectrum holds for a large variety of natural scenes (Field 1987; Burton & Moorhead 1987; Ruderman & Bialek 1994). In the presence of natural visual stimulation, a hypothetical linear neuronal element with transfer function proportional to the spatial frequency u would counterbalance the uneven spatial power density of natural images and produce a spatially decorrelated output. It has been proposed that neurons in the retina may operate as the filter of Figure 1 at low spatial frequencies (Atick & Redlich 1992). By taking psychophysical measurements of human contrast sensitivity to represent a cumulative envelope of the frequency responses of retinal cells, Atick and Redlich (1992) have argued that these neurons may eliminate the broad spatial correlation of natural scenes and yet enable robust representations in the presence of noise. However, data from neurophysiological recordings with both the cat and the monkey have shown that the frequency responses of neurons in the retina and the LGN deviate substantially from linearity at

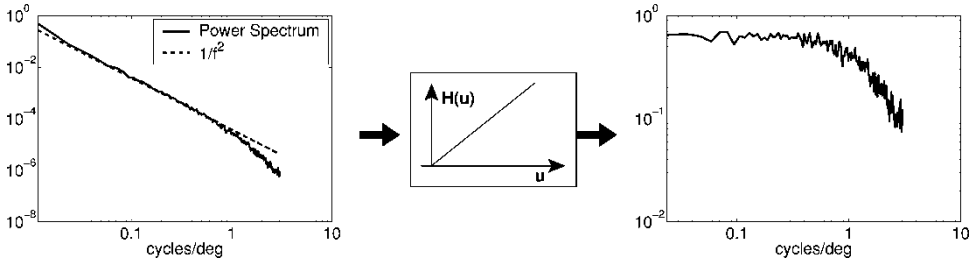

Figure 1. Whitening of natural visual input produced by a linear filter with transfer function $H(u)$ proportional to the spatial frequency u. The power spectrum of images of natural scenes is well approximated by $1/u^2$. Filtering natural images with $H(u) \propto u$ produces a spatially decorrelated output with flat power spectrum. The two graphs show the spectral density functions of the input and output signals.

low spatial frequencies (So & Shapley 1981; Linsenmeier et al. 1982; Derrington & Lennie 1984; Croner & Kaplan 1995; Cheng et al. 1995). This deviation is not consistent with a decorrelation of their neural responses.

It is important to observe that the approach of Figure 1 focuses exclusively on the spatial domain and neglects the dynamics of neural activity. During natural viewing, cell responses are affected not only by the spatial characteristics of the stimulus, but also by the motion of the retinal image due to the natural instability of visual fixation. This study investigates the influence of fixational instability on the output correlation of spatiotemporal filters with spatial characteristics that are not optimal for decorrelation.

Measuring correlated activity in a model of the LGN

In this study, we modeled the responses of geniculate cells to visual stimuli and examined the second-order statistics of LGN activity under different viewing conditions. For each investigated condition, we summarized the spatial structure of correlated activity by means of the average width S_α of pools of coactive units. As illustrated in Figure 2, S_α was estimated on the basis of the levels of correlation between the responses of cells with receptive fields at various separations:

$$\hat{c}_{\alpha\alpha}(\mathbf{x}) = \langle \alpha_{\mathbf{y}}(t)\alpha_{\mathbf{z}}(t)\rangle_{\mathcal{I},\Xi,T} \tag{1}$$

where $\alpha_{\mathbf{y}}(t)$ and $\alpha_{\mathbf{z}}(t)$ are the responses of two ON-center units with receptive fields centered at $\mathbf{y}$ and $\mathbf{z}$, and $\mathbf{x} = \mathbf{y} - \mathbf{z}$ is the separation between the receptive field centers. This average was evaluated over time T, ensembles of stimuli $\mathcal{I}$, and (when present) sequences of eye movements Ξ. The average width of pools of coactive units was defined as $S_\alpha = 2D$, where D was the minimum separation at which the mean level of correlation, averaged over all possible orientations ϕ, fell below a specified percentage of the value measured at separation zero:

$$\langle \hat{c}_{\alpha\alpha}(\mathbf{x})|_{||\mathbf{x}||=D}\rangle_\phi \leq \epsilon \hat{c}_{\alpha\alpha}(\mathbf{0})$$

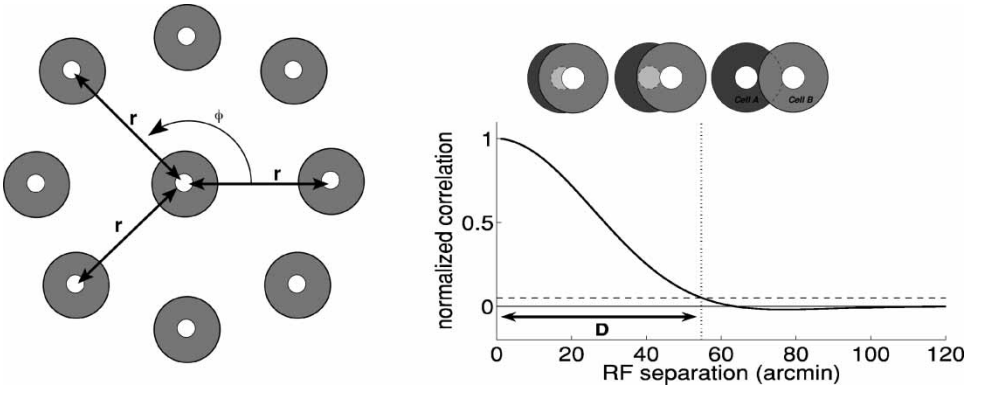

Figure 2. Estimating the spatial extent of correlated activity. (Left) Levels of correlation were measured between pairs of cells with receptive fields located at all possible separations r and relative angles ϕ. (Right) The spatial extent of correlated activity was defined as the double of the minimum separation D at which the radial average of the correlation was less than 5% of the level measured for perfectly overlapping receptive fields.

The results presented in the following sections were obtained with $\epsilon = 0.05$. In this paper, we refer to $\hat{c}_{\alpha\alpha}(\mathbf{x})$ as the correlation map or the pattern of correlated activity, and to S_α as the spatial extent of correlation measured in the considered condition.

Modeling the responses of LGN cells

Geniculate responses to visual stimuli were estimated by means of the model proposed by Cai et al. (1997). This model gives the deviation in the mean instantaneous firing rate with respect to the level of spontaneous activity on the basis of the convolution between the visual input I and the cell spatiotemporal kernel h_α:

$$\alpha(t) = [h_\alpha(\mathbf{x}, t) \star I(\mathbf{x}, t)]_\theta = \left[\int_0^t \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_\alpha(x', y', t') I(x - x', y - y', t - t')\, dx'\, dy'\, dt' \right]_\theta$$

where the operator $[x]_\theta$ indicates rectification with threshold θ ($[x]_\theta = x - \theta$ if $x > \theta$, and $[x]_\theta = 0$ if $x \leq \theta$). The spatiotemporal kernel $h_\alpha(\mathbf{x}, t)$ was given by sum of two space-time separable terms representing the contributions from the center and periphery of the receptive field:

$$h_\alpha(\mathbf{x}, t) = h_c(\mathbf{x}, t) + h_s(\mathbf{x}, t) = g_c(t)\, f_c(\mathbf{x}) + g_s(t)\, f_s(\mathbf{x}).$$

For both the center and the surround, the spatial element was a 2D Gaussian function

$$f_c(\mathbf{x}) = A_c e^{-\frac{|\mathbf{x}|^2}{2\sigma_c^2}}, \quad f_s(\mathbf{x}) = A_s e^{-\frac{|\mathbf{x}|^2}{2\sigma_s^2}}$$

and the temporal elements possessed identical biphasic profiles with the response of the surround delayed with respect to that of the center:

$$g_c(t) = k_1 \Gamma(t; c_1, t_1, n_1) - k_2 \Gamma(t; c_2, t_2, n_2), \quad g_s(t) = g_c(t - t_D)$$

where

$$\Gamma(t; t_0, c, n) = \frac{[c(t - t_0)]^n e^{-c(t - t_o)}}{n^n e^{-n}}.$$

Spatial and temporal parameters followed neurophysiological measurements to replicate the responses of ON-center nonlagged X cells in the cat. Spatial parameters replicated the typical kernels of geniculate cells with receptive fields located at various angles of visual eccentricity within 5° and 25° (Linsenmeier et al. 1982). Temporal parameters followed the data by Cai et al. (1997). The results presented in this paper were obtained with $t_1 = t_2 = 0$, $n_1 = n_2 = 2$, $k_1 = 1$, $k_2 = 0.6$, $c_1 = 60\ \mathrm{s}^{-1}$, $c_2 = 40\ \mathrm{s}^{-1}$.

Estimating correlated activity

Although rectification is a critical component of geniculate responses, previous simulations in which the rectification threshold of geniculate cells was systematically varied to eliminate up to 50% of the dynamic range of responses have shown a modest influence of the level of rectification on the patterns of correlated activity (Rucci et al. 2000). For this reason, in this study we neglected the rectification operated by geniculate units. This assumption greatly simplified the mathematical analysis, as it enabled the use of linear system theory. In the absence of rectification, the correlation map $\hat{c}_{\alpha\alpha}$ can be obtained on the basis of the spatiotemporal correlation function of the visual input $c_{SS}(\mathbf{x}, t)$ (the correlation between the luminance of pixels at a separation $\mathbf{x} = (x, y)$ measured at a time lag t), or equivalently

the input power spectrum $R_{SS}(\mathbf{u}, w)$, where $\mathbf{u}$ and w represent the spatial and temporal frequencies:

$$c_{\alpha\alpha}(\mathbf{x}, t) = \mathcal{F}^{-1}\{\bar{H}_\alpha H_\alpha R_{SS}\} \tag{2}$$

where $\mathcal{F}^{-1}$ represents the inverse Fourier transform, $H_\alpha(\mathbf{u}, w)$ is the spatiotemporal Fourier transform of the kernel $h_\alpha(\mathbf{x}, t)$ of a geniculate cell, and the bar indicates complex conjugation. The correlation map $\hat{c}_{\alpha\alpha}$ in Equation 1 is the section at time $t = 0$ of the spatiotemporal correlation function: $\hat{c}_{\alpha\alpha}(\mathbf{x}) = c_{\alpha\alpha}(\mathbf{x}, 0)$.

In the absence of rectification, the response of a geniculate unit $\alpha(t)$ is determined by the sum of the outputs of the center and surround linear filters, each receiving the same input signal $S(\mathbf{x}, t)$:

$$\alpha(t) = [h_c(\mathbf{x}, t) + h_s(\mathbf{x}, t)] \star S(\mathbf{x}, t)$$

Thus, the spatiotemporal power spectrum of LGN activity is given by:

$$R_{\alpha\alpha} = (|H_c|^2 + |H_s|^2 + \bar{H}_c H_s + H_c \bar{H}_s) R_{SS}(\mathbf{u}, w)$$

where $H_c(\mathbf{u}, w)$ and $H_s(\mathbf{u}, w)$ represent the Fourier transform of the center and surround kernels, h_c and h_s, respectively. Under the model assumption of space time separability of the contributions of the center and surround, $H_c(\mathbf{u}, w) = F_c(\mathbf{u})G_c(w)$ and $H_s(\mathbf{u}, w) = F_s(\mathbf{u})G_s(w)$, we obtain:

$$R_{\alpha\alpha} = |G(w)|^2[|F_c(\mathbf{u})|^2 + |F_s(\mathbf{u})|^2 + 2F_c(\mathbf{u})F_s(\mathbf{u})\cos(wt_D)]R_{SS}(w, \mathbf{u}) \tag{3}$$

where $|G| = |G_c| = |G_s|$ and we have taken into consideration the fact that, due to the symmetry of the spatial kernels, F_c and F_s are real functions.

This expression of $R_{\alpha\alpha}$ can be further simplified in the presence of input stimuli with power spectra that are separable in their spatial and temporal components, $R_{SS}(\mathbf{u}, w) = R^T_{SS}(w)R^S_{SS}(\mathbf{u})$. In this case, separation of the space and time domains gives the following structure of correlated activity:

$$\begin{aligned} c_{\alpha\alpha}(\mathbf{x}, t) = {} & \mathcal{F}_T^{-1}\left\{|G(w)|^2 R^T_{SS}(w)\right\}\mathcal{F}_S^{-1}\left\{[|F_c(\mathbf{u})|^2 + |F_s(\mathbf{u})|^2]R^S_{SS}(\mathbf{u})\right\} \\ & + 2\mathcal{F}_T^{-1}\left\{|G(w)|^2\cos(wt_D)R^T_{SS}(w)\right\}\mathcal{F}_S^{-1}\left\{F_c(\mathbf{u})F_s(\mathbf{u})R^S_{SS}(\mathbf{u})\right\} \end{aligned} \tag{4}$$

where $\mathcal{F}_T^{-1}$ and $\mathcal{F}_S^{-1}$ indicate the operations of inverse Fourier Transform in time and space.

Correlated activity in the absence of retinal image motion

To establish a reference baseline, we first analyzed the structure of correlated activity in the absence of eye movements, that is, when the model was passively presented with visual stimulation. This condition simulated the experimental procedure of retinal stabilization, in which the stimulus is displayed at a fixed location on the retina and the normal motion of the retinal image is eliminated. In this paper, we focused on the specific case of static images, images for which the temporal component of the power spectrum is $R^T_{SS}(w) = 2\pi\delta(w)$. In this case, substitution of the input power spectrum in Equation 4 yields a correlation map:

$$\hat{c}_{\alpha\alpha}(\mathbf{x}) = g_o^2\mathcal{F}_S^{-1}\left\{[|F_c(\mathbf{u}) + F_s(\mathbf{u})|^2]R^S_{SS}(\mathbf{u})\right\} \tag{5}$$

where g_o is the integral of the temporal impulse response of the center and surround. Not surprisingly, in the absence of retinal motion, the spatial distribution of coactive geniculate units depended exclusively on the spatial properties of the stimulus and cell kernels. The dynamic characteristics of cell responses contributed only to the proportionality constant g_o.

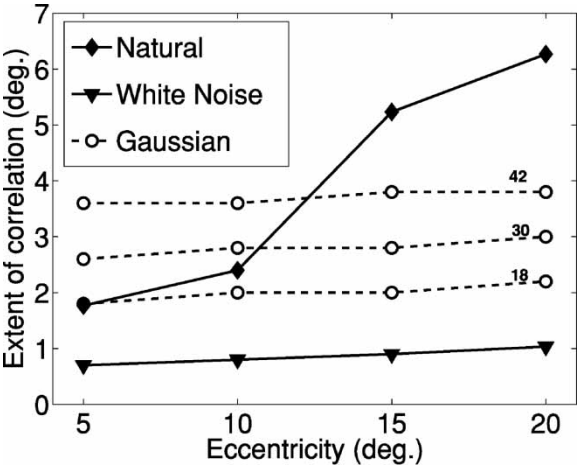

Figure 3. Spatial extent of correlated activity in the model with various types of visual input. Visual stimuli were presented in the absence of the retinal motion produced by fixational instability. The average widths of pools of coactive units are shown for units that replicated the mean characteristics of LGN units located at several angles of visual eccentricity. Results are shown for three types of visual stimulation: natural images (Natural), spatially uncorrelated input (White Noise), and images with Gaussian spatial correlation (Gaussian). The numbers on the curves indicate the standard deviation of the Gaussian correlation expressed in minutes of arc.

Figure 3 shows the spatial extent of correlated activity measured from Equation 5 for various types of stimuli. Data points represent the average widths of pools of coactive units present at different angles of visual eccentricity. In the case of spatially uncorrelated visual input, such as when the model was driven by patterns of white noise, $R^S_{SS}(\mathbf{u}) = 1$ in Equation 5, and the structure of correlated activity was determined exclusively by the spatial characteristics of geniculate units. The responses of two geniculate units were strongly correlated only when a substantial overlap was present between the ON and OFF subregions in their receptive fields. As illustrated in Figure 3, the spatial extent of correlation was smaller than 1° at all the considered angles of visual eccentricity.

To examine the influence of input spatial correlations on the structure of geniculate activity, we first analyzed the case of input signals with Gaussian second-order statistics:

$$c_{SS}(\mathbf{x}, t) = \frac{1}{2\pi\sigma_I} e^{-|\mathbf{x}|^2/2\sigma_I^2}.$$

This type of visual input allows the derivation from Equation 2 of an explicit formula for the structure of correlated activity:

$$\begin{aligned} c_{\alpha\alpha}(r) \propto & \frac{A_c^2\sigma_c}{\sqrt{2+\sigma_I^2/\sigma_c^2}} e^{-r^2/2(2\sigma_c^2+\sigma_I^2)} + \frac{A_s^2\sigma_s}{\sqrt{2+\sigma_I^2/\sigma_s^2}} e^{-r^2/2(2\sigma_s^2+\sigma_I^2)} \\ & + \frac{2A_cA_s\sigma_c\sigma_s}{\sqrt{\sigma_c^2+\sigma_s^2+\sigma_I^2}} e^{-r^2/2(\sigma_c^2+\sigma_s^2+\sigma_I^2)} \end{aligned} \tag{6}$$

where $r = |\mathbf{x}|$. Figure 3 shows the spatial extent of correlation obtained from Equation 6 for different values of σ_I. It is clear that in the model, small increments in the width of input correlations led to significantly wider pools of simultaneously active geniculate units. The

mean spatial extent of correlated activity over all the considered angles of visual eccentricity increased from 2 deg with $\sigma_I = 18$ arcmin to 3.8 deg with $\sigma_I = 42$ arcmin.

Although the second-order statistics of natural images is not Gaussian, images of natural scenes are known to possess broad spatial correlations. Figure 3 shows the spatial extent of correlated activity measured in the model when an estimate of the power spectrum of natural images was applied to Equation 5. This estimate was obtained over a set of images extracted from a public domain database (van Hateren & van der Schaaf 1998). Its radial mean was best interpolated by $S(u) \propto u^{-2.02}$, which is consistent with previous measurements (Field 1987; Burton & Moorhead 1987; Ruderman & Bialek 1994). Similar to the case of Gaussian correlations, in the presence of natural images the model also responded with wide pools of coactive units. As shown by Figure 3, the spatial extent of correlation measured around 15 deg of visual eccentricity was 5.2 deg, a value almost six times wider than that obtained in the presence of uncorrelated input.

Figure 4 explains the origins of the broad output correlations measured in the presence of spatially correlated input. The average spatial kernels measured by Linsenmeier et al. (1982) at 5° and 15° of visual eccentricity are shown in both the space (Figure 4a) and frequency domains (Figure 4b). These are the kernels that were used to obtain the results shown in Figure 3. As illustrated by Figure 4b, the contrast sensitivity functions of these geniculate units deviate significantly from a linear proportionality with spatial frequency. The deviation at high spatial frequencies is to be expected because of the cut-off frequency of geniculate units. This deviation has actually been proposed to contribute to the robustness of the neural code in the presence of noise (Atick & Redlich 1992). It is the departure from linearity in the low frequency range that is responsible for the failure of these kernels in decorrelating cell responses. The broad spatial correlations of natural images are a consequence of the high power of harmonics with low spatial frequencies. An elevated sensitivity to low spatial frequencies causes geniculate responses to be correlated over a wide range of separations.

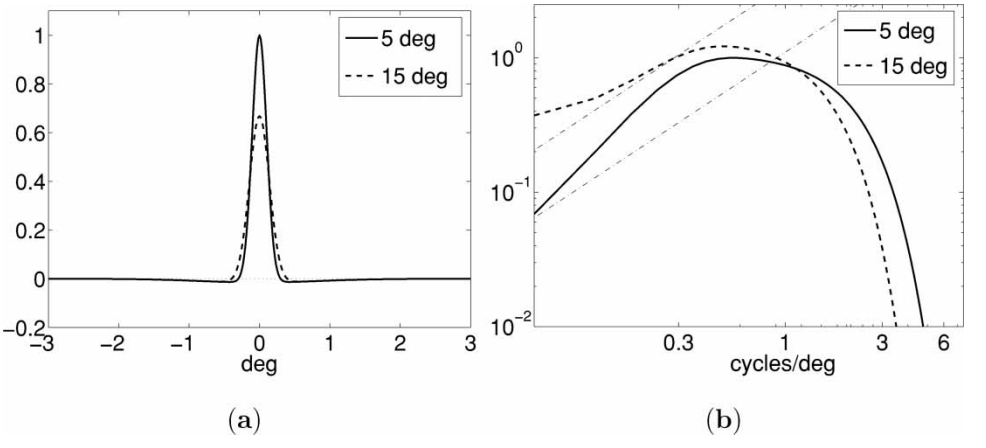

Figure 4. The response characteristics of geniculate cells are not optimal to decorrelate neural responses in the presence of natural images. (*a*) Sections of the spatial receptive fields of modeled units. The two curves show the receptive field amplitudes as a function of space for units located around 5° and 25° of visual eccentricity. (*b*) Contrast sensitivity functions of the two units. Note that at low spatial frequencies both responses deviate from linearity. The dashed lines show the frequency responses of filters that produce maximal decorrelation. Receptive field parameters are from (Linsenmeier et al. 1982).

Correlated activity during fixational instability

The results of Figure 3 are a direct consequence of the wide spatial correlations that characterize natural scenes. These broad correlations activated simultaneously geniculate cells with non-overlapping receptive fields. In the analysis of Figure 3, however, visual input was presented to the model without considering the retinal motion produced by the physiological instability of visual fixation. Under natural viewing conditions, the spatiotemporal structure of the signals entering the eyes depends on both the statistical characteristics of the visual scene and the movements performed by the eye during the acquisition of visual information.

To analyze the effect of fixational instability on the statistics of geniculate activity, it is useful to approximate the input image in a neighborhood of a fixation point $\mathbf{x}_0$ by means of its Taylor series (see also Casile & Rucci 2005):

$$I(\mathbf{x}) \approx I(\mathbf{x}_0) + \nabla I(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)^T + o(|\mathbf{x} - \mathbf{x}_0|^2) \tag{7}$$

If the self-motion of the retinal image is sufficiently small, high-order derivatives can be neglected, and the input to a location $\mathbf{x}$ on the retina can be approximated during visual fixation by its first-order expansion:

$$S(\mathbf{x}, t) \approx I(\mathbf{x}) + \boldsymbol{\xi}^T(t) \cdot \nabla I(\mathbf{x}) = I(\mathbf{x}) + \tilde{I}(\mathbf{x}, t) \tag{8}$$

where $\boldsymbol{\xi}(t) = [\xi_x(t), \xi_y(t)]$ is the trajectory of the center of gaze during the period of fixation, t is the time elapsed from fixation onset, $I(\mathbf{x})$ is the visual input at $t = 0$, and $\tilde{I}(\mathbf{x}, t) = \frac{\partial I(\mathbf{x})}{\partial x}\xi_x(t) + \frac{\partial I(\mathbf{x})}{\partial y}\xi_y(t)$ is the dynamic fluctuation in the visual input produced by fixational instability.

Equation 8 allows an analytical estimation of the power spectrum of the signal entering the eye during fixational instability. Since according to Equation 8 the retinal input $S(\mathbf{x}, t)$ can be approximated by the sum of the two contributions I and $\tilde{I}$, its power spectrum R_{SS} consists of three terms:

$$R_{SS} \approx R_{II} + R_{\tilde{I}\tilde{I}} + 2R_{I\tilde{I}} \tag{9}$$

In this paper we modeled fixational instability by means of an ergodic process with zero moments of the first-order and uncorrelated components along the two axes, that is, $\langle \boldsymbol{\xi} \rangle_T = \mathbf{0}$ and $R_{\xi_x\xi_y}(t) = 0$. With these assumptions, $R_{I\tilde{I}}$ in Equation 9 is zero, and the power spectrum of the visual input is simply given by:

$$R_{SS} \approx R_{II} + R_{\tilde{I}\tilde{I}} \tag{10}$$

where R_{II} is the power spectrum of the stimulus, and $R_{\tilde{I}\tilde{I}}$ depends on both the stimulus and fixational instability.

Substitution of Equation 10 into Equation 3 yields the power spectrum of geniculate activity:

$$R_{\alpha\alpha} \approx |G|^2[|F_c|^2 + |F_s|^2 + 2F_cF_s\cos(wt_D)]R_{II} + |G|^2[|F_c|^2 + |F_s|^2 + 2F_cF_s\cos(wt_D)]R_{\tilde{I}\tilde{I}}$$

That is, similar to the retinal input, the power spectrum of geniculate activity can be approximated by the sum of two separate elements:

$$R_{\alpha\alpha} \approx R^S_{\alpha\alpha} + R^D_{\alpha\alpha} \tag{11}$$

with

$$R^S_{\alpha\alpha} = |G|^2[|F_c|^2 + |F_s|^2 + 2F_cF_s\cos(wt_D)]R_{II}$$
$$R^D_{\alpha\alpha} = |G|^2[|F_c|^2 + |F_s|^2 + 2F_cF_s\cos(wt_D)]R_{\tilde{I}\tilde{I}}$$

Only $R^D_{\alpha\alpha}$ depends on fixational instability. The first term, $R^S_{\alpha\alpha}$ is determined by the power spectrum of the stimulus and the characteristics of geniculate cells, but does not depend on the motion of the eye during the acquisition of visual information.

The pattern of correlation between the responses of geniculate cells, $\hat{c}_{\alpha\alpha}(\mathbf{x})$ (see Equation 1), can be derived from Equation 11:

$$\hat{c}_{\alpha\alpha}(\mathbf{x}) \approx \hat{c}^S_{\alpha\alpha}(\mathbf{x}) + \hat{c}^D_{\alpha\alpha}(\mathbf{x}) \tag{12}$$

where

$$\hat{c}^S_{\alpha\alpha}(\mathbf{x}) = \mathcal{F}^{-1}\left\{ R^S_{\alpha\alpha}(\mathbf{u}, w)\right\}\Big|_{t=0} \quad \text{and} \quad \hat{c}^D_{\alpha\alpha}(\mathbf{x}) = \mathcal{F}^{-1}\left\{ R^D_{\alpha\alpha}(\mathbf{u}, w)\right\}\Big|_{t=0}$$

In the presence of static images, $R_{II}(w) = 2\pi\delta(w)$, and $R^S_{\alpha\alpha}$ and $R^D_{\alpha\alpha}$ provide, respectively, a static and a dynamic contribution to the spatiotemporal correlation of geniculate activity. The first term gives a correlation map identical to that obtained with presentation of the stimulus in the absence of fixational instability (see Equation 5):

$$\hat{c}^S_{\alpha\alpha}(\mathbf{x}) = g_o^2 \mathcal{F}^{-1}\left\{(|F_c + F_s|^2) R^S_{II}(\mathbf{u})\right\} \tag{13}$$

To evaluate $\hat{c}^D_{\alpha\alpha}(\mathbf{x})$ we first need to determine $R_{\tilde{I}\tilde{I}}(\mathbf{u}, w)$. From Equation 8, it follows that

$$\tilde{I}(\mathbf{u}, w) = j u_x I(\mathbf{u}) \xi_x(w) + j u_y I(\mathbf{u}) \xi_y(w)$$

and, given our assumption of uncorrelated components of fixational instability, approximating the power spectrum via finite Fourier Transform yields:

$$R_{\tilde{I}\tilde{I}}(\mathbf{u}, w) = \lim_{T\to\infty} \left\langle \frac{1}{T} |\tilde{I}_T(\mathbf{u}, w)|^2 \right\rangle_{\xi,\mathcal{I}} = R_{\xi\xi}(w) R_{II}(\mathbf{u}) |\mathbf{u}|^2 \tag{14}$$

where $\tilde{I}_T$ is the Fourier Transform of a signal of duration T, and we have also assumed that the second-order statistics of fixational instability are identical along the two axes. It is clear that the presence of the term $\mathbf{u}^2$ in Equation 14 compensates for the scaling invariance of natural images. Since for natural images $R_{II}(\mathbf{u})$ varies proportionally to $\mathbf{u}^{-2}$, the product $R_{II}(\mathbf{u})|\mathbf{u}|^2$ whitens R_{II} by producing a power spectrum $R_{\tilde{I}\tilde{I}}$ that remains virtually constant at all frequencies.

Use of this new expression for $R_{\tilde{I}\tilde{I}}$ in Equation 11 gives a dynamic power spectrum of geniculate activity:

$$R^D_{\alpha\alpha} = |G|^2 R_{\xi\xi}(w)(|F_c|^2 + |F_s|^2) R_{II}(\mathbf{u})|\mathbf{u}|^2 + |G|^2 \cos(w t_D) R_{\xi\xi}(w) 2 F_c F_s R_{II}(\mathbf{u})|\mathbf{u}|^2$$

and, correspondingly, a dynamic correlation map:

$$\begin{aligned}\hat{c}^D_{\alpha\alpha}(\mathbf{x}) = {} & \mathcal{F}^{-1}\{|G(w)|^2 R_{\xi\xi}(w)\}\big|_{t=0} \mathcal{F}^{-1}\{(|F_c|^2 + |F_s|^2) R_{II}(\mathbf{u})|\mathbf{u}|^2\} \\ & + \mathcal{F}^{-1}\{|G(w)|^2 R_{\xi\xi}(w)\cos(w t_D)\}\big|_{t=0} \mathcal{F}^{-1}\{2 F_c F_s R_{II}(\mathbf{u})|\mathbf{u}|^2\}\end{aligned} \tag{15}$$

To summarize, Equation 12 shows that fixational instability adds the dynamic correlation $\hat{c}^D_{\alpha\alpha}$ to the pattern of correlated activity $\hat{c}^S_{\alpha\alpha}$ obtained with presentation of the stimulus under conditions of retinal stabilization. Equation 15 shows that this contribution depends on $R_{\tilde{I}\tilde{I}}$ (and not R_{II}), a signal that discards the broad correlation of natural images.

To examine the influence of the two terms $\hat{c}^D_{\alpha\alpha}$ and $\hat{c}^S_{\alpha\alpha}$ on the structure of correlated activity, Figure 5 shows their ratio at separation zero, $\rho_{DS} = \hat{c}^D_{\alpha\alpha}(0)/\hat{c}^S_{\alpha\alpha}(0)$, with presentation of natural images and for various parameters of fixational instability. Fixational instability was assumed to possess Gaussian temporal correlation ($R_{\xi\xi}(w)$ in Equation 15), with standard deviation σ_T and amplitude σ_S. Figure 5a shows the effect of varying the

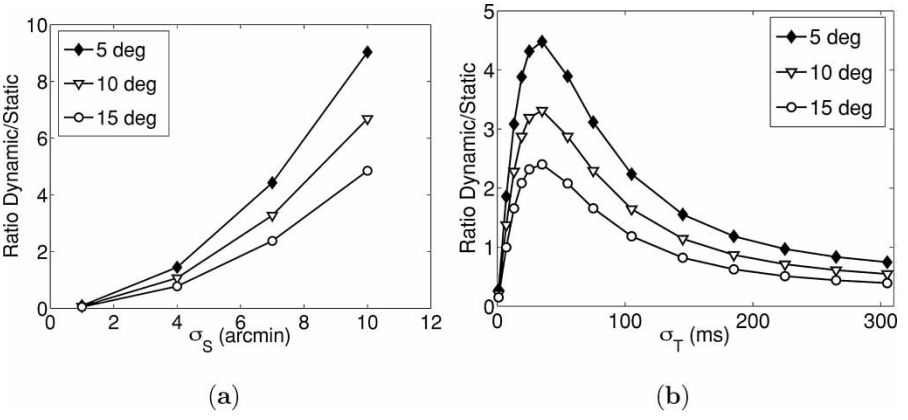

Figure 5. Influence of the characteristics of fixational instability on the patterns of correlated activity during presentation of natural images. The two graphs show the ratio ρ_{DS} between the peaks of the two terms $\hat{c}^D_{\alpha\alpha}$ and $\hat{c}^S_{\alpha\alpha}$ in Equation 12. Fixational instability was assumed to have a Gaussian correlation with standard deviation σ_T and amplitude σ_S. (*a*) Effect of varying σ_S ($\sigma_T = 35$ ms). (*b*) Effect of varying σ_T ($\sigma_S = 7$ arcmin). The three curves in each graph represent results obtained at different angles of visual eccentricity.

amplitude of fixational instability. To remain within the range of validity of the Taylor approximation in Equation 8, only small amplitude values were considered. As shown by Figure 5a, the larger the instability of visual fixation, the larger the contribution of the dynamic term $\hat{c}^D_{\alpha\alpha}$ with respect to $\hat{c}^S_{\alpha\alpha}$. Except for very small values of σ_S, ρ_{DS} was always larger than one, indicating that $\hat{c}^D_{\alpha\alpha}$ influenced the structure of correlated activity more strongly than $\hat{c}^S_{\alpha\alpha}$. Figure 5b shows the impact of varying σ_T, which defined the temporal window over which the fixational jitter was correlated. Note that ρ_{DS} was a non-monotonic function of σ_T. For a range of σ_T corresponding to intervals shorter than the typical duration of visual fixation, $\hat{c}^D_{\alpha\alpha}$ was significantly larger than $\hat{c}^S_{\alpha\alpha}$. Thus, fixational instability strongly influenced correlated activity in the model when it moved the direction of gaze within a range a few arcmin and was correlated over a fraction of the period of visual fixation. This range of parameters is consistent with the instability of fixation observed in many species.

Figure 6 shows the spatial extent of correlation measured in the model when images of natural scenes were examined in the presence of fixational instability. In this example, the parameters of fixational instability were $\sigma_S = 10$ arcmin and $\sigma_T = 30$ ms. In addition to the width of pools of coactive units resulting from Equation 12, Figure 6 also shows the correlation extents produced by the two components $\hat{c}^S_{\alpha\alpha}$ and $\hat{c}^D_{\alpha\alpha}$. Whereas $\hat{c}^S_{\alpha\alpha}$ produced patterns of correlation identical to those observed in the absence of fixational instability, $\hat{c}^D_{\alpha\alpha}$, due to its dependence on the whitened power spectrum $R_{\tilde{I}\tilde{I}}$, was similar to the correlation map measured with uncorrelated input. In this map, units were simultaneously active only when their receptive fields largely overlapped. Responses were, instead, anti-correlated when the ON subregion of one cell overlapped the OFF subregion of the other unit. Since, as shown in Figure 5, $\hat{c}^D_{\alpha\alpha}$ provided a larger contribution than $\hat{c}^S_{\alpha\alpha}$, the global structure of correlated activity was strongly influenced by the dynamic term. Thus, the spatial extent of correlation was similar to that measured in the presence of white noise.

Figure 7 analyzes the origins of the decorrelation of neural activity produced by fixational instability. Two main elements contributed to this effect: (a) the presence of a spatially uncorrelated component in the retinal input during the physiological instability of visual

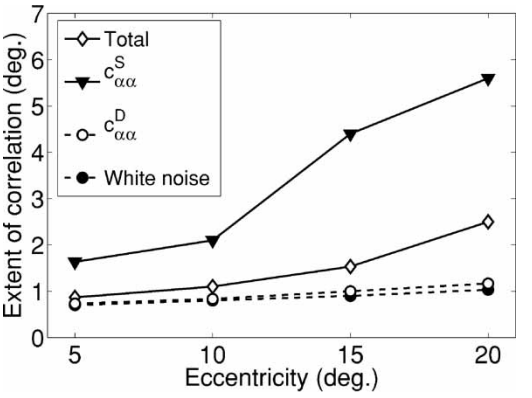

Figure 6. Spatial extent of correlation obtained from Equation 12 when natural images were examined in the presence of fixational instability. The curve labeled as "Total" represents the correlation extent given by $\hat{c}_{\alpha\alpha}$ in Equation 12. In addition to the total correlation, also the width of pools of units with correlated activity produced by the static and dynamic terms $\hat{c}^S_{\alpha\alpha}$ and $\hat{c}^D_{\alpha\alpha}$ are shown. The spatial extent measured with presentation of white noise is plotted here for comparison. Results are shown for parameters of the model that replicated the receptive fields of geniculate cells located at various angles of visual eccentricity.

fixation ($R_{\tilde{I}\tilde{I}}$ in Equation 10), and (b) the strong sensitivity of geniculate units to this input component (Figure 5). It is important to note that the first of these two elements derived from the way fixational instability interacted with the statistics of natural images. It is the particular spectral density function of natural images that ensured that R_{II} had a flat spectrum in Equation 14. Figure 7a shows the effect of altering the power spectrum of natural images by low-pass filtering. As expected, the spatial extent of correlation increased when input images contained only low spatial frequencies. This increment is due to the attenuation of harmonics in the medium and high frequency bands that provide uncorrelated input to units with receptive fields in different locations.

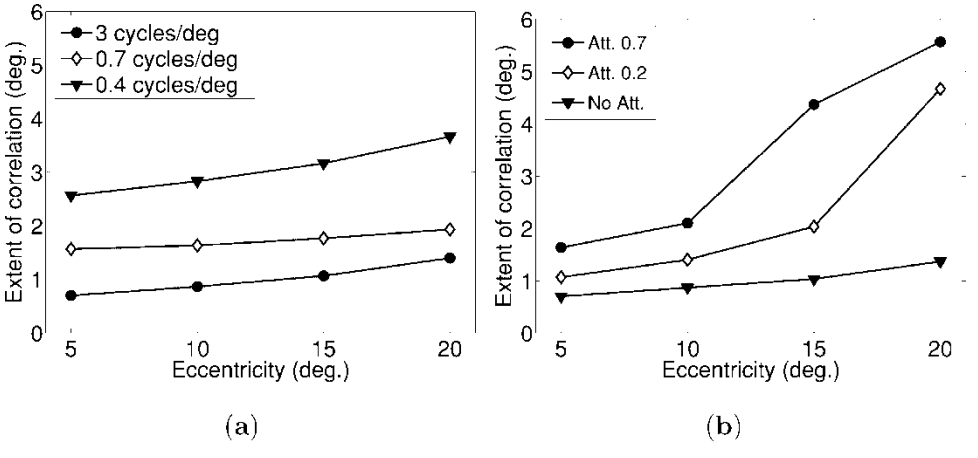

Figure 7. Influence of the spatial frequencies of the stimulus and the temporal frequencies of fixational instability on neural decorrelation. Data points indicate the spatial extent of correlation measured with presentation of natural images. (*a)* Effect of low-pass filtering the input images. The curves show results obtained with low-pass spatial filters with different cut-off frequency. (*b*) Effect of band-pass filtering fixational instability. The curves show results obtained with band-pass temporal filters that eliminated different frequency ranges around the unit's preferred frequency.

Since in natural images most power is concentrated at low spatial frequencies, the uncorrelated fluctuations in the input signals generated by fixational instability possess small amplitudes. In the model, these small input modulations were amplified by the response characteristics of geniculate cells. As expressed by Equation 15, levels of correlation were multiplied by a gain factor, which weighted the matching between the temporal frequencies of fixational instability $R_{\xi\xi}(w)$ and the temporal sensitivity of geniculate cells $G(w)$. Figure 7b shows the effect of band-pass filtering fixational instability in order to eliminate the temporal frequencies at which geniculate units are most responsive. The three curves in Figure 7 represent the results obtained with filters that suppressed bands with different widths centered around the optimal frequency of the modeled units. For each filter, the size of the suppressed band is indicated by the corresponding attenuation index. This index determines the critical amplitude in the normalized frequency response above which temporal frequencies were eliminated. For example, an attenuation index equal to a indicates elimination of all frequencies that evoked responses with amplitudes greater than a fraction $1-a$ of the peak amplitude (the amplitude measured with the optimal frequency). It is clear from Figure 7 that elimination of the temporal frequencies to which modeled geniculate units were most sensitive greatly affected the structure of correlated activity.

These results show that there are at least two distinct ways in which the visual system can use fixational instability to precisely control the level of decorrelation in neural responses. The first method is by acting on the spatial scale of instability as shown in Figure 5a. Within the limits of validity of the Taylor approximation in Equation 7, the larger the area of instability the higher the contribution of the dynamic term. The second possibility is by altering the temporal characteristics of fixational instability. This can be achieved by properly adjusting the frequency band of the motion of the retinal image relative to the sensitivity of geniculate units or, equivalently, by changing the temporal window over which fixational instability is correlated.

Discussion

In this paper, we used neural modeling to study the second-order statistics of LGN activity when images of natural scenes were examined in the presence and absence of the physiological instability of visual fixation. In the model, when visual input was presented in the absence of retinal image motion, a condition that simulated the stabilization of the stimulus on the retina, the broad correlations of natural images activated wide ensembles of geniculate units. On the contrary, during the instability of visual fixation, input spatial correlations had little influence on neural responses, and the structure of correlated activity was similar to that observed with presentation of spatially uncorrelated input. These results support a role for fixational instability in decorrelating LGN activity during natural viewing.

It has been argued that a decorrelation of neural responses in the early stages of the visual system is beneficial for the encoding of visual information (Atick 1992; Atick & Redlich 1992). According to the redundancy reduction hypothesis, statistically independent neural responses enable efficient representations with relatively small neuronal populations (Barlow 1961). While at the level of the cortex, the numbers of neurons and their interconnections are so extraordinarily large that redundant representations can be afforded (Barlow 2002), a compact neural code appears advantageous in the neural pathway that transmits information from the retina to the cortex. Neurophysiological studies have shown that the spatial characteristics of neurons in the retina and LGN *per se* are not compatible with a decorrelation of neural responses. In both the cat and the monkey, the frequency responses of cells in these areas deviate significantly from linearity in the low spatial frequency range (So

& Shapley 1981; Linsenmeier et al. 1982; Derrington & Lennie 1984; Croner & Kaplan 1995; Cheng et al. 1995). This deviation was responsible for the broad spatial correlations observed in our model when natural images were presented in the absence of retinal motion. During natural viewing, however, cell responses depend on both their spatial and temporal characteristics. The results of this study show that when the dynamics of neural responses are taken into account, it is not necessary for the contrast sensitivity functions of individual neurons to counterbalance the input power spectrum in order to decorrelate neural activity. In the model, fixational instability decorrelated cell responses even if their contrast sensitivity functions deviated significantly from a linear proportionality with spatial frequency.

Two main elements contributed to decorrelating the responses of the model during fixational instability. The first element was the presence of a dynamic component $R_{\tilde{I}\tilde{I}}$ in the power spectrum of the retinal input, which lacked spatial correlations with presentation of natural images. This component appears to be tuned to the statistics of natural images. It is spatially uncorrelated only in the presence of visual input with a power spectrum that declines as u^{-2} with spatial frequency. In other words, it is a feature of natural images that, although the intensity values of nearby pixels are correlated, changes in intensity around each pixel are uncorrelated. This property is not satisfied by an arbitrary image. In a spatial grating, for example, the intensity changes at any two locations are highly correlated. During visual fixation, neurons receive input from the small regions of the visual field covered by the jittering of their receptive fields. In the presence of natural images, although the inputs to cells with nearby receptive fields are on average correlated, the fluctuations in these input signals produced by fixational instability are not correlated. These spatially uncorrelated fluctuations are characteristic of natural images and do not occur with other types of correlated visual stimulation.

It is important to realize that the presence of a decorrelated component in the visual input is a direct consequence of the spectral density function of natural images and does not originate from some level of high-frequency noise affecting the images. $R_{\tilde{I}\tilde{I}}$ in Equation 10 corresponds to the power spectrum of the first-order term in the Taylor expansion of Equation 8. This term is the main determinant of how the input stimulus changes around a given location. As shown by Figure 7a, a substantial degree of decorrelation was still present even when natural images were low-pass filtered with cut-off frequency as low as 0.7 cycles/deg.

The second important element that contributed to the decorrelation of neural activity was the dynamic interaction between fixational instability and the temporal sensitivity of geniculate units. In the model, units were highly sensitive to uncorrelated fluctuations in the input signals produced by fixational instability. This elevated level of responding originated from the overlap between the spectrum of temporal sensitivity of modeled units and the temporal frequencies provided by fixational instability. In both the cat and the monkey, geniculate cells tend to respond maximally to input harmonics within a range of 1–10 Hz (Usrey & Reid 2000; Saul & Humphrey 1990). The optimal frequency for X cells in the cat is around 2–4 Hz (Saul & Humphrey 1990). A frequency characterization of fixational instability under natural viewing conditions has not been performed. Nonetheless, low temporal frequency harmonics should be expected to possess high power. In humans, the power spectrum of isolated periods of ocular drift and tremor declines as $1/f^2$ in the frequency range 0–40 Hz (Eizenman et al. 1985). This spectrum is presumably broadened by fixational saccades and small movements of the head and body that occur during natural viewing. The instability of fixation of the cat is similar to that of humans. Drift velocities tend to be twice as high as they are in humans, and small saccades, although not as small as those of humans, occur frequently (Winterson & Robinson 1975). Therefore, under natural viewing conditions,

fixational instability should provide temporal frequencies to which cat's LGN neurons are highly sensitive.

According to the theory described in this study, the self-motion of the retinal image attenuates cell sensitivity to the low spatial frequencies of the stimulus and enhances neural responses to high spatial frequencies. As a consequence of such modulations, the prediction emerges that fixational instability might contribute to the neural encoding of local features in a scene. These features typically depend on high spatial frequency harmonics. Interestingly, this prediction is in striking contrast to the prediction of the "neural refreshing" hypothesis, which maintains a role for fixational instability in enhancing (and not discarding) low spatial frequency harmonics. These frequencies appear to fade rapidly under conditions of retinal stabilization. The two hypotheses could be confronted by analyzing visual performances under normal viewing and retinal stabilization conditions with stimuli in which high spatial frequency targets are partly masked by low spatial frequency noise. Contrary to the "neural refreshing" hypothesis, this study predicts that low-frequency noise will have a reduced impact during natural viewing, when the physiological self-motion of the retinal image enhances cell responses to the targets.

It should be observed that the mathematical analysis of this paper implicitly assumed a steady-state condition of visual fixation. This assumption is intrinsic to the procedure followed for evaluating the structure of correlated activity on the basis of the power spectrum of visual stimulation. Under natural viewing conditions, saccades separate periods of fixation every few hundred milliseconds. Given that for the majority of geniculate neurons most of the impulse response dynamics occurs within 150 ms, the analysis of this paper strictly applies to a later period of visual fixation, the period following the initial 150 ms transitory interval when neural responses have adjusted to the stimulus brought in by the saccade. Neurons respond strongly to the onset of new stimuli. Thus, the power spectrum of visual stimulation R_{II} may be expected to dominate the structure of correlated activity during the transitory period immediately following a saccade. It is a prediction of this study that initially correlated geniculate responses might become progressively less correlated during the period of visual fixation. These considerations have important repercussions for the way visual information is then processed by later stages of the visual system, as processing needs to be synchronized with the onset of visual fixation.

As in any modeling study, the results of this paper depend on the degree of accuracy with which biological signals were replicated. The model of geniculate neuron used in this paper was originally derived on the basis of reverse correlation experiments, and it has been shown to accurately reproduce cell responses to visual stimulation (Cai et al. 1997). All parameters of the model followed neurophysiological data. Nonetheless, a number of factors were not present in the model, including sources of correlated activity in the LGN other than shared visual stimulation (Alonso et al. 1996) and extraretinal influences (Reppas et al. 2002). Considerations of this sort emphasize the need to validate the predictions of our study with *in vivo* experiments. In any case, it is important to observe that the first element of the decorrelation observed in this study, the spatially uncorrelated retinal input component present with natural images, is independent of the model. Furthermore, the results of the mathematical analysis of this paper are consistent with those of previous simulations investigating the effect of chronic exposure to fixational instability on visual development (Rucci et al. 2000; Rucci & Casile 2004; Parsons & Rucci 2005). These previous simulations have produced results consistent with physiological data irrespective of a number of factors including the fine characteristics of simulated eye movements, the degree of modulation of geniculate activity due to saccades, and the degree of rectification in the responses of geniculate cells.

Studies that investigated vision under conditions of retinal stabilization, in which fixational instability is eliminated, have shown that images tend to fade over a period of several seconds or minutes in the absence of retinal motion (Ditchburn & Ginsborg 1952; Riggs & Ratliff 1952; Steinman & Levinson 1990). Image fading has often been held as a justification for the existence of retinal jittering. According to this view, the physiological motion of the retinal image is necessary to refresh neuronal responses and prevent the disappearance of a stationary scene. This claim, however, remains controversial given the typically brief durations of natural visual fixation (Arend & Timberlake 1986; Ditchburn 1987; Steinman & Levinson 1990; Coppola & Purves 1996). Other theories have proposed a more elaborate influence of fixational instability on neural activity than that of simply refreshing neural responses. It was originally proposed by Marshall and Talbot (1942) that the tremor of the eye might contribute to hyperacuity, a proposal that was reexamined by a recent computational study (Hennig & Wörgötter 2004). More recently, Ahissar and Arieli (2001) have argued that fixational instability is a critical element in the neural encoding of visual information. Although these proposals differ in suggested mechanisms, they are close to the theory described in this paper in arguing for an important role of fixational instability in structuring neural activity. The novel contributions of the proposed theory are the specific mechanisms of neural coding, which have been cast within the framework of the redundancy reduction hypothesis and the establishment of a link between fixational instability and the statistics of natural images.

Studies investigating the influence of the structure of the environment on visual development and performance often neglect the ubiquitous presence of behavior during the acquisition of information. During natural viewing, the spatiotemporal structure of the signals entering the eyes depends not only on the visual scene, but also on the movement performed by the observer. Eye movements are always present during natural vision. Consistent with the results of recent neurophysiological investigations (Leopold & Logothetis 1998; Snodderly et al. 2001; Greschner et al. 2002; Martinez-Conde et al. 2002; Olveczky et al. 2003), this study suggests that even a minute motor activity like fixational instability may have a profound influence on the statistics of neural responses. Further studies are needed to analyze the impact of other dimensions of the visual stimulus, such as motion and color, on neural activity in the presence of a constantly moving retina.

Acknowledgements

We thank Gaelle Desbordes, Eric Schwartz, and Robert Steinman for helpful comments and discussions. This material is based upon work supported by the National Institute of Health under Grant EY15732-01 and the National Science Foundation under Grants CCF-0432104 and CCF-0130851.

References

Ahissar E, Arieli A. 2001. Figuring space by time. Neuron 32:185–201.

Alonso JM, Usrey WM, Reid RC. 1996. Precisely correlated firing in cells of the lateral geniculate nucleus. Nature 383:815–819.

Arend LE, Timberlake GT. 1986. What is psychophysically perfect image stabilization? Do perfectly stabilized images always disappear? J Opt Soc Am 3:235–241.

Atick JJ. 1992. Could information theory provide an ecological theory of sensory processing? Network 3:213–251.

Atick JJ, Redlich A. 1992. What does the retina know about natural scenes? Neural Comp 4:449–572.

Attneave F. 1954. Informational aspects of visual perception. Psychol Rev 61:183–193.

Barlow HB. 1961. The coding of sensory messages. In: Thorpe WH and Zangwill OL, editors, Current Problems in Animal Behaviour. Cambridge: Cambridge University Press, pp. 331–360.

Barlow HB. 2002. Redundancy reduction revisited. Network Comput Neural Syst 12:241–253.
Burton GJ, Moorhead IR. 1987. Color and spatial structure in natural scenes. Applied Optics 26:157–170.
Cai D, DeAngelis GC, Freeman RD. 1997. Spatiotemporal receptive field organization in the lateral geniculate nucleus of cats and kittens. J Neurophysiol 78:1045–1061.
Casile A, Rucci M. 2005. A theoretical analysis of the influence of fixational instability on the development of thalamo-cortical connectivity. Submitted manuscript and CNS Technical Report CAS/CNS-TR-2005-001.
Cheng H, Chino YM, Smith EL, Hamamoto J, Yoshida K. 1995. Transfer characteristics of lateral geniculate nucleus X neurons in the cat: effects of spatial frequency and contrast. J Neurophysiol 74:2548–2557.
Coppola D, Purves D. 1996. The extraordinary rapid disappearance of entopic images. Proc Natl Acad USA 23:8001–8004.
Croner LJ, Kaplan E. 1995. Receptive fields of P and M ganglion cells across the primate retina. Vision Res 35:7–24.
Derrington AM, Lennie P. 1984. Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. J Physiol 357:219–240.
Ditchburn RW. 1955. Eye movements in relation to retinal action. Optica Acta 1:171–176.
Ditchburn RW. 1987. Comment to: What is psychophysically perfect image stabilization? Do perfectly stabilized images always disappear? J Opt Soc Am 4:405–406.
Ditchburn RW Ginsborg B. 1952. Vision with a stabilized retinal image. Nature 170:36–37.
Eizenman M, Hallett P, Frecker RC. 1985. Power spectra for ocular drift and tremor. Vision Res 25:1635–1640.
Field D. 1987. Relations between the statistics of natural images and the response properties of cortical cells. J Opt Soc America A 4:2379–2394.
Field D. 1994. What is the goal of sensory coding? Neural Comp 6:559–601.
Greschner M, Bongard M, Rujan P, Ammermüller J. 2002. Retina ganglion cell synchronization by fixational eye movements improves feature estimation. Nature 5:341–347.
Gur M, Beylin A, Snodderly DM. 1997. Response variability of neurons in primary visual cortex (V1) of alert monkeys. J Neurosci 17:2914–2920.
Hennig MH, Wörgötter F. 2004. Eye micro-movements improve stimulus detection beyond the Nyquist limit in the peripheral retina. In: Thrun S, Saul L, Schölkopf B, editors. Advances in Neural Information Processing Systems 16. Cambridge, MA: MIT Press.
Leopold DA, Logothetis NK. 1998. Microsaccades differentially modulate neural activity in the striate and extrastriate visual cortex. Exp Brain Res 123:341–345.
Linsenmeier RA, Frishman LJ, Jakiela HG, Enroth-Cugell C. 1982. Receptive field properties of X and Y cells in the cat retina derived from contrast sensitivity measurements. Vision Res 22:1173–1183.
Marshall W, Talbot S. 1942. Recent evidence for neural mechanisms in vision leading to a general theory of sensory acuity. Biological Symposia Visual Mechanisms 7:117–164.
Martinez-Conde S, Macknik SL, Hubel DH. 2000. Microsaccadic eye movements and firing of single cells in the striate cortex of macaque monkeys. Nat Neurosci 3:251–258.
Martinez-Conde S, Macknik SL, Hubel DH. 2002. The function of bursts of spikes during visual fixation in the awake primate lateral geniculate nucleus and primary visual cortex. Proc Natl Acad Sci USA 99:13920–13925.
Martinez-Conde S, Macknik SL, Hubel DH. 2004. The role of fixational eye movements in visual perception. Nature Rev Neurosci 5:229–240.
Murakami I, Cavanagh P. 1998. A jitter after-effect reveals motion-based stabilization of vision. Nature 395:798–801.
Olveczky B, Baccus S, Meister M. 2003. Segregation of object and background motion in the retina. Nature 423:401–408.
Parsons O, Rucci M. 2005. Modeling the possible influences of eye movements on the refinement of cortical direction selectivity. Submitted manuscript and CNS Technical Report CAS/CNS-TR-04-013.
Ratliff F, Riggs LA. 1950. Involuntary motions of the eye during monocular fixation. J Exp Psychol 40:687–701.
Reppas JB, Usrey WM, Reid RC. 2002. Saccadic eye movements modulate visual responses in the lateral geniculate nucleus. Neuron 35:961–974.
Riggs LA, Ratliff F. 1952. The effects of couteracting the normal movements of the eye. J Opt Soc Am 42:872–873.
Rucci M, Casile A. 2004. Decorrelation of neural activity during fixational instability: possible implications for the refinement of V1 receptive fields. Visual Neurosci 21:725–738.
Rucci M, Desbordes G. 2003. Contributions of fixational eye movements to the discrimination of briefly presented stimuli. J Vision 3:852–64.
Rucci M, Edelman GM, Wray J. 2000. Modeling LGN responses during free-viewing: a possible role of microscopic eye movements in the refinement of cortical orientation selectivity. J Neurosci 20:4708–4720.
Ruderman DL, Bialek W. 1994. Statistics of natural images: Scaling in the woods. Physical Rev Lett 73:814–817.

Saul AB, Humphrey AL. 1990. Spatial and temporal response properties of lagged and nonlagged cells in cat lateral geniculate nucleus. J Neurophysiol 64:206–224.

Skavenski AA, Hansen RM, Steinman RM, Winterson BJ. 1979. Quality of retinal image stabilization during small natural and artificial body rotations in man. Vision Res 19:675–683.

Snodderly DM, Kagan I, Gur M. 2001. Selective activation of visual cortex neurons by fixational eye movements: implications for neural coding. Vis Neurosci 18:259–277.

So YT, Shapley R. 1981. Spatial tuning of cells in and around lateral geniculate nucleus of the cat: X and Y relay cells and perigeniculate interneurons. J Neurophysiol 45:107–120.

Steinman R, Levinson J. 1990 The role of eye movements in the detection of contrast and spatial detail. In: Kowler E, editior. Eye Movements and their Role in Visual and Cognitive Processes. Oxford: Elsevier Science, pp 115–212.

Steinman RM, Haddad GM, Skavenski AA, Wyman D. 1973. Miniature eye movement. Science 181:810–819.

Usrey WM, Reid RC. 2000. Visual physiology of the lateral geniculate nucleus in two species of new world monkey: Saimiri sciureus and Aotus trivirgatis. J Physiol 523:755–769.

van Hateren JH, van der Schaaf A. 1998. Independent component filters of natural images compared with simple cells in primary visual cortex. Proc R Soc Lond B 265:359–366.

Winterson BJ, Robinson DA. 1975. Fixation by the alert but solitary cat. Vision Res 15:1349–1352.