ONLINE METHODS

Geometry of V1

We assume a columnar architecture of the primary visual cortex (V1) with the cortical hypercolumns packed hexagonally in cortical space (**Fig. 2a**). The computational units within each hypercolumn have receptive fields centered at the same retinal location but each tuned to different orientations and spatial scales. For our purposes, we are concerned mainly with the orientation tuning of these units. We thus use 8 broadband oriented filters to extract the corresponding orientation energy in the image patch under the RF ($\theta = [0,1,\ldots,7] \times \frac{\pi}{8}$ radians; **Fig. 4a**). The average diameter of the receptive fields increase linearly with eccentricity with a slope of 0.1[24].

We can show that this geometry captures the logarithmic cortical magnification as follows. Let d_0 be the RF diameter (in degrees) at eccentricity $\phi = 0$, s be the slope of the linear function relating RF diameters to eccentricity and γ be the proportion of RF diameter overlap between adjacent hypercolumns (assumed constant across all eccentricities). Within a V1 hypercolumn, the eccentricity ϕ of an RF at the cortical distance p, measured in units of hypercolumns from the location on V1 that represents the center of the fovea, is given by:

$$\phi(p) = \frac{\alpha}{\beta - 1}\left(\beta^{p-1} - 1\right) \qquad (1)$$

where $\alpha = d_0(1-\gamma)$ and $\beta = 1 + s(1-\gamma)$. For our model, $d_0 = 0.01$, $s = 0.1$[24] and $\gamma = 0.3$ (average value in REF [49]). Away from the immediate vicinity of the center of gaze, Equation 1 can be written in approximate form as

$$\phi(p) \approx \frac{\alpha}{\beta(\beta-1)}\beta^p \tag{2}$$

Inverting the function, we can express p as a function of ϕ as

$$p(\phi) = \frac{\ln(\phi)}{\ln(\beta)} - K \tag{3}$$

where $K = \ln\left(\frac{\alpha}{\beta(\beta-1)}\right) / \ln(\beta)$. Equation 3 thus gives the logarithmic cortical magnification outside of the immediate center of gaze.

If we assume the critical spacing in the visual field for crowding to be $\Delta\phi = b\phi$, where b is Bouma's constant (about 0.5), then the critical spacing in the cortex is

$$\begin{aligned}
\Delta p &= p(\phi + \Delta\phi) - p(\phi) \\
&= \frac{\ln(\phi + \Delta\phi) - \ln(\phi)}{\ln(\beta)} \\
&= \frac{\ln(\phi + b\phi) - \ln(\phi)}{\ln(\beta)} \\
&= \frac{\ln(1 + b)}{\ln(\beta)}
\end{aligned} \tag{4}$$

Using $b = 0.5$ we get $\Delta p \approx 6$. That is, the critical spacing for crowding corresponds to six hypercolumns in V1, independent of eccentricity. This is in agreement with the anatomical extent of lateral (long-range horizontal) connections in V1[18] and with the estimated extent of "combining fields" in V1[10] if each hypercolumn is roughly 1 mm on the cortex.

Conversely, Equation 4 shows that if a computational unit in a particular hypercolumn has lateral connections to all computational units in neighboring hypercolumns up to an isotropic extent of a constant number of hypercolumns on the cortex, then the resulting spatial interaction in the visual field must follow Bouma's Law of linear scaling ($\Delta\phi = b\phi$). We will refer to the set of hypercolumns (blue circles in **Fig. 2a**) to which a reference hypercolumn (red circle in **Fig. 2a**) has lateral connections, as the lateral interaction zone. In our model, we set the radius of the lateral interaction zone to 6 hypercolumns.

Saccadic eye movements

For the eye-movement simulations, the saccadic velocity profile was modeled as follows. Let A be the saccade amplitude (the distance between successive fixations in degrees of visual angle), T the duration and $v(t)$ the velocity profile of a saccade. $v(t)$ must satisfy the following conditions[50]:

$$\begin{aligned} v(0) &= 0 \\ v(T) &= 0 \\ v_{\text{peak}} &= v\left(\tfrac{T}{2}\right) = k\sqrt{A} \\ \left.\frac{dv}{dt}\right|_{t=\frac{T}{2}} &= 0 \end{aligned} \quad (5)$$

where k is a constant. A sinusoidal velocity profile of the following form satisfies the constraints in Equation 5 in the range $[0 \le t \le T]$:

$$v(t) = k\sqrt{A}\sin\frac{\pi t}{T} \quad (6)$$

Since $A = \int_0^T v(t)\,dt$, we have:

$$T = \frac{\pi}{2k}\sqrt{A} \quad (7)$$

The distance traversed, $D(\tau)$, in time τ is therefore given by

$$\begin{aligned} D(\tau) &= \int_0^\tau v(t)\,dt \\ &= \frac{A}{2}\left(1-\cos\frac{2k\tau}{\sqrt{A}}\right) \end{aligned} \qquad (8)$$

The distribution of saccade amplitudes along the radial axis from the fovea was modeled as an exponential distribution[31] with the following p.d.f. : $f(x)=\lambda e^{-\lambda x}$, $\lambda=\frac{1}{7.6}$. The distribution along the iso-eccentric axis was assumed to be uniform.

Eye-movement simulations and image statistics

Using the distribution of saccade amplitudes and the corresponding velocity profile described above, we simulated saccadic eye movements in which the visual stimulus presented to the system was a random clutter of uppercase letters (Palatino font) at various sizes and orientations. For computational tractability we calculated the outputs of the set of 8 broadband oriented filters within each hypercolumn at discrete time points in the interval $[0 \ldots T]$. Each filter measures the contrast energy along a given orientation in the image patch that is in the receptive field of the hypercolumn. Let $r[t]$ denote the response of one of the filters at time t. The cumulative response of the filter over the time course of the eye movement is

$$\mathbf{r}=\sum_{t=0}^{T} r[t]e^{-t/\lambda} \qquad (9)$$

where the modulation of spatial attention during its overlap with saccadic eye movement (**Fig. 2b**) is modeled as an exponential decay function with a time constant λ, a free parameter of the model. Such a characterization captures the probabilistic distribution of the overlap period. For the purpose of calculating joint image statistics, the cumulative filter response, $\mathbf{r}$, is first converted into a firing probability p with a saturating non-linearity:

$$p = \tanh(k\mathbf{r}) \tag{10}$$

Let $\Theta_{i,\mathrm{R}}$ be a random variable associated with a filter with orientation θ_i in the reference hypercolumn R. $\Theta_{i,\mathrm{R}}$ is equal to 1 if the cell has fired within a temporal window, else it is zero. For simplicity, the temporal window used in our simulations was the entire duration of a saccade. The joint probability distribution $P(\Theta_{i,\mathrm{R}},\Theta_{j,\mathrm{N}})$, between the oriented filter in the reference hypercolumn and another oriented filter in a neighboring hypercolumn (N) can be calculated by accumulating and averaging the joint firing probabilities across many eye-movement traces (30000 in our simulations). To obtain robust estimates of the joint probability distribution we used the bootstrap procedure. For any saccade trace, the probabilities are accumulated only if both the reference and the neighboring hypercolumn are under the spotlight of attention. Finally, the statistical dependence between $\Theta_{i,\mathrm{R}}$ and $\Theta_{j,\mathrm{N}}$ can be calculated in terms of the pair-wise mutual information

$$I(\Theta_{i,\mathrm{R}};\Theta_{j,\mathrm{N}}) = \sum_{\substack{\Theta_{i,\mathrm{R}}=\{0,1\}\\ \Theta_{j,\mathrm{N}}=\{0,1\}}} P(\Theta_{i,\mathrm{R}},\Theta_{j,\mathrm{N}})\log_2 \frac{P(\Theta_{i,\mathrm{R}},\Theta_{j,\mathrm{N}})}{P(\Theta_{i,\mathrm{R}})P(\Theta_{j,\mathrm{N}})} \tag{11}$$

The mutual information is zero when the two random variables are statistically independent.

For a reference hypercolumn R, pooled mutual information (pooled across all orientations) between R and a neighboring hypercolumn N is defined as

$$
\begin{aligned}
\mathbf{I}_{\mathrm{SC}}(\mathrm{R};\mathrm{N}) &= \sum_i \sum_j I_{\mathrm{SC}}\left(\Theta_{i,\mathrm{R}};\Theta_{j,\mathrm{N}}\right) \\
\mathbf{I}_{\mathrm{V}}(\mathrm{R};\mathrm{N}) &= \sum_i \sum_j I_{\mathrm{V}}\left(\Theta_{i,\mathrm{R}};\Theta_{j,\mathrm{N}}\right)
\end{aligned} \tag{12}
$$

where I_{SC} and I_{V} are the pairwise mutual information for the saccade-confounded and veridical conditions respectively (Equation 11). We express the gross difference between the saccade-confounded and veridical statistics in term of the normalized difference between saccade-confounded and veridical mutual information:

$$
\Delta I(\mathrm{R};\mathrm{N}) = \frac{\mathbf{I}_{\mathrm{SC}}(\mathrm{R};\mathrm{N}) - \mathbf{I}_{\mathrm{V}}(\mathrm{R};\mathrm{N})}{\mathbf{I}_{\mathrm{V}}(\mathrm{R};\mathrm{N})} \tag{13}
$$

This normalized difference when plotted in visual space for all neighboring hypercolumns maps the amplitude and spatial extent of inappropriate feature integration for a reference hypercolumn. Image features from hypercolumns with negative difference (reduced mutual information in the saccade-confounded statistics as compared with the true statistics) would have weaker interactions and thus be only loosely bound to the reference features, leading to an under-integration of features. Conversely, features from hypercolumns with positive difference (excessive mutual information in the saccade-confounded statistics) would strongly influence the reference, leading to excessive and possibly erroneous feature integration.