

# Supporting Information

Segal et al. 10.1073/pnas.1412059112

## 1. General Definitions

The objective of this study is to reveal the effects of eye motion during fixational instability on the structure of the correlations of the information streaming from the retina to the brain. The information to be processed by the retina comes in the form of the intensity distribution over the retina surface (2D space), which we denote  $I(x, t)$ . Here  $x$  is the position of the photoreceptor on the retina and  $t$  is time. In reality the photoreceptors have finite size and their positions are described by a discrete set of  $x_i$ . However, in theoretical analysis we shall use the continuum approximation (i.e.,  $x$  is continuous). We describe the information that the retina transfers upstream by the continuous response  $O(x, t)$ , where now  $x$  is the position of the retinal ganglion cell. Again, the positions of the centers of the receptive fields are in reality a discrete set  $x_i$ , but for analytical treatment we shall use the continuous description. The response of the cell can be summarized by the firing rate as function of time, which is also treated as continuous in this analysis.

As a characteristic of the synchronization of the cell responses we use the adopted lifetime correlation coefficient. Let us assume that we observe the retinal response during a very long time  $T$  (lifetime). In other words, we have both input  $I(x, t)$  and output  $O(x, t)$  for  $0 \leq t \leq T$  (for simplicity we use one-dimensional notation). We define the lifetime mean  $M(x)$  and SD  $S(x)$  in the following way (1):

$$M(x) = \frac{1}{T} \int_0^T O(x, t) dt \quad [\text{S1}]$$

$$S^2(x) = \frac{1}{T} \int_0^T [O(x, t) - M]^2 dt. \quad [\text{S2}]$$

Respectively, we define the lifetime correlation, covariance, and correlation coefficient as follows:

$$C(x, y) = \frac{1}{T} \int_0^T O(x, t) O(y, t) dt \quad [\text{S3}]$$

$$\begin{aligned} \text{Cov}(x, y) &= \frac{1}{T} \int_0^T [O(x, t) - M(x)][O(y, t) - M(y)] dt \\ &= C(x, y) - M(x)M(y) \end{aligned} \quad [\text{S4}]$$

$$CC(x, y) = \frac{\text{Cov}(x, y)}{S(x)S(y)}. \quad [\text{S5}]$$

Let the whole time of observation be split in  $n$  trials of the duration  $\Delta t$  each. We shall define the trial mean, correlation, and covariance, respectively, in a similar way:

$$M_i(x) = \frac{1}{\Delta t} \int_{t_i}^{t_i+\Delta t} O(x, t) dt \quad [\text{S6}]$$

$$C_i(x, y) = \frac{1}{\Delta t} \int_{t_i}^{t_i+\Delta t} O(x, t) O(y, t) dt \quad [\text{S7}]$$

$$\begin{aligned} \text{cov}_i(x, y) &= \frac{1}{\Delta t} \int_{t_i}^{t_i+\Delta t} [O(x, t) - M_i(x)][O(y, t) - M_i(y)] dt \\ &= C_i(x, y) - M_i(x)M_i(y), \end{aligned} \quad [\text{S8}]$$

where  $i$  denotes the number of the trial starting at  $t_i$  and ending at  $t_i + \Delta t$ .

Now let the trial correlation coefficient be

$$CC_i(x, y) = \frac{\text{COV}_i(x, y)}{S(x)S(y)} \quad [\text{S9}]$$

$$\text{COV}_i(x, y) = \frac{1}{\Delta t} \int_{t_i}^{t_i+\Delta t} [O(x, t) - M(x)][O(y, t) - M(y)] dt \quad [\text{S10}]$$

$$\text{COV}_i(x, y) = \text{cov}_i(x, y) + \delta M_i(x)\delta M_i(y), \quad [\text{S11}]$$

where  $\delta M_i(x) = M_i(x) - M(x)$  is the difference between the trial mean and the lifetime mean. Note the difference between  $\text{COV}_i$  and  $\text{cov}_i$ . The latter is the true trial covariance, whereas the former mixes the trial response with the lifetime mean.

From the last definition it is clear that knowledge of the lifetime behavior is needed to determine the trial correlation coefficient  $CC_i$  defined by Eq. S9. The mean trial correlation coefficient is then

$$\langle CC_i(x, y) \rangle = \frac{1}{T} \sum_i CC_i(x, y) \Delta t = \frac{C(x, y) - M(x)M(y)}{S(x)S(y)} = CC(x, y), \quad [\text{S12}]$$

where  $CC(x, y)$  is defined in Eq. S5. Here we used the identities

$$\frac{1}{T} \sum_i \int_{t_i}^{t_i+\Delta t} (\dots) dt = \frac{1}{T} \int_0^T (\dots) dt \quad [\text{S13}]$$

$$M(x) = \frac{1}{T} \sum_i M_i(x) \Delta t. \quad [\text{S14}]$$

Thus, calculation of the trial correlation coefficients  $CC_i$  and averaging it later is equivalent to calculation of the correlation coefficient over the lifetime observation, without splitting in trials. It is clear that  $\langle CC_i(x, x) \rangle = CC(x, x) = 1$  by construction.

The quantity of ultimate interest is the dependence of the correlation coefficient on the distance between the cells:

$$CC(r) \equiv \frac{1}{2\pi L_x L_y} \int d^2 x d\varphi CC(x, x+r), \quad [\text{S15}]$$

where the spatial size of the retina,  $L_x \times L_y$ , is assumed sufficiently large to exclude the influence of the edges. Here we restored the 2D notation to remind the additional averaging over directions ( $\varphi$ ).

One important point is that during the lifetime observation a cell is exposed to a large statistical ensemble of stimuli. If the cells are identical and there is no systematic bias of the stimuli, the lifetime mean and SD should be the same for all cells, that is,  $M(\mathbf{x}) = M$  and  $S(\mathbf{x}) = S$ . Thus, statistically,  $CC(r) = \tilde{C}(r)/\tilde{C}(0)$ , where

$$\tilde{C}(r) = \langle \langle [O(\mathbf{x}, t) - M][O(\mathbf{x} + \mathbf{r}, t)] \rangle_{\mathbf{x}} \rangle_{\varphi}, \quad [\text{S16}]$$

where  $M = \langle O(\mathbf{x}, t) \rangle_t$  and  $\langle (\dots) \rangle$  stands for averaging, for brevity, and the corresponding index denotes over which variable the averaging is done (time, position, or direction).

For the below theoretical treatment it is assumed that the response can be represented as a functional of the input. We shall also neglect any noise in the activity of the retinal ganglion cells and assume that we are within the linear regime. That is, we assume that the time averaged input is zero, which accounts for the adaptation of the retina to the static illumination, and that the relation between the input and the response can be written as follows:

$$O(\mathbf{x}, t) = \int dt' dx' RF(\mathbf{x}, \mathbf{x}', t, t') I(\mathbf{x}', t'). \quad [\text{S17}]$$

Here the receptive field  $RF(\mathbf{x}, \mathbf{x}', t, t')$  takes into account that the response of a cell depends on the input accumulated over some area and on the temporal development of the intensity. In general, there are no restrictions on  $RF$  except that  $RF = 0$  for  $t < t'$ , which states that the retina is a causal system. If the retinal properties do not depend on time (e.g., no retinal adaptation) then  $RF(t, t') = RF(t - t')$ , which seems reasonable because in our analysis we measure the retinal response in relatively long experiments. Furthermore, if all cells were identical and the retina were infinite, one would have  $RF = RF(t - t', \mathbf{x} - \mathbf{x}')$ . This assumption is of course not correct because the retina is finite and the effect of the edges can destroy the translational invariance of  $RF$ . However, the edge effects can be reduced and probably neglected by considering a sufficiently large retina part where the densities of photoreceptors and ganglion cells are sufficiently high.

Additional discussion should address the issue that cells in the retina are not identical. Thus, proper analysis to be compared with the experiment should take into account the variability of the cellular responses. That is, in general,  $RF = RF(t - t', \mathbf{x} - \mathbf{x}', \mathbf{x})$ . However, as a first-step approximation, here we ignore the differences between cells. Furthermore, throughout the analysis we take an even more simplifying approximation of using the separable kernel where  $RF = T(t - t')X(\mathbf{x} - \mathbf{x}')$ , with  $T(t - t') = 0$  for  $t < t'$ . Although this approximation is useful for our analysis, one should keep in mind that it does not take into account the center-surround feature of retinal ganglion cells. For the latter one can try  $RF = \sum_i T_i(t - t')X_i(\mathbf{x} - \mathbf{x}')$  as a simple model.

With all above, we assume that the intensity field is transformed to the field of the firing rates as follows:

$$O(\mathbf{x}, t) = \int dt' dx' T(t - t')X(\mathbf{x} - \mathbf{x}')I(\mathbf{x}', t'), \quad [\text{S18}]$$

where the background is subtracted from both  $I$  and  $O$ . It is worth noting that if the cells differ by only the amplitude of their response and are identical by the functional form, that is, if

$$O(\mathbf{x}, t) = A(\mathbf{x}) \int dt' dx' T(t - t')X(\mathbf{x} - \mathbf{x}')I(\mathbf{x}', t'), \quad [\text{S19}]$$

the multiplicative factors  $A(\mathbf{x})$  eventually drop in all above calculations and the resulting correlations have the same form as in the case where  $A$  does not depend on  $\mathbf{x}$ .

We are interested in the comparison of the responses for two type of stimuli. First is the scenario when the retinal output is directly derived by the input correlation. This is obtained by flashing and holding the image over the photoreceptors layer, which simulates a stabilized fixation (SF) condition. Second is the scenario where the retinal output is determined by the interplay between fixational eye movements and image statistics (denoted FEM) over the same statistical ensemble of images (discussed below). The intensity distribution for both is given by

$$I_{SF}(\mathbf{x}, t) = I_0(\mathbf{x})H(t) \quad [\text{S20}]$$

$$I_{FEM}(\mathbf{x}, t) = I_0(\mathbf{x} + \mathbf{s}(t)), \quad [\text{S21}]$$

where  $I_0(\mathbf{x})$  is the intensity distribution of the static image,  $H(t)$  is the step function, and  $\mathbf{s}(t)$  represents the translational motion of the image on the retina owing to the eye movements.

## 2. Analysis of the Response to the SF Condition

The response to the SF input  $I(\mathbf{x}, t) = I_0(\mathbf{x})H(t)$  takes the form

$$O(\mathbf{x}, t) = L(t)F(\mathbf{x}) \quad [\text{S22}]$$

$$L(t) = \int_0^t dt' T(t - t') \quad [\text{S23}]$$

$$F(\mathbf{x}) = \int dx' X(\mathbf{x} - \mathbf{x}')I_0(\mathbf{x}'). \quad [\text{S24}]$$

Respectively,

$$\langle O \rangle(\mathbf{x}) = \bar{L}F(\mathbf{x}), \quad \langle L \rangle = \frac{1}{\tau} \int_0^{\tau} L(t) dt \quad [\text{S25}]$$

$$\tilde{O}(\mathbf{x}, t) = (L(t) - \langle L \rangle)F(\mathbf{x}) \equiv \tilde{L}(t)F(\mathbf{x}). \quad [\text{S26}]$$

The covariance therefore takes the following form:

$$\text{Cov}(\mathbf{x}, \mathbf{y}) = GF(\mathbf{x})F(\mathbf{y}) \quad [\text{S27}]$$

$$G = \int_0^{\tau} dt \tilde{O}^2(t) \quad [\text{S28}]$$

$$F(\mathbf{x}) = \int dx' X(\mathbf{x} - \mathbf{x}')I_0(\mathbf{x}'). \quad [\text{S29}]$$

Here and hereafter  $G$  will denote any unimportant constant factor. The covariance for the SF can be easily calculated using the Fourier transform of the spatial part:

$$\text{Cov}(\mathbf{x}, \mathbf{y}) = G \int dk dk' X(\mathbf{k})I_0(\mathbf{k})e^{i\mathbf{k}\mathbf{x}}X(\mathbf{k}')I_0(\mathbf{k}')e^{i\mathbf{k}'\mathbf{y}} \quad [\text{S30}]$$

$$\tilde{C}(r) = \int dx C(\mathbf{x}, \mathbf{x} + \mathbf{r}) = G \int dk |X(\mathbf{k})|^2 |I_0(\mathbf{k})|^2 e^{i\mathbf{k}\mathbf{r}}, \quad [\text{S31}]$$

where for each  $f(\mathbf{x})$  we define

$$f(\mathbf{x}) = \int f(\mathbf{k})e^{i\mathbf{k}\mathbf{x}} d\mathbf{k}, \quad [\text{S32}]$$

retaining, for brevity, the same symbol for the function and its Fourier transform.

Eq. S31 describes the correlation function for a particular retinal response and a particular image spectrum (trial correlation). Statistical averaging is done assuming  $|X(\mathbf{k})|^2 = |X(k)|^2$  (symmetric receptive fields), and average  $|I_0(\mathbf{k})|^2 = |I_0(k)|^2$ . When averaging over the directions

$$\tilde{C}(r) = \int_0^{2\pi} \tilde{C}(r, \phi) d\phi \quad [\text{S33}]$$

$$= G \int_0^\infty k dk \int_0^{2\pi} d\phi |X(k)|^2 |I_0(k)|^2 e^{ikr \cos \phi}, \quad [\text{S34}]$$

where  $\mathbf{k}r = kr \cos \phi$ . Using

$$e^{ikr \cos \phi} = \sum_n J_n(kr) e^{in(\pi/2 - \phi)} \quad [\text{S35}]$$

one gets

$$\tilde{C}_{SF}(r) = G \int_0^\infty k dk |X(k)|^2 |I_0(k)|^2 J_0(kr) \quad [\text{S36}]$$

$$CC_{SF}(r) = \frac{\int_0^\infty k dk |X(k)|^2 |I_0(k)|^2 J_0(kr)}{\int_0^\infty k dk |X(k)|^2 |I_0(k)|^2}. \quad [\text{S37}]$$

Eq. S37 provides the general expression for the cross-correlation for the SF input.

### 3. Analysis of the Response to FEM

We shall consider the stationary response to the input from eye movements,  $I(\mathbf{x}, t) = I_0(\mathbf{x} + \mathbf{s}(t))$ . The stationarity means that  $\langle O(\mathbf{x}, t) \rangle = 0$  for the measurement interval  $0 \leq t \leq \tau$ . In this case  $O = 0$ , whereas  $\langle O(\mathbf{y}, t) O(\mathbf{x}, t) \rangle$  remains finite and independent of time. Respectively, the correlation takes the form

$$C(\mathbf{x}, \mathbf{y}) = \langle O(\mathbf{y}, t) O(\mathbf{x}, t) \rangle = \frac{1}{\tau} \int_0^\tau O(\mathbf{y}, t) O(\mathbf{x}, t) dt. \quad [\text{S38}]$$

One can write

$$I(\mathbf{x}, t) = \int d\mathbf{k} I_0(\mathbf{k}) e^{i\mathbf{k}(\mathbf{x} + \mathbf{s}(t))} \quad [\text{S39}]$$

$$O(\mathbf{x}, t) = \int d\mathbf{k} \int d\mathbf{x}' dt' X(\mathbf{x} - \mathbf{x}') T(t - t') I_0(\mathbf{k}) e^{i\mathbf{k}(\mathbf{x}' + \mathbf{s}(t'))} \quad [\text{S40}]$$

$$O(\mathbf{x}, t) = \int d\mathbf{k} \int dt' T(t - t') X(\mathbf{k}) I_0(\mathbf{k}) e^{i\mathbf{k}(\mathbf{x} + \mathbf{s}(t'))}. \quad [\text{S41}]$$

Respectively, one has

$$C(\mathbf{x}, \mathbf{y}) = \int dt dt' dt'' dx' dx'' dk' dk'' T(t - t') T(t - t'') \cdot X(\mathbf{x} - \mathbf{x}') X(\mathbf{y} - \mathbf{x}'') I_0(\mathbf{k}') I_0(\mathbf{k}'') \cdot e^{i\mathbf{k}'(\mathbf{x}' + \mathbf{s}(t'))} e^{i\mathbf{k}''(\mathbf{x}'' + \mathbf{s}(t''))} \quad [\text{S42}]$$

$$C(\mathbf{x}, \mathbf{y}) = \int d\mathbf{k}' d\mathbf{k}'' M(\mathbf{k}', \mathbf{k}'') e^{i\mathbf{k}'\mathbf{x}} e^{i\mathbf{k}''\mathbf{y}} F(\mathbf{k}') F(\mathbf{k}'') \quad [\text{S43}]$$

$$F(\mathbf{k}) = X(\mathbf{k}) I_0(\mathbf{k}) \quad [\text{S44}]$$

$$M(\mathbf{k}', \mathbf{k}'') = \int dt q(\mathbf{k}', t) q(\mathbf{k}'', t) \quad [\text{S45}]$$

$$q(\mathbf{k}, t) = \int dt' T(t - t') e^{i\mathbf{k}\mathbf{s}(t')} \quad [\text{S46}]$$

$$= \int dt' d\omega T(\omega) e^{i\mathbf{k}\mathbf{s}(t') - i\omega(t - t')}. \quad [\text{S47}]$$

One also has

$$\tilde{C}(r) = \int d\mathbf{k} e^{i\mathbf{k}r} M(\mathbf{k}, -\mathbf{k}) |F(\mathbf{k})|^2 \quad [\text{S48}]$$

$$M(\mathbf{k}, -\mathbf{k}) = \int dt |q(\mathbf{k}, t)|^2 = \int d\omega |q(\mathbf{k}, \omega)|^2 \quad [\text{S49}]$$

$$q(\mathbf{k}, \omega) = T(\omega) \left[ \int dt e^{i\mathbf{k}\mathbf{s}(t) + i\omega t} \right] \quad [\text{S50}]$$

$$|q(\mathbf{k}, \omega)|^2 = |T(\omega)|^2 Q(\mathbf{k}, \omega) \quad [\text{S51}]$$

$$Q(\mathbf{k}, \omega) = \int dt dt' e^{i\mathbf{k}(\mathbf{s}(t) - \mathbf{s}(t')) + i\omega(t - t')}. \quad [\text{S52}]$$

$Q(\mathbf{k}, \omega)$  corresponds to the attenuation factor of ref. 2. Thus, one has

$$\tilde{C}_{FEM}(r) = G \int d\mathbf{k} d\omega e^{i\mathbf{k}r} |T(\omega)|^2 Q(\mathbf{k}, \omega) |X(\mathbf{k})|^2 |I_0(\mathbf{k})|^2. \quad [\text{S53}]$$

In the above expressions  $T(\omega)$  plays the role of a band-pass filter. Indeed, we assume no response to completely static inputs. However, for any physical system the response should drop to zero when  $\omega \rightarrow \infty$ . Averaging over angles and statistical ensemble with the assumption  $Q(\mathbf{k}, \omega) = Q(k, \omega)$  would give

$$\tilde{C}_{FEM}(r) = G \int_0^\infty k^3 dk R(k) |X(k)|^2 |I_0(k)|^2 J_0(kr), \quad [\text{S54}]$$

where

$$R(k) = \int d\omega |T(\omega)|^2 Q(k, \omega). \quad [\text{S55}]$$

Eq. S41 can be greatly simplified if the amplitude of the eye movements is small compared with the receptive field size, that is,  $X(\mathbf{k})$  is significantly nonzero only for  $|\mathbf{k}\mathbf{s}| \ll 1$ . In this case, expanding the exponent and taking into account that the response to the static (time independent) input vanishes, one gets

$$O(\mathbf{x}, t) = \int d\mathbf{k} d\mathbf{x}' X(\mathbf{k}) I_0(\mathbf{k}) S(t) \quad [\text{S56}]$$

$$S(t) = \int dt' T(t - t') s(t'). \quad [\text{S57}]$$

Respectively,

$$\tilde{C}_{FEM}(r) = \sum_{ij} G \int dk d\omega e^{ikr} k_i k_j \bar{S}_{ij} |X(k)|^2 |I_0(k)|^2, \quad [\text{S58}]$$

where  $i, j = 1, 2$  denote the two orthogonal directions at the retina and

$$\bar{S}_{ij} = \int dt S_i(t) S_j(t). \quad [\text{S59}]$$

Averaging over directions and a statistical ensemble is straightforward. Let  $k_x = k \cos \alpha$ ,  $k_y = k \sin \alpha$ ,  $\mathbf{k}r = kr \cos \varphi$ . Averaging over  $\alpha$  and  $\varphi$  is independent. Assuming again isotropy of the receptive fields and average image spectra, one gets

$$\tilde{C}_{FEM}(r) = G \int k^3 dk |X(k)|^2 |I_0(k)|^2 J_0(kr). \quad [\text{S60}]$$

#### 4. Comparison Between the Response Patterns

We are interested in the comparison of  $CC(r) = \tilde{C}(r)/\tilde{C}(0)$  for both cases:

$$CC_{SF}(r) = \frac{\int_0^\infty k dk |X(k)|^2 |I_0(k)|^2 J_0(kr)}{\int_0^\infty k dk |X(k)|^2 |I_0(k)|^2} \quad [\text{S61}]$$

$$CC_{FEM}(r) = \frac{\int_0^\infty k^3 dk R(k) |X(k)|^2 |I_0(k)|^2 J_0(kr)}{\int_0^\infty k^3 dk R(k) |X(k)|^2 |I_0(k)|^2}. \quad [\text{S62}]$$

If the amplitude of the eye movements is substantially smaller than the receptive field size the latter expression simplifies to the following:

$$CC_{FEM}(r) = \frac{\int_0^\infty k^3 dk |X(k)|^2 |I_0(k)|^2 J_0(kr)}{\int_0^\infty k^3 dk |X(k)|^2 |I_0(k)|^2}, \quad [\text{S63}]$$

which is used in the main text for the comparison with the experiment.

For the receptive field shape in the form of an isotropic Gaussian,

$$X(r) = \frac{1}{r_0^2} \exp(-r^2/2r_0^2) \Rightarrow X(k) = \exp(-k^2/2k_0^2), \quad [\text{S64}]$$

where  $k_0 = 1/r_0$ ; the wave numbers well above  $k_0$  in the input spectrum do not contribute noticeably.

The natural spectrum is believed to have the shape  $|I_0(k)|^2 = 1/k^2$  for  $k_{min} < k < k_{max}$ . There should be a cutoff or a faster drop at  $k > k_{max}$ . This short wavelength range is of no importance, though. The lower end of the spectrum, however, significantly affects the correlations. Indeed, for  $k \ll k_0$  and  $kr \ll 1$  one has  $J_0(kr) \approx 1$ . If  $k_{min} \ll k_0$ , it is the lowest end of the spectrum that dominates in Eq. S61 and the correlation remains large up to large distances. In particular, the presence of  $k = 0$  (nonzero mean of the input) would result in long-range correlations. The presence of  $k^2$  in Eq. S63 efficiently removes this dependence on the low wavenumber end of the spectrum, that is, on the largest scales of the input. Fig. S1 shows the normalized covariance for SF and eye movements for several values of  $k_{min}$ . Eye movements effectively limit the effects on the correlation to the scales approximately within an order of magnitude around the receptive field size. For this visualization the spectrum was assumed flat  $|I_0(k)|^2 = |I_0(k_{min})|^2$  for  $0 \leq k \leq k_{min}$ .

This can be understood as follows: let  $I_0(x) = \bar{I} + \tilde{I}$ , where  $\bar{I}$  is the spatial mean of the intensity,

$$\bar{I} = \langle I_0 \rangle \equiv \frac{1}{L_x L_y} \int I_0(x) dx, \quad [\text{S65}]$$

whereas  $\langle \tilde{I} \rangle = 0$ . In other words, we decompose the input onto full-field flash and zero mean input. The corresponding output is  $O = \bar{O} + \tilde{O}$ , where

$$\bar{O}(x) = L(t) \int X(x-x') \bar{I} dx' \quad [\text{S66}]$$

$$\tilde{O}(x) = L(t) \int X(x-x') \tilde{I}(x') dx'. \quad [\text{S67}]$$

It is easy to see that  $\langle \tilde{O} \rangle = 0$  while  $\bar{O} = \text{const}$ . The correlation would be

$$C(r) = \langle O(x) O(x+r) \rangle = \bar{O}^2 + \langle \tilde{O}(x) \tilde{O}(x+r) \rangle. \quad [\text{S68}]$$

Respectively,

$$CC(r) = \frac{\bar{O}^2 + \langle \tilde{O}(x) \tilde{O}(x+r) \rangle}{\bar{O}^2 + \langle \tilde{O}(x)^2 \rangle}. \quad [\text{S69}]$$

Therefore, even if  $\langle \tilde{O}(x) \tilde{O}(x+r) \rangle \rightarrow 0$  for  $|r| \rightarrow \infty$ , the normalized correlation

$$CC(|r| \rightarrow \infty) \Rightarrow \frac{\bar{O}^2}{\bar{O}^2 + \langle \tilde{O}(x)^2 \rangle}. \quad [\text{S70}]$$

The larger is ratio  $\bar{O}^2 / \langle \tilde{O}(x)^2 \rangle$  the closer is the correlation to unity. In other words, the long-range correlation is determined by the full-field flash amplitude. It is worth noting that the time average output is assumed to be zero here, whereas there are no restrictions on the spatial mean. The center-surround antagonism reduces the output corresponding to the full-field flash. If  $\int X(x) dx = 0$  then  $\bar{O} = 0$  and the correlations at large distances can be expected to drop to zero, otherwise long range correlations persist.

#### 5. Whitening and Correlations

For the totally spatially uncorrelated input (white noise)

$$\langle I_0(x) I_0(x+r) \rangle \propto \delta(r), \quad |I_0(k)|^2 = \text{const} \quad [\text{S71}]$$

the retinal response in the case of SF induces correlations corresponding to the shape of the receptive field:

$$\tilde{C}_{SF,WN}(r) = G \int_0^\infty k dk |X(k)|^2 J_0(kr). \quad [\text{S72}]$$

That is, uncorrelated input is transformed in a correlated signal transferred further to the brain. These correlations are known to the brain because they depend only on the retina design. Thus, ignoring these correlations is easy and can be done using a single built-in algorithm.

For the natural images, in the case of SF, the output sent to the brain would have correlations that mix the retina design and the intrinsic image correlations corresponding to the spectrum  $|I_0(k)|^2 \propto k^{-2}$  (here, to simplify the discussion we assume that  $|I_0(k)|^2 = 0$  outside  $k_{min} < k < k_{max}$ ):

$$\tilde{C}_{SF,natural}(r) = G \int_{k_{min}}^{k_{max}} k^{-2} k dk |X(k)|^2 J_0(kr). \quad [\text{S73}]$$

Thus, decoding would require taking into account this mixture. Small-amplitude eye movements essentially convert the spectrum of the natural images into the (truncated) white noise spectrum. Thus, eye movements remove the correlations due to the image and retain only the correlations due to the receptive fields (compare (37) with  $|I_0(\mathbf{k})|^2 = \text{const}$  with (60) with  $|I_0(\mathbf{k})|^2 \propto k^{-2}$ ). When doing so the finite size of the receptive field kills the shortest scales and the eye movements kill the largest scales of the image.

## 6. Beyond the Linear Model

The above theory clearly shows that correlations drop more quickly with the distance for eye movements than for SF. However, the theory makes use of a number of approximations that are broken when we deal with real cells: infinite retina, identical cells, continuum description, and linearity of the response. At this stage we have no analytical approach that would allow us to get rid of these approximations, in particular linearity of the response, and calculate the correlations. To access the effects of finiteness of the retina, discreteness of the cells, and nonlinearity we develop a simple numerical model. This model should be considered as an extension of the theoretical analysis and not as modeling of real cells measured in the described experiment. In what follows we show the lifetime correlation coefficient, calculated in accordance with Eqs. S12 and S15.

Input for the model is given on the grid of the size  $L_{grid} \times L_{grid}$  pixels,  $L_{grid} = 400$ , when each  $200 \times 200$  quadrant is a natural image. Each image is randomly chosen from a large set of natural images and rotated by  $n \times 90^\circ$ ,  $n = 0, 1, 2, 3$ , according to the position of the quadrant. In this way averaging over natural images and directions is achieved. The input is changed so that 10 trials are combined in the lifetime observation. Identical cells (total number  $n_{cells} = 21^2$ ) are first positioned on a subgrid with the spacing 15 pixels and then randomly shifted (uniform distribution with a maximum shift of about 20 pixels). Spatial shape of the receptive field is taken as a difference of two Gaussians (center and surround):

$$RF(\mathbf{r}) = \frac{1}{L_c^2} \exp(-r^2/2L_c^2) - a \frac{1}{L_s^2} \exp(-r^2/2L_s^2). \quad [S74]$$

The coefficient  $a$  controls the integral  $\int RF(\mathbf{r})d\mathbf{r}$  over the whole plane. For  $a = 1$  this integral vanishes, which corresponds to no response to a uniform intensity. Fig. S2 shows the distribution of the receptive fields on the grid. Hereafter all results are for  $L_c = 10$  if not specified otherwise. This value of the center width is chosen to ensure substantial overlap of the receptive fields for neighboring cells.

Fig. S3 shows the number of cells vs. distance between the cells normalized on the RF center size  $L_c$ . The number of cells grows linearly up to about the distance of 10 receptive fields. Beyond this distance the effects of the finite size and edges may become substantial.

The linear output of each cell is taken instantaneous for SF:

$$O_L(\mathbf{x}, t) = \int RF(\mathbf{x} - \mathbf{y}, t) I(\mathbf{y}, t) dt. \quad [S75]$$

For identical cells the temporal response is not important in the case of SF. Nonlinearity can be implemented in a number of simple ways:

- i) The actual nonlinear output is made positive for both positive and negative linear output (75) but taking absolute values  $O_{NL} = |O_L|$  (absolute value).
- ii) The negative values of the linear output are converted to zeroes but the positive values are unchanged:  $O_{NL} = 0$  if  $O_L > 0$  and  $O_{NL} = O_L$  if  $O_L < 0$ .
- iii) The positive values of the linear output are converted to zeroes but the positive values are unchanged:  $O_{NL} = 0$  if  $O_L < 0$  and  $O_{NL} = O_L$  if  $O_L > 0$  (rectifying function).

These implementations take into account that the firing rate is always positive, although the input may be both positive and negative. Cases ii and iii roughly describe only off or only on cells. In what follows we show the numerical results for the off case, because it resembles more closely the experimentally found cell features. The differences with other nonlinearities are only quantitative (the value of the correlation at large distances) but not qualitative. For the SF mode the retina starts firing after adaptation to the gray background (we follow the experimental setup), so that the actual input is the intensity with the gray background subtracted. In the case of eye movements the retina adapts to the still image. Eye movements are simulated with 30 random circular shifts, in the range  $-10: +10$  in each direction, of the whole grid. The walk of an image on the retina is shown in Fig. S4.

Fig. S5 shows the scatter plots for the cross-correlation. The left panel compares SF (black points) and FEM (red circles) obtained in the numerical model. The middle panel compares the model obtained (black) and experimentally found (red circles) correlations for SF, and the right panel shows similar comparison for FEM.

Fig. S6 shows the cross-correlation for SF (blue curves) and FEM (black curves). The cross-correlations are obtained within the numerical model described above. In both cases (SF and FEM) the response is nonlinear (off only). For SF the active time is only a part of the trial time. All means are calculated here with the ratio of SF active-to-trial time of 0.2. The lower blue curve corresponds to  $a = 1$ , when the center-surround antagonism cancels the response to a uniform background, and the upper blue curve corresponds to  $a = 0.9$ . The long-range correlations grow with the decrease of  $a$ . The cross-correlations for FEM are not sensitive to the surround-to-center amplitude ratio  $a$ . Pitkow and Meister (3) note that the center-surround antagonism does not have to be ideal, that is,  $a \neq 1$ , in general. We see that the amplitude affects the baseline for SF but does not affect the correlation for eye movements. Thus, eye movements effectively cancel the sensitivity to the relative strength of the center and surround. To show that different forms of nonlinearities do not change the above results we present in Fig. S7 the results of the numerical model for the two choices of nonlinearities: absolute value and rectifying function (discussed above). Although the values of the correlation at large distances are affected, the FEM correlations are always lower than SF correlations.

## 7. Experimentally Determined Correlation Patterns

Fig. S8 shows the mean values of correlation for pairs of cells with similar receptive field spacing. The correlation for SF decreased slowly with increasing distance and remained significant for distant cells. In the presence of fixational eye movements the level of correlations dropped, as proposed by the theory. Only pairs of the same type (Off-Off) were used. The vertical line represents the mean receptive field size in all panels. The correlation pattern when we use only time windows of 1 s to 2 s of both the SF and FEM segment was calculated. We found that in the SF segment the retina fires spontaneously and the activity is not derived by the stimulus.

1. Rucci M (2008) Fixational eye movements, natural image statistics, and fine spatial vision. *Network* 19(4):253–285.  
 2. Kuang X, Poletti M, Victor JD, Rucci M (2012) Temporal encoding of spatial information during active visual fixation. *Curr Biol* 22(6):510–514.

3. Pitkow X, Meister M (2012) Decorrelation and efficient coding by retinal ganglion cells. *Nat Neurosci* 15(4):628–635.









