

Continuum Dynamics Model of the Primary Visual Cortex for Contour Detection

Vijay Singh,¹ Martin Tchernookov,¹ Rebecca Butterfield,² and Ilya Nemenman^{1,2}

¹*Department of Physics, Emory University, Atlanta, GA 30322, USA*

²*Department of Biology, Emory University, Atlanta, GA 30322, USA*

We model the neural dynamics in the primate primary visual cortex in terms of a continuous director field that describes the average rate and the average orientational preference of active neurons at a particular point in the cortex. This representation has a correspondence to the Landau - de Gennes order parameter for nematic liquid crystal in two dimensions. Our linear-nonlinear dynamical model incorporates long range connectivity patterns that enforce context dependence present in the visual field. The model can distinguish large contiguous objects from the background clutter by suppressing the clutter and by filling-in occluded elements of object contours. This results in high-precision, high-recall detection of large objects in cluttered scenes.

PACS numbers: 84.35.+i, 87.85.dm, 87.19.lj, 42.30.Tz

To recognize an object in a visual scene, humans and other primates process visual signals relayed through the retina [1] in the ventral stream of the cortex. Contour detection is a crucial part of this process (Fig. 1). It is carried out at early stages of the processing, in the area of the brain called V1 or the primary visual cortex [2]. V1 consists of hundreds of millions of neurons organized topographically into columns of $\sim 10^4 \dots 10^5$ neurons each. Neurons in each column receive inputs from a localized part of the visual field (called classical, or feedforward receptive field). They are directionally selective, responding primarily to oriented edges within their receptive fields [3, 4]. Computational vision models that account for such receptive fields of individual neurons [5–10] typically incorporate them within feedforward hierarchical structures similar to the cortex [11, 12]. Such feedforward models account for the visual processes on short time scales, and they achieve the error rate as low as $\sim 10 - 20\%$ on typical object detection tasks [10, 13].

It is believed that, *in vivo*, the error rate is reduced by orders of magnitude by contextual information that influences local processing, and is not captured in classical models [14, 15]. These collective, recurrent dynamics span large spatiotemporal scales and are supported by thousands of axons laterally connecting distant columns [16]. These interactions are believed to suppress the clutter present in the visual field, while simultaneously binding contours across occlusions [17].

The goal of this paper is to build a model of the primary visual cortex that simultaneously achieves these two contradictory tasks: removing clutter and filling in occlusion gaps. For this, we elaborate on the proposal that captures important properties of lateral connectivity among V1 neurons [18, 19]. Specifically, we incorporate the Hebbian constraint that neurons that are excited simultaneously by the same long, low-curvature contours should activate each other [18]. Since the number of neurons in V1 is ~ 100 million, with each neuron having $\gtrsim 10^3$ connections, some of which extend for many millimeters, we do not focus on individual neurons, as most

models do, but represent the activity as a *coarse-grained, continuous neural field*, which we model as a complex-valued field on the complex plane, $W(z)$. The magnitude and the phase of W represent the level of excitation and the orientation of the contour element at point z . The purpose of the coarse graining is to identify which features of the neural structure and dynamics are essential for contour recognition, and which may be omitted.

The complex field crucially distinguishes our model from individual neurons models, as well as from most other coarse-grained models. Indeed, typical continuum approaches represent the activity as a real function of three variables (position in the visual plane and the directional sensitivity) [20, 21], and do not account for the parity symmetry of the neural field. Thus our model has fewer degrees of freedom and is simpler. Other models that used a similar complex field representation [22, 23] have focused on development, rather than on the visual performance of the cortex. Thus an ability of a reduced, lower-dimensional model to solve visual tasks is far from certain. Here we show that our model can *reconstruct object contours with very high precision and recall* even in the presence of dense clutter and large occlusions.

The model – The dynamical variables in our model are the neural firing rate $s(x, y)$, $s \geq 0$, and the dominant orientation preference Θ of active neurons, both averaged over a microscopic patch of the cortex, which still contains many thousands of neurons. Such averaging is traditional in, for example, fluid dynamics, where continuous dynamics is sought from discrete agents. The neural activity is invariant under parity (i. e., an edge or its π rotation results in the same activity). Further, two equal edges at one point oriented $\pi/2$ apart lead to cross orientation suppression and no dominant orientation at the point [24]. So the fields s and Θ are combined into a time varying complex field $W(z, t)$ in a somewhat uncommon way, forming an object called a *director* [25]: $W(z, t) = s \times e^{i2\Theta}$. The magnitude of this field is the average of the firing rate, and the argument is twice the average orientation preference of the dominant neurons

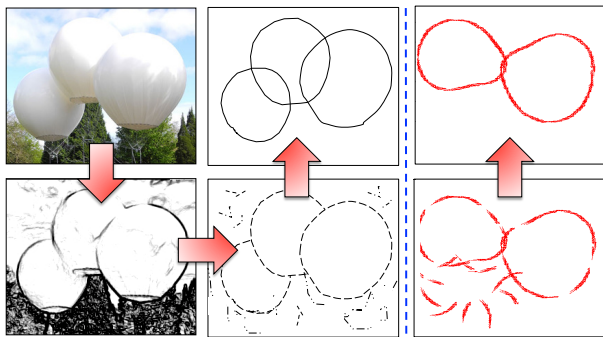


FIG. 1: Contour Reconstruction Task: A 2d image (left top; credit: <http://9bytz.com/balloon-bridge/>) is recorded as a field of contrast by the retina and the LGN (left bottom). V1 neurons respond to regions of contrast changes in a direction-selective manner, performing edge detection (middle bottom). The information from edges is integrated to reconstruct long contours (middle top). In this paper, we model the visual process starting from edges in V1; sample input (bottom) and output (top) to our model are on the right.

at a point $z = |z|e^{i\theta} = x + iy$ [22, 23]. We similarly coarse-grain the input images, identifying the dominant orientation at every point (see *Methods*). This orientation field serves as the input to the model.

Neurophysiological and psychophysical experiments [14, 15, 17, 26, 27] and theoretical considerations [18] suggest that neurons in V1 are laterally connected, such that active neurons excite nearby neurons with collinear or large-radius co-circular directional preference. Conceptually, simultaneous input from several collinear or co-circular neurons can excite other neurons that might otherwise not be getting enough excitation from the visual field due to occlusion or noise, cf. Fig. 2(a). At the same time, neurons responding to high spatial frequency clutter elements will not get sufficient lateral excitation, and their activity will decay. These collective dynamics integrate information over large spatial scales.

We can represent these phenomena in a traditional linear-nonlinear model, where the neural field at a point z is affected by a combination of lateral synaptic inputs:

$$\frac{dW(z, t)}{dt} = F_{\delta_{\text{th}}}[I(z, t)] - r(z, t) + j(z, t). \quad (1)$$

Here $F_{\delta_{\text{th}}}$ is some sigmoidal function of the excitatory input $I(z, t)$, $r(z, t)$ describes the inhibitory contribution to the field, and $j(z, t)$ is the stimulus.

The excitatory input, $I(z, t)$, combines synaptic input from all points z' in its interaction region ‘Ex’

$$I(z, t) = \int_{\text{Ex}} d^2z' K[z - z'|W(z, t)] W^*(z', t), \quad (2)$$

where $K[z - z'|W(z, t)]$ is the excitatory interaction kernel between the fields at point z' and z , when the field at z is $W(z, t)$. The kernel for an arbitrary orientation

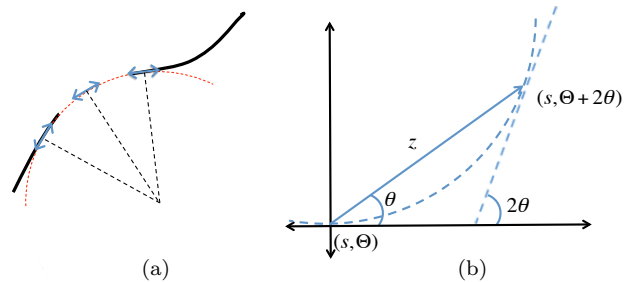


FIG. 2: (a) Neurons send excitatory signals along tangential co-circular directions. Thus neurons in occluded gaps may get enough excitatory input along smooth contours. (b) Co-circularity condition: The orientation at two points is said to be co-circular if they are tangential to the circle connecting the two points. If the orientation preference at the origin is along the real axis, the co-circular edge at a point $z = |z|e^{i\theta}$ has the orientation 2θ . Multiplication by $e^{i2\theta}$ can be written as: $e^{i2\theta} = (e^{i\theta})^2 = (z/|z|)^2 = (z/z^*)^2$.

of $W(z)$ can be defined by an appropriate rotation of the kernel defined for $W = 1$ (parallel to the real axis):

$$K[z - z'|W] = K \left[(z - z') e^{-i \frac{\arg(W)}{2}} |1 \right]. \quad (3)$$

Co-circular excitation may be represented as

$$K[z|1] = \left(\frac{z}{z^*} \right)^2 \times \exp \left\{ -\frac{|z|^2}{2\sigma^2} - \mu \frac{|\text{Im}(z)|}{[\text{Re}(z)]^2} \right\}. \quad (4)$$

The first term, derived in Fig. 2(b), determines the field direction at z that is co-circular to the field at z' , and is thus excited. The σ term in the exponent determines the spatial range of the excitation. Finally, the μ term determines the smallest radius for which substantial co-circular excitations still exist, giving the kernel its characteristic bowtie shape (*Methods*) [18].

We define the input nonlinearity using a complex step function: $F_{\delta_{\text{th}}}(I) = \frac{I}{|I|} \times A H(|I| - \delta_{\text{th}})$, where H is the Heaviside step function, and A is a constant that determines the maximum excitation strength. Smoother sigmoidal nonlinearities had little effect on the results presented below. Thus if the total excitatory input is higher than the threshold δ_{th} , then the field $W(z)$ gets a positive increment in the direction of the total input. For this to happen, the excitatory contribution from a large part of the neural field must align in the same direction, representing coincidence detection. While importance of this phenomenon in vision is unclear, it is crucial in the context of auditory signal processing [28]. The thresholding also suppresses clutter-induced spurious excitations, as it is unlikely that the excitatory input from short clutter elements becomes higher than the threshold in the absence of contextual support from long contours.

The inhibitory term r in our model represents two distinct phenomena: local relaxation, which depends only

on the local value of the field at a particular point [29], and global inhibition [30], which keeps the activity of the entire neural field in check (presumably through intermediate inhibitory neurons, not modeled explicitly):

$$r(z) = \gamma_l W(z) + \gamma_g H(|W(z)|) \frac{W(z)}{|W(z)|} \int_{\text{In}} d^2 z' |W(z')|. \quad (5)$$

Here γ_l and γ_g determine the rates of local and global inhibition, and ‘In’ stands for the range of global inhibitory interactions. Combined with the non-linear excitation, this linear inhibition produces bimodal asymptotic field values. This makes it easier to define which neuron is ‘active’ and which one is not.

Test data set – The time evolution of this model was studied numerically using randomly generated computer images, whose orientation fields were presented as inputs $j(z)$ in Eq. (1) (Fig. 3). The images contained one or two large curvilinear closed contours (targets or amoebas) with gaps that emulated object occlusion. The images also contained small clutter elements that were obtained from reshuffling fragments of another amoeba, and hence were locally indistinguishable from targets. The goal was to filter out the clutter while retaining and enhancing the targets. Since we intended to capture the biological processes in V1 only, we chose synthetic rather than natural images. There the known generation rules ensured that the targets and the clutter were only distinguishable by their context of belonging to a large contour, and not by other features such as texture, curvature, or continuity. Generation of images largely followed [19], and is described in the *Methods*. Finally, we ensured that essentially the same edges did not belong to the target and the clutter at the same time – this made scoring the model performance unambiguous, but without artificially inflating the performance itself. We tested the model with a temporally constant input $j(z, t) = j(z)$, as well as instantaneous inputs $j(z, t) = j(z)\delta(t)$. Instantaneous inputs model exposures of as little as 10-20ms in speed-sight-experiments, which is nonetheless often sufficient for object detection (see [19] and references therein). The results are similar in both cases, and in the rest of the paper we present results from the instantaneous case.

Results —The model, Eq. (1), was solved numerically for specific visual inputs $j(t)$; the numerical integration procedures can be found in the *Methods*. Figure 3 shows the time evolution of the neural field $W(z, t)$ for a sample input image, and for an image used in psychophysics experiments with human subjects [31]. The gaps present in amoeba target get filled, while the clutter decays with time, resulting in emergence of long contours. Note also that spurious activity appears around contours at large simulation times, so that the best performance happens at an intermediate time.

The performance of the model was quantified in terms of precision, P , and recall, R . Precision determines the

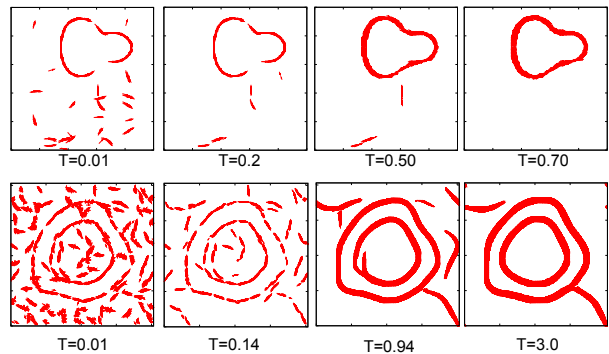


FIG. 3: (top) Time evolution of the neural field for a sample image. The magnitude (line width) and the direction of the field are plotted at every point where the strength of the field is higher than a cutoff (0.35). Here and elsewhere in this work we use $A = 5, \delta_{\text{th}} = 5, \sigma = 7.9, \mu = 15, \gamma_g = 0.012, \gamma_l = 1$, which optimizes performance according to a genetic algorithm. Dynamics removes the clutter and fills in the occlusion gaps. However, spurious activity (widening lines) appears for large simulation times, so that the best performance is obtained for intermediate times. (bottom) Performance of the model on an image used in psychophysics experiments [31]; like human subjects, the model can identify long contours.

fraction of the total field activity integrated over the image that matches the actual target contour (visible and occluded). Recall gives the fraction of the target contour that has been recovered. $P = 1$ means there is no clutter, and $R = 1$ means all parts of the contour have been identified. For a successful contour detection, we must have $R, P \rightarrow 1$ simultaneously. Both P and R depend on the cutoff used to decide which neurons are considered active (larger cutoff degrades clutter faster, but slows down occlusion filling), and on the time of the simulation. Hence different cutoffs and times must be explored.

Figure 4 gives the variation of precision and recall at various cutoffs at a particular time during the simulation. The neural field at $t = 0$ has $(R, P) = (0.75, 0.5)$, on average. That is about 25% of the target is occluded, and the total lengths of the clutter and the target segments are about the same. At t as small as 0.25 (with the simulation time step $\Delta t = 0.01$), P, R are above 0.9 simultaneously for a large set of cutoff values (1% – 42%). Since we present the stimulus instantaneously only, its effect eventually decreases with time. Thus there is a time that optimizes performance; i.e., at which the precision vs. recall curve majorates the same curves for other times. For the data-set in Fig. 4, this optimal time is $t = 0.40 \times 1/\gamma_l$ (40 numerical iterations), where the curve reach $R \approx 0.97$ and $P \approx 0.95$ simultaneously.

Performance of the model depends only weakly on the *ad hoc* details of the simulations and the data. For example, we also defined the threshold parameter not as an absolute value, but as a fraction of the maximum activity of the field at a given time point; this did not

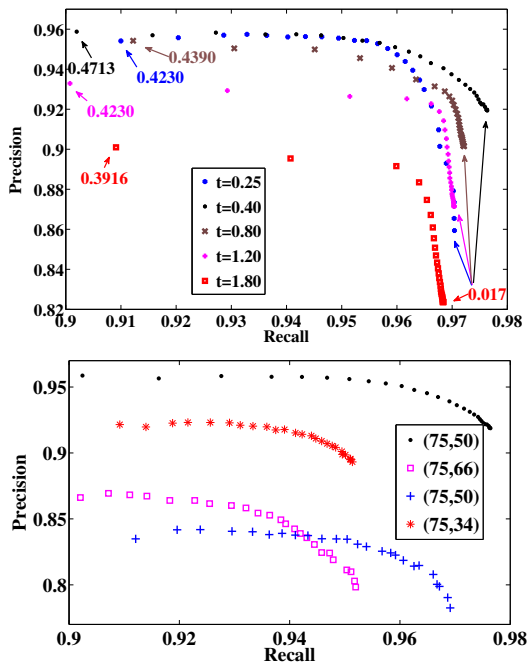


FIG. 4: (top) P vs R averaged over 500 randomly generated images at various simulation times starting with $(R, P) = (0.75, 0.5)$. The numbers indicate cutoff values for a specific data point at the corresponding simulation time. Note the weak dependence on the cutoff. The simulation lengths of $t = 0.40 \times 1/\gamma_1$ (black dots) produces the curve with the best precision and recall combination. (bottom) P vs R with different starting values of precision and recall averaged over 100 randomly generated images, but with the same model parameters. Legend indicates the initial (R, P) . The black dots are the same as in the top panel. Red *’s correspond to a lower precision (more clutter), compared to the black dots. Blue +’s stand for the same initial (R, P) as black, but with the target partitioned into more shorter segments (a larger number of occlusions). Pink \square ’s correspond to higher initial precision (less clutter), but the clutter elements are longer and harder to suppress.

change the precision-recall curves much. Similarly, different amounts of initial clutter had only a moderate effect if the length of the clutter elements remained the same (Fig. 4, bottom). This is because the time scale of the clutter decay depends on the size of the segments, and not on their number. For longer segments, the decay takes longer, and hence the optimal processing time increases. The optimal processing time also increases with the linear dimension of the occlusions present in the target amoebas and with the number of occlusions (Fig. 4, bottom). However, for all of these cases, the maximum precision and recall remains simultaneously high.

Discussion – We developed a *continuum, coarse-grained model* of the primary visual cortex to study contour detection in complex images. The model describes neural activity as a parity-symmetric continuous director field. The model incorporates some experimen-

tally observed properties of the visual neural dynamics, namely non-linear excitation, thresholding, cross orientation suppression, local relaxation, global suppression, and, crucially, co-circular excitatory connectivity [18], which brings long-range context to local edge detection.

The model identifies long object contours in computer-generated images with the simultaneous recall and precision of over 90% for many conditions. This happens even though large parts of objects are occluded (potentially lowering recall), and clutter is present (potentially decreasing precision). The model fills in the occlusions and filters out the clutter based on the presence or absence of co-circular contextual edge support. The ability to *fill in the occlusions* particularly distinguishes our approach from the previous work on co-circular excitatory feedback [18, 19]. It remains to be seen to which extent the performance is affected by more natural statistics of images, and by the presence of stochasticity and synaptic plasticity in neural dynamics.

The model has been developed to identify computational primitives needed to detect long contours. Thus its ability to perform on par or better than agent-based three-dimensional models (two spatial dimensions and one orientation preference dimension) [19] illustrates that *discreteness of neurons and existence of the orientation preference as an independent variable are not crucial* for this V1 function. The reduced dimensionality makes for a more efficient computational implementation. Thus we propose to augment practical feedforward models of object detection, such as [10], with the laterally connected layer developed in this work. We expect this to lead to improvements in object recognition performance.

The model makes predictions that can be tested experimentally, such as regarding the amount of neural excitation in V1 as a function of the computation time and the duration of exposure to an image. Additionally, the model predicts that the neural activity localizes to long contours with time, which can be tested with certain imaging technologies. Finally, the model can be used to predict the dependence of the contour detection performance on the statistical structure of images and on the exposure time. Testing such predictions in psychophysics experiments [19] will be a subject of the future work.

Finally, we notice that the neural field $W(z) = s(x, y) \times e^{i2\Theta(x, y)}$ can be mapped exactly onto the Landau - de Gennes order parameter for a two-dimensional nematic liquid crystal

$$Q_{\mu\nu} = \frac{1}{2} s \begin{pmatrix} -\sin(2\Theta) & \cos(2\Theta) \\ \cos(2\Theta) & \sin(2\Theta) \end{pmatrix}. \quad (6)$$

This may help solve a crucial difficulty in implementing an artificial laterally-interacting neural model: the computational cost of long-range communication. Indeed, one can think of materials with symmetry and dynamical properties such that the neural computation and the

communication are performed by the intrinsic dynamics of the material itself. Potential implementations can include polarizable liquid crystals with long-range magnetic interactions, polar colloidal materials, or heterogeneous solid state materials with long-range connectivity. The liquid crystal analogy suggests the use of the well-developed repertoire of theoretical physics to understand the impact of different terms in the model neural dynamics, Eq. (1). One can hope that the renormalization group treatment of this dynamics can reveal the terms in the interaction kernel K that are relevant for its long-time, long-range aspects.

Acknowledgements: We thank G Kenyon, V Gintautas, M Ham, L Bettencourt, and P Goldbart for stimulating discussions. This work was partially supported by grants from the Army Research Office and the James S. McDonnell Foundation.

METHODS

Co-circularity Kernel

The shape of the co-circular kernel, Eq. (3) is illustrated in Fig. 5, alongside with the dynamics of a single kernel driven by this kernel.

Image generation

Targets – The amoeba targets are generated by choosing a center at a random point in the image, and then drawing the amoeba around this point in polar coordinates, with the radius as a superposition of periodic functions with different radial frequencies, $\rho(\phi) = \sum_{k=0}^n a_k \sin(k\phi + \phi_k)$. The Fourier coefficients a_k are generated randomly from a normal distribution ($\sigma = 1$), with $k \leq n = 3$, and the phases ϕ_k are uniformly distributed between 0 and 2π . To create amoebas that are about the same size, the coefficients are further constrained such that the minimum and the maximum radii of the resulting amoeba and their ratios obey $0.2L < R_{\min} < R_{\max} < 0.3L$, $0.4 < \frac{R_{\min}}{R_{\max}} < 0.6$, where L is the image size. The current then $j(z) = \delta(z - z_e)e^{i2\Theta}$, for every point z_e within 1 lattice spacing away from any point on the amoeba contour, where Θ is tangential to the contour at that point. While generating an amoeba we also determine an exclusion region around it of 8 lattice sites. There clutter elements (see below) with orientations parallel to the closest amoeba segment are not allowed. Without such exclusion, it is impossible to assign an edge to either the clutter or the amoeba unambiguously, and thus the performance scoring (but not the performance itself) is ambiguous.

Occlusions – We simulate occlusions in real-world images by removing parts of the amoeba. A random number

of 2-4 segments with random angular length combining to the total of $\sim 30\%$ of the amoeba length are chosen at random positions along the amoeba contour. Within the chosen segments, the current $j(z)$ is then set to zero.

Clutter – We need the clutter to be indistinguishable from the targets by curvature, brightness, and other local statistics. Thus clutter is generated by first generating an amoeba as described above, partitioning it into small segments, and then randomly shuffling and rotating the segments to break long-range contour continuity. Specifically, the model cortex is divided into 5×5 square regions, which are then randomly permuted. The center-of-mass (CoM) of an image within each region is computed, and the dominant angular orientation is determined. Then each region is rotated around its CoM by a random angle, subject to a constraint that the resulting dominant orientations of neighboring regions are different. The constraint ensures that the clutter does not form long range target-like structures.

Combined images – One or two targets and clutter resulting from breakup of one or two additional targets were then superimposed together to form test images, see Fig. 3, top, for an example. Clutter in the exclusion zones along the amoeba contours was removed.

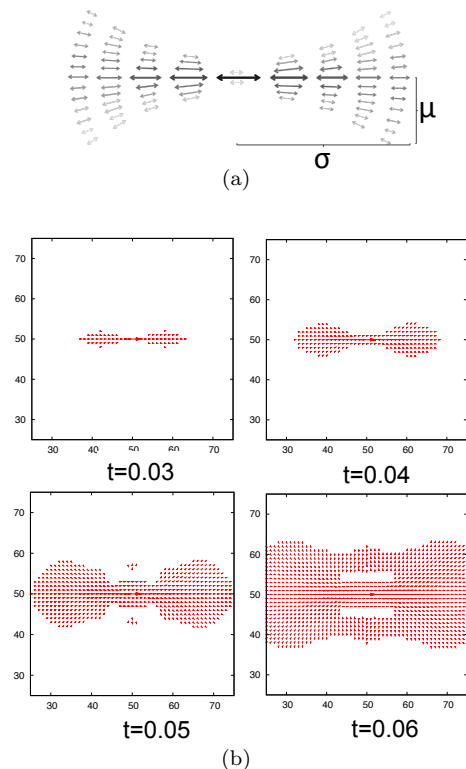


FIG. 5: Schematic shape of interaction kernel $K[z - z'|1]$. Arrows represent the orientation preference and darkness and size represent the magnitude. (b) Results of dynamics with the kernel K with the current $j(z, t) = \delta(z)\delta(t)$. Here, as always, $\mu = 15$, $\sigma = 7.9$.

Transforming pixel images – Images used in psychophysics experiment (Fig. 3, bottom) were imported into MATLAB and then converted to grey scale using `rgb2grey`. The resulting matrix was then thresholded and converted into a binary matrix. A 2D Gabor filter was used to find the edges in this bitmap image. For each point in the image, we find the convolution of a Gabor filter ($\sigma_{\text{smaller}} = 10$ pixel, $\sigma_{\text{longer}} = 100$ pixel, convolution range = 20 pixel \times 20 pixel) with the image at $(360/n)$ angles where $n = 100$. The direction with the maximum convolution is taken as the orientation of the visual field at the point, and the result of the convolution as the field magnitude. The image thus processed is presented as an input for simulation.

Simulations

The time evolution of the model is studied on a square lattice of a linear size $L = 100$ with periodic boundary conditions using Euler iteration method. The lattice discretization is done for simulation purposes, and should not be viewed as a representation of discrete neurons.

In each iteration cycle we first calculate the total input I at each point z from all other points z' in the excitation region ‘Ex’ using a precomputed interaction kernel $K[z - z'|1]$ on a $4L \times 4L$ kernel lattice. Square discretization destroys the angular symmetry of the kernel evaluated at an arbitrary z . The following procedure restored the symmetry. First, to calculate the contribution from z' to $I(z)$, the kernel lattice is superimposed on the image lattice with the origin of the kernel lattice at point z of the image lattice. Next the kernel lattice is rotated by $\frac{\arg(W(z'))}{2}$ with respect to the image. Then the contribution from the point z' to $I(z)$ is $W^*(z') \times K(0, z'')$, where z'' is the point on kernel lattice closest to z' . The total input $I(z)$ is then the sum of contributions from all points z' in the excitatory interaction region ‘Ex’. After the input is calculated, if $|I(z)| > \delta_{\text{th}}$, then the field is incremented $W(z, t + \Delta t) \leftarrow W(z, t) + A \frac{I(z)}{|I(z)|} \Delta t$, where Δt is the time step. To account for degradation, we finally set $W(z, t + \Delta t) \leftarrow W(z, t + \Delta t) \times \exp[-r(z) \times \Delta t / W(z, t + \Delta t)]$, where $r(z)$ is as in Eq. (5). To the first order in Δt , this is equivalent to the dynamics in Eq. (1). However, this exponential form removes the large fluctuations in $r(z)$ when $W(z) \approx 0$.

In our simulations, the excitation range ‘Ex’ is 3σ , where σ is the effective spatial range of the kernel $K[z - z'|W(z)]$. For global inhibition range ‘In’ is the entire lattice. The model is easily modified to restrict the suppression to a smaller inhibition region.

We first chose the parameter μ to be similar to the curvature of a typical amoeba, and σ was chosen to be larger than the extent of the occluded amoeba segments. γ_1 and γ_g were initially determined using steady state analysis of

the model, which leads to $(N\gamma_g + \gamma_1)W_0 \sim F_\delta(\infty)$, where N is the typical number of points with non zero field, and F thresholding function as defined in Eq. 1. Setting $\gamma_1 = 1$ and $W_0 = 1$, we thus constrain all other parameters. After starting with these values, genetic algorithm was used to optimize the model for maximum simultaneous precision and recall. The final optimized values of the parameters used for simulations presented here were: $A = 5, \delta_{\text{th}} = 5, \sigma = 7.9, \mu = 15, \gamma_g = 0.012, \gamma_1 = 1$.

The code was implemented in C, compiled with the gcc v. 4.7, and optimized with OpenMP libraries. Simulations were performed on a computer with Intel i7 2600k (clock speed 3.4 GHz). The simulation time for 250 iteration cycles for one image took about 10s. All model dynamics times were measured in units of $1/\gamma_1$, which was set to 1 in our simulations.

-
- [1] D Felleman and D Van Essen. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex*, 1:1, 1991.
 - [2] O Creutzfeldt and H Nothdurft. Representation of complex visual stimuli in the brain. *Naturwissenschaften*, 65:307, 1978.
 - [3] D Hubel and T Wiesel. Receptive fields and functional architecture of monkey striate cortex. *J Physiol*, 195(1), 1968.
 - [4] D Hubel and T Wiesel. Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *J Physiol*, 160:106, 1962.
 - [5] G Wallis, E Rolls, et al. Invariant face and object recognition in the visual system. *Progr Neurobiol*, 51:167, 1997.
 - [6] L Itti, C Koch, and E Niebur. A model of saliency-based visual attention for rapid scene analysis. *Pattern Analysis and Machine Intelligence, IEEE Trans*, 20:1254, 1998.
 - [7] Y LeCun, L Bottou, Y Bengio, and P Haffner. Gradient-based learning applied to document recognition. *Proc IEEE*, 86:2278, 1998.
 - [8] S Thorpe, A Delorme, and R van Rullen. Spike-based strategies for rapid processing. *Neural Netw*, 14:715, 2001.
 - [9] Y-L Boureau, F Bach, Y LeCun, and J Ponce. Learning mid-level features for recognition. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, page 2559, 2010.
 - [10] T Serre, A Oliva, and T Poggio. A feedforward architecture accounts for rapid categorization. *Proc Natl Acad Sci USA*, 104:6424, 2007.
 - [11] M Riesenhuber and T Poggio. Hierarchical models of object recognition in cortex. *Nat Neurosci*, 2:1019, 1999.
 - [12] M Riesenhuber and T Poggio. Models of object recognition. *Nat Neurosci*, 3:1199, 2000.
 - [13] Y LeCun, F Huang, and L Bottou. Learning methods for generic object recognition with invariance to pose and lighting. In *Computer Vision and Pattern Recognition, Proc 2004 IEEE Comp Soc Conf*, volume 2, page II, 2004.
 - [14] D Stettler, A Das, J Bennett, and C Gilbert. Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron*, 36:739, 2002.

- [15] A Angelucci, J Levitt, E Walton, J Hupe, J Bullier, and J Lund. Circuits for local and global signal integration in primary visual cortex. *J Neurosci*, 22(19), 2002.
- [16] M Colonnier, J O’Kusky, et al. [number of neurons and synapses in the visual cortex of different species]. *Revue canadienne de biologie/editee par l’Universite de Montreal*, 40(1), 1981.
- [17] D Field, A Hayes, and R Hess. Contour integration by the human visual system: Evidence for a local "association field". *Vision Res*, 33:173, 1993.
- [18] P Parent and S Zucker. Trace inference, curvature consistency, and curve detection. *Pattern Analysis and Machine Intelligence, IEEE Trans*, 11:823, 1989.
- [19] V Gintautas, M Ham, B Kunsberg, et al. Model cortical association fields account for the time course and dependence on target complexity of human contour perception. *PLoS Comp Biol*, 7:e1002162, 2011.
- [20] J Zweck and L Williams. Euclidean group invariant computation of stochastic completion fields using shiftable-twistable functions. *J Math Imaging Vision*, 21(2):135–154, 2004.
- [21] P Bressloff and J Cowan. The functional geometry of local and horizontal connections in a model of v1. *J Physiol*, 97(2):221–236, 2003.
- [22] F Wolf and T Geisel. Spontaneous pinwheel annihilation during visual development. *Nature*, 395:73, 1998.
- [23] F Wolf and T Geisel. Universality in visual cortical pattern formation. *J Physiol*, 97(2):253–264, 2003.
- [24] G DeAngelis, J Robson, I Ohzawa, and R Freeman. Organization of suppression in receptive fields of neurons in cat visual cortex. *J Neurophysiol*, 68:144, 1992.
- [25] P de Gennes, J Prost, and R Pelcovits. The physics of liquid crystals. *Phys Today*, 48:70, 1995.
- [26] C Gilbert and T Wiesel. Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *J Neurosci*, 9:2432, 1989.
- [27] I Kovacs and B Julesz. A closed curve is much more than an incomplete one: Effect of closure in figure-ground segmentation. *Proc Natl Acad Sci USA*, 90:7495, 1993.
- [28] L Jeffress et al. A place theory of sound localization. *J Comp Physiol Psychol*, 41:35, 1948.
- [29] P Dayan and L Abbott. *Theoretical Neuroscience*. MIT Press, 2005.
- [30] T Miconi and R VanRullen. The gamma slideshow: object-based perceptual cycles in a model of the visual cortex. *Front Hum Neurosci*, 4:205, 2010.
- [31] C Altmann, H Bülthoff, and Z Kourtzi. Perceptual organization of local elements into global shapes in the human visual cortex. *Current Biology*, 13(4):342–349, 2003.