Building a mechanistic model of the development and function of the primary visual cortex

James A. Bednar
Institute for Adaptive and Neural Computation,
The University of Edinburgh
10 Crichton St, EH8 9AB, Edinburgh, UK

Abstract

Researchers have used a very wide range of different experimental and theoretical approaches to help understand mammalian visual systems. These approaches tend to have quite different assumptions, strengths, and weaknesses. Computational models of the visual cortex, in particular, have typically implemented either a proposed circuit for part of the visual cortex of the adult, assuming a very specific wiring pattern based on findings from adults, or else attempted to explain the long-term development of a visual cortex region from an initially undifferentiated starting point. Previous models of adult V1 have been able to account for many of the measured properties of V1 neurons, while not explaining how these properties arise or why neurons have those properties in particular. Previous developmental models have been able to reproduce the overall organization of specific feature maps in V1, such as orientation maps, but are generally formulated at an abstract level that does not allow testing with real images or analysis of detailed neural properties relevant for visual function. In this review of results from a large set of new, integrative models developed from shared principles and a set of shared software components, I show how these models now represent a single, consistent explanation for a wide body of experimental evidence, and form a compact hypothesis for much of the development and behavior of neurons in the visual cortex. The models are the first developmental models with wiring consistent with V1, the first to have realistic behavior with respect to visual contrast, and the first to include all of the demonstrated visual feature dimensions. The goal is to have a comprehensive explanation for why V1 is wired as it is in the adult, and how that circuitry leads to the observed behavior of the neurons during visual tasks.

1 Introduction

Understanding how we see remains an elusive goal, despite more than half a century of intensive work using a wide array of experimental and theoretical techniques. Because each of these techniques has different assumptions, strengths, and weaknesses, it can be difficult to establish clear principles and conclusive evidence. To make significant progress in this area, it is important to consider how the existing data can be synthesized into a coherent explanation for a wide variety of phenomena.

Computational models of the primary visual cortex (V1) could provide a platform for achieving such a synthesis, integrating results across levels to provide an overall explanation for the main body of results. However, existing models typically fall into one of two categories with different aims, neither of which achieves this goal: (1) narrowly constrained models of specific aspects of adult cortical circuitry or function, or (2) abstract models of large-scale visual area development, accounting for only a few of the response properties of individual neurons within these areas.

Existing models of type (1) (e.g. [1, 2, 23]) have been able to show how a variety of specialized circuits (often mutually incompatible) can account for most of the major observed functional properties of V1 neurons, but do not attempt to show how a single, general-purpose circuit could explain most of them at the same time. Because each specific phenomenon can often be explained by many different specialized models, it can be difficult or impossible to distinguish between different explanations. Moreover, just showing an example of how the property can be implemented does little to explain why neurons are arranged in this way, and what the circuit might contribute to the process of vision.

Similarly, many existing models of type (2) have been able to account for the large-scale organization of V1, such as its arrangement into topographic orientation maps (reviewed in refs. [30, 35, 75]). Yet because the developmental models are formulated at an abstract level, they address only a few of the observed properties of V1 neurons (such as their orientation or eye preference), and it is again difficult to decide between the various explanations. Thus despite the many thousands of experimental and computational papers about the visual cortex, methods for integrating, interpreting, and evaluating the overall body of evidence to build a coherent explanation remain frustratingly scarce.

This paper outlines and reviews a large set of closely interrelated computational models of the visual cortex that together are beginning to form a consistent, biologically grounded, computationally simple explanation of the bulk of V1 development and function. Specifically, the models develop:

1. Neurons with receptive fields (RFs) selective for retinotopy (X,Y), orientation (OR), ocular dominance
2. Preferences for each of these organized into realistic spatially topographic maps

3. Lateral connections between these neurons that reflect the structure of the maps

4. Realistic surround modulation effects, including their diversity, caused by interactions between these neurons

5. Contrast-gain control and contrast-invariant tuning for the individual neurons, ensuring that they retain selectivity robustly

6. Both simple and complex cells, to account for the major response types of V1 neurons

7. Long-term and short-term plasticity (e.g. aftereffects), emerging from mechanisms originally implemented for development

Together, these phenomena arguably represent the bulk of the generally agreed stimulus-driven response properties of V1 neurons. Accounting for such a diverse set of phenomena could have required an extremely complex model, e.g. the union of the many previously proposed models for each of the individual phenomena. Yet our results show that it is possible to account for all of these using only a small set of plausible principles and mechanisms, within a consistent biologically grounded framework:

1. Single-compartment (point) firing-rate (non-spiking) RGC, LGN, and V1 neurons

2. Hardwired subcortical pathways to V1 including the main LGN (lateral geniculate nucleus) or RGC (retinal ganglion cell) types that have been identified

3. Initially isotropic, topographic connectivity within and between neurons in layers in V1

4. Natural images and spontaneous activity patterns that lead to V1 responses

5. Hebbian learning with normalization for V1 neurons

6. A large number of parameters associated with each of these mechanisms

Properties not necessary to explain the phenomena above, such as spiking and detailed neuronal morphology, have been omitted, to clearly focus on the most relevant aspects of the system. The overall hypothesis is that much of the complex structure and properties observed in the visual cortex emerges from interactions between relatively simple but highly interconnected computing elements, with connection strengths and patterns self-organizing in response to visual input and other sources of neural activity. Through visual experience, the geometry and statistical regularities of the visual world become encoded into the structure and connectivity of the visual cortex, leading to a complex functional cortical architecture that reflects the physical and statistical properties of the visual world.

At present, many of the results have been obtained independently in a wide variety of separate projects performed by different collaborators at different times. However, all of the models share the same underlying principles outlined above, and all are implemented using the same simulator and a small number of underlying components. This review shows how each of these modelling studies, previously reported separately, fits into a consistent and compact framework for explaining a very wide range of data. The models are the first developmental models of V1 maps with wiring consistent with V1, and the first to have realistic behavior with respect to visual contrast, and together they account for all of the various spatial feature dimensions for which topographic maps have been reported using imaging in mammals.

Preliminary work into developing an implementation combining each of the models into a single, working model visual system is also reported, although doing so is still a long-term work in progress. The unified model will include all of the visual feature dimensions, as well as all of the major sources of connectivity that affect V1 neuron responses. The goal is to have the first comprehensive, mechanistic explanation for why V1 becomes wired as it is in the adult, and how that circuitry leads to the observed behavior of the neurons during visual tasks. That is, the model will be the first that starts from an initially undifferentiated state, to wire itself into a collection of neurons that behave, at a first approximation, like those in V1. Because such a model starts with no specializations (at the cortical level) specific to vision and would organize very differently when given different inputs, it would also represent a general explanation for the development and function of sensory and motor areas throughout the cortex.

2 Material and Methods

All of the models whose results are presented here are implemented in the Topographica simulator, and are freely available along with the simulator from
www.topographica.org. This section describes the complete, unified architecture under development, with each currently implemented model representing a subset or simplification of this architecture (as specified for each set of results below). The proposed model is a generalization of the GCAL (gain-controlled, adaptive, laterally connected) map model [46], extended to cover results from a large family of related models. The unified GCAL model is intended to represent the visual system of the macaque monkey, but relies on data from studies of cats, ferrets, tree shrews, or other mammalian species where clear results are not yet available from monkeys.

Sheets and projections

Each Topographica model consists of a set of sheets of neurons and projections (sets of topographically mapped connections) between them. For each model discussed in this paper, there are sheets representing the visual input (as a set of activations in photoreceptor cells), the transformation from the photoreceptors to inputs driving V1 (expressed as a set of LGN cell activations), and neurons in V1. Figure 1 shows the sheets and connections in the unified GCAL model, which is described for the first time here as a single, combined model.

Each sheet is implemented as a two-dimensional array of firing-rate neurons. The Topographica simulator allows parameters for sheets and projections to be specified in measurement units that are independent of the specific grid sizes used in a particular run of the simulation. To achieve this, Topographica sheets provide multiple spatial coordinate systems, called sheet and matrix coordinates. Where possible, parameters (e.g. sheet dimensions or connection radii) are expressed in sheet coordinates, expressed as if the sheet were a continuous neural field rather than a finite grid. In practice, of course, sheets are always implemented using some finite matrix of units. Each sheet has a parameter called its density, which specifies how many units (matrix elements) in the matrix correspond to a length of 1.0 in continuous sheet coordinates, which allows conversion between sheet and matrix coordinates.

In all simulations shown, sheets of V1 neurons have dimensions (in sheet coordinates) 1.0×1.0. If every other sheet in the model were to have a 1.0×1.0 area, units near the border of higher-level sheets like V1 would have afferent connections that extend past the border of lower-level sheets like the RGC/LGN cells. This cropping of connections will result in artifacts in the behavior of units near the border. To avoid such artifacts, lower-level sheets have areas larger than 1.0×1.0. (Alternatively, one could avoid cropping by imposing periodic boundary conditions, but doing so would create further artifacts by combining unrelated portions of the visual field into the connection fields of V1 neurons.) In figure 1 each sheet is plotted at the same scale in terms of degrees of visual angle covered, and thus the photoreceptor and RGC/LGN sheets appear larger. Sheet dimensions were chosen to ensure that each unit in the receiving sheet has a complete set of connections, where possible, minimizing edge effects in the RGC/LGN and V1 [49]. When sizes are scaled appropriately [10], results are independent of the density used, except at very low densities or for simulations with complex cells, where complexity increases with density (as described below). Larger areas can be simulated easily [10], but require more memory and simulation time.

A projection to an m × m sheet of neurons consists of m² separate connection fields, one per target neuron, each of which is a spatially localized set of connections from neurons in an input sheet near the corresponding topographic location of the target neuron. Figure 1 shows one sample connection field (CF) for each projection, visualized as an oval of the corresponding radius on the input sheet (drawn to scale), connected by a cone to the neuron on the target sheet. The connections and their weights determine the specific properties of each neuron in the network, by differentially weighting RGC/LGN inputs of different types and/or locations. Each of the specific types of sheets and projections is described in the following sections.

Images and photoreceptor sheets

The unified GCAL model contains six input sheets, representing the Long, Medium, and Short (LMS) wavelength cone photoreceptors in the retinas of the left and right eyes (illustrated in figure 1). The density of photoreceptors is uniform across the sheets, because only a relatively small portion of the visual field is being modeled, but the non-uniform spacing from fovea to periphery can be added for a model with a larger input sheet. Given a color image appearing in one eye, the estimated activations of the LMS sheets are calculated from the cone sensitivity functions for each photoreceptor type [73] [74] using calibrated color images [55], following the method described in ref. [27]. Input image pairs (left, right) were generated by choosing one image randomly from a database of single calibrated images, selecting a random patch within the image, a random nearly horizontal offset between patterns in each eye (as described in ref. [61]), a random direction of motion translation with a fixed speed (described in ref. [12]), and a random brightness difference between the two eyes (described in ref. [49]). These modifications are intended as a simple model of motion and eye differences, to allow development of direction preference, ocular dominance, and disparity maps, until suitable full-motion stereo calibrated-color video datasets of natural scenes are available. Simulated retinal waves can also be used as inputs, to provide initial RF and map structure before eye opening, but are not required for RF or map development in the model [13].
**Subcortical sheets**

The subcortical pathway from the photoreceptors to the thalamorecipient cells in V1 is represented as a set of hardwired subcortical cells with fixed connection fields (CFs) that determine the response properties of each cell. These cells represent the complete processing pathway to V1, including circuitry in the retina (including the retinal ganglion cells), optic nerve, lateral geniculate nucleus, and optic radiations to V1. Because the focus of the model is to explain cortical development given its thalamic input, the properties of these RGC/LGN cells are kept fixed throughout development, for simplicity and clarity of analysis.

Each distinct RGC/LGN cell type is grouped into a separate sheet, each of which contains a topographically organized set of cells with identical properties but responding to a different region of the retinal photoreceptor input sheet. Figure 1 shows examples of each of the different response types suitable for the development of the full range of V1 RFs: SF1 (achromatic cells with large receptive fields, as in the magnocellular pathway), SF2 (achromatic cells with small receptive fields), BY cells (color opponent cells with blue (short) cone photoreceptor center, and red (long) and green (short) surround), and similarly for medium-center (GR) and long-center (RG) chromatic L/M RFs. Each such opponent cell comes in two types, On (with an excitatory center) and Off (with an excitatory surround).

All of these cells have Difference-of-Gaussian RFs, and thus perform edge enhancement at a particular size scale (SF1 and SF2) and/or color enhancement (e.g. RG and GR). Each of these RF types has been reported in the macaque...
retina or LGN [26, 43, 79], and together they cover the range of typically reported spatial response types. Additional such cell classes can easily be added as needed, e.g. to provide more than two sizes of RGC/LGN cells with different spatial frequency preferences [58], although doing so increases memory and computation requirements. Not all of the cell types currently included are necessarily required for these results, but so far they have been found to be sufficient.

For each of the RGC/LGN sheets, multiple projections with different delays connect them to the V1 sheet. These delays represent the different latencies in the lagged vs. non-lagged cells found in cat LGN [66, 86], and allow V1 neurons to become selective for the direction of motion. Lagged cells could instead be implemented using separate LGN sheets, as in ref [12], if differences from non-lagged cells other than temporal delay need to be taken into account. Many other sources of temporal delays would also lead to direction preferences, but have not been tested specifically.

Apart from these delays, the detailed temporal properties of the subcortical neural responses and of signal propagation along the various types of connections elsewhere in the network have not been modelled. Instead, the model RGC/LGN neurons have a constant, sustained output, and all connections in each projection have a constant delay, independent of the physical length of that connection. Modelling the subcortical temporal response properties and simulating non-uniform delays would greatly increase the number of timesteps needed to simulate each input presentation, but should otherwise be feasible in future work.

Cortical sheets

Many of the simulations use only a single V1 sheet for simplicity, but in the full unified model, V1 is represented by three cortical sheets. First, cells with direct thalamic input are labeled V1 L4 in figure 1 and nominally correspond to pyramidal simple cells in macaque V1 layer 4Cβ. Second, pyramidal cells in layer 2/3 (labeled V1 L2/3E) receive input from a small topographically corresponding (columnar) region of V1 L4. These cells (as discussed below) become complex-cell–like, i.e. relatively insensitive to spatial phase, by pooling across nearby L4 simple cells of different preferred spatial phases [5]. The third V1 sheet L2/3I models inhibitory interneurons in layer 2/3. Together, these sheets will be used to show how simple cells can develop in V1 L4, how complex cells can develop in L2/3, and how interactions between these cells can vary depending on contrast, which affects the balance between excitation and inhibition [4].

The behavior of the V1 sheets is primarily determined by the projections to, within, and between them. Figure 2 illustrates and describes each of these projections. The overall pattern of projections was chosen based on anatomical tracing of the V1 microcircuit in cat [18], providing the first model of V1 self-organization where the long-range connections are excitatory [4, 45], as found in animals [34]. Each of these projections is initially non-specific, and becomes selective only through the process of self-organization (described below), which increases some connection weights at the expense of others.

Activation

At each training iteration, a new retinal input image is presented and the activation of each unit in each sheet is updated in a series of steps. One training iteration represents one visual fixation (for natural images) or a snapshot of the relatively slowly changing spatial pattern of spontaneous activity (e.g. for retinal waves [87]). I.e., an iteration consists of a constant retinal activation, followed by recurrent processing at the LGN and cortical levels. For one iteration, assume that input is drawn on the photoreceptors at time $t_0$.
and the connection delay (constant for all projections) is defined as $\delta t = 0.05$ (roughly corresponding to 10-20 milliseconds). Then at $t + 0.05$ the RGC/LGN cells compute their responses, and at $t + 10$ the thalamic output is delivered to V1, where it similarly propagates through the cortical sheets.

Images are presented to the model by activating the retinal photoreceptor units. The activation value $\Psi_{i,p}$ of unit $i$ in photoreceptor sheet $P$ is given by the calibrated-color estimate of the L, M, or S cone activation in the chosen image at that point.

For each model neuron in the other sheets, the activation value is computed based on the combined activity contributions to that neuron from each of the sheet’s incoming projections. The activity contribution from a projection is recalculated whenever its input sheet activity changes, after the corresponding connection delay. For unit $j$ in the target sheet, the activity contribution $C_{jp}$ to $j$ from projection $p$ is a dot product of the relevant input with the weights:

$$C_{jp}(t + \delta t) = \sum_{i \in F_{j,p}} X_{is_p}(t)\omega_{ij,p}$$

(1)

where $X_{is}$ is the activation of unit $i$ on this projection’s input sheet $s_p$, taken from the set of all input neurons from which target unit $j$ receives connections in that projection (its connection field $F_{j,p}$), and $\omega_{ij,p}$ is the connection weight from $i$ to $j$ in that projection. Across all projections, multiple direct connections between the same pair of neurons are possible, but each projection $p$ contains at most one connection between $i$ and $j$, denoted by $\omega_{ij,p}$.

For a given subcortical or cortical unit $j$ in the separate models reported in this paper (except those containing complex cells), the activity $\eta_j(t + \delta t)$ is calculated from a rectified weighted sum of the activity contributions $\eta_{jp}(t + \delta t)$:

$$\eta_j(t) = f\left(\sum_p \gamma_p C_{jp}(t)\right)$$

(2)

$f$ is a half-wave rectifying function with a variable threshold point $\theta$ dependent on the average activity of the unit, as described in the next subsection. Each $\gamma_p$ is an arbitrary multiplier for the overall strength of connections in projection $p$. The $\gamma_p$ values are set in the approximate range 0.5 to 3.0 for excitatory projections and -0.5 to -3.0 for inhibitory projections. For afferent connections, the $\gamma_p$ value is chosen to map average V1 activation levels into the range 0 to 1.0 by convention, for ease of interconnecting and analyzing multiple sheets. For lateral and feedback connections, the $\gamma_p$ values are then chosen to provide a balance between feedforward, lateral, and feedback drive, and between excitation and inhibition; these parameters are crucial for making the network operate in a useful regime.

For the full unified model, RGC/LGN neuron activity is computed similarly to equation (2), except that they have a separate divisive lateral inhibitory projection:

$$\eta_{jp} = f\left(\sum_p \gamma_p C_{jp}(t) - \gamma_p C_{jp}(t) + k\right)$$

(3)

where $L$ stands for one of the RGC/LGN sheets. Projection $S$ consists of a set of isotropic Gaussian-shaped lateral inhibitory connections (see equation [7], evaluated with $u = 1$), and $p$ ranges over all the other projections to that sheet. $k$ is a small constant to make the output well-defined for weak inputs. The divisive inhibition implements the contrast gain control mechanisms found in RGC and LGN neurons [3, 4, 20, 32].

For the unified model and individual models with complex cells, cortical neuron activity is computed similarly to equation (2), except to add firing-rate–fluctuation noise and exponential smoothing of the recurrent dynamics:

$$\eta_{jV}(t + \delta t) = \lambda f\left(\sum_p \gamma_p C_{jp}(t + \delta t)\right) + (1 - \lambda)\eta_{jV}(t) + \sigma_n x$$

(4)

where $V$ stands for one of the cortical sheets, $p$ ranges over all projections to that sheet, $\lambda = 0.5$ is a time constant parameter that defines the strength of smoothing of the recurrent dynamics in the network, $x$ is a normally distributed zero-mean unit-variance random variable, and $\sigma_n$ scales $x$ to determine the amount of noise. The smoothing ensures that the system remains numerically stable, without spurious oscillations caused by simulating only discrete time steps, for the relatively coarse time steps that are used here for computational efficiency.

For any of the models, each time the activity is computed using equation (2), (3), or (4), the new activity values are sent to each of the outgoing projections, where they arrive after the projection delay (typically 0.05). The process of activity computation then begins again, with a new contribution $C$ computed as in equation (1) leading to new activation values by equation (2), (3), or (4). Activity thus spreads recurrently throughout the network, and can change, die out, or be strengthened, depending on the parameters.

With typical parameters that lead to realistic topographic map patterns, initial blurry patterns of afferent-driven activity are sharpened into well-defined “activity bubbles” through locally cooperative and more distantly competitive lateral interactions [49]. Nearby neurons are thus influenced to respond more similarly, while more distant neurons receive net inhibition and thus learn to respond to different input patterns. The competitive interactions “sparsefy” the cortical response into patches, in a process that can be compared to the explicit sparseness constraints in non-mechanistic models [38, 56], while the local facilitatory interactions encourage spatial locality so that smooth topographic maps will be developed.
Whenever the time reaches an integer multiple (e.g. 1.0 or 2.0), the V1 response is used to update the threshold point \( \theta \) of V1 neurons (using the adaptation process described in the next section) and to update the afferent weights via Hebbian learning (as described in the following section). Both adaptation and learning could also be performed at each settling step, but doing so would greatly decrease computational efficiency. Because the settling (sparsification) process typically leaves only small patches of the cortical neurons responding strongly, those neurons will be the ones that learn the current input pattern, while other nearby neurons will learn other input patterns, eventually covering the complete range of typical input variation. Overall, through a combination of the network dynamics that achieve sparsification along with local similarity, plus Hebbian learning that leads to feature dynamics that achieve sparsification along with local variation. Overall, through a combination of the network dynamics that achieve sparsification along with local similarity, plus Hebbian learning that leads to feature dynamics that achieve sparsification along with local variation.

**Homeostatic adaptation**

For this model, the threshold for activation of each neuron is a very important quantity, because it directly determines how much the neuron will fire in response to a given input. To set the threshold, each neuron unit \( j \) in V1 calculates a smoothed exponential average of its activity \( \bar{\eta}_j \):

\[
\bar{\eta}_j(t) = (1 - \beta)\eta_j(t) + \beta\bar{\eta}_j(t - 1) \tag{5}
\]

The smoothing parameter \( \beta = 0.999 \) determines the degree of smoothing in the calculation of the average. \( \bar{\eta}_j \) is initialized to the target average V1 unit activity \( \mu \), which for all simulations is \( \bar{\eta}_A(0) = \mu = 0.024 \). The threshold is updated as follows:

\[
\theta(t) = \theta(t - 1) + \lambda(\bar{\eta}_j(t) - \mu) \tag{6}
\]

where \( \lambda = 0.0001 \) is the homeostatic learning rate. The effect of this scaling mechanism is to bring the average activity of each V1 unit closer to the specified target. If the activity in a V1 unit moves away from the target during training, the threshold for activation is thus automatically raised or lowered in order to bring it closer to the target. Note that an alternative rule with only a single smoothing parameter (rather than \( \beta \) and \( \lambda \)) could be formulated, but the rule as presented here makes it simple for the modeler to set a desired target activity \( \mu \), and is in any case relatively insensitive to the values of the smoothing parameters.

**Learning**

Initial connection field weights are random within a two-dimensional Gaussian envelope. E.g., for a postsynaptic (target) neuron \( j \) located at sheet coordinate \((0,0)\), the weight \( \omega_{ij,p} \) from presynaptic unit \( i \) in projection \( p \) is:

\[
\omega_{ij,p} = \frac{1}{Z_p}u \exp\left(-\frac{x^2 + y^2}{2\sigma_p^2}\right) \tag{7}
\]

where \((x, y)\) is the sheet-coordinate location of the presynaptic neuron \( i \), \( u \) is a scalar value drawn from a uniform random distribution for the afferent and lateral inhibitory projections \((p = A, I)\), \( \sigma_p \) determines the width of the Gaussian in sheet coordinates, and \( Z_p \) is a constant normalizing term that ensures that the total of all weights \( \omega_{ij} \) to neuron \( j \) in projection \( p \) is 1.0. Weights for each projection are only defined within a specific maximum circular radius \( r_p \); they are considered zero outside that radius.

In every iteration, each connection weight \( \omega_{ij} \) from unit \( i \) to unit \( j \) is adjusted using a simple Hebbian learning rule. This rule results in connections that reflect correlations between the presynaptic activity and the postsynaptic response. Hebbian connection weight adjustment for unit \( j \) is dependent on the presynaptic activity \( \eta_i \), the post-synaptic response \( \eta_j \), and the Hebbian learning rate \( \alpha \):

\[
\omega_{ij,p}(t) = \frac{\omega_{ij,p}(t - 1) + \alpha\eta_i\eta_j}{\sum_k (\omega_{kj,p}(t - 1) + \alpha\eta_j\eta_k)} \tag{8}
\]

Unless it is constrained, Hebbian learning will lead to ever-increasing (and thus unstable) values of the weights. The weights are constrained using divisive post-synaptic weight normalization (equation 8), which is a simple and well understood mechanism. All afferent connection weights from RGC/LGN sheets are normalized together in the model, which allows V1 neurons to become selective for any subset of the RGC/LGN inputs. Weights are normalized separately for each of the other projections, to ensure that Hebbian learning does not disrupt the balance between feedforward drive, lateral and feedback excitation, and lateral and feedback inhibition. Subtractive normalization with upper and lower bounds could be used instead, but it would lead to binary weights [50, 51], which is not desirable for a firing-rate model whose connections represent averages over multiple physical connections. More biologically motivated homeostatic mechanisms for normalization such as multiplicative synaptic scaling [78] or a sliding threshold for plasticity [17] could be implemented instead, but these have not been tested so far.

**3 Results**

In each of the subsections below, results are shown from previously reported simulations using partial or simplified versions of the full unified GCAL model described above. Models were typically run for 10,000 iterations (where an iteration corresponds to one complete image presentation, e.g. a visual saccade), with structured feature maps and receptive fields gradually appearing as more input patterns...
are presented. By 10,000, the maps and connection fields have come to a “dynamic equilibrium”, where they are stable as long as the input statistics are stationary [49], and these are the results that are shown here.

### 3.1 Feature maps

For each of the topographic feature maps reported for V1 in imaging experiments, figure 3 shows a typical imaging result from an animal, plotted above a typical final result from a simulation using the above framework and including that feature value. For OR, DR, DY, and CR, black indicates neurons not selective for the indicated feature, and bright colors indicate highly selective neurons. For CR the color in the plot indicates the preferred color; the others are false-color plots showing the feature value for each neuron using the color key adjacent to that panel. Maps were measured by collating responses to sine gratings covering ranges of values of each feature [49], reproducing typical methods of map measurement in animals [19]. Each of these specific simulations was done with an earlier model named LISSOM [49], which is similar to the full GCAL model described above but simplified to include only a single cortical layer, with direct inhibitory long-range lateral connections, and with modeller-determined thresholds rather than the homeostatic mechanisms described above (equation 6). Each result is for a model including only a subset of the subcortical sheets shown in figure 1. Specifically, the model (X,Y) and OR map simulations use only a single pair of monochromatic On, Off RGC/LGN sheets at a fixed size, the OD and DY maps are similar to OR but include an additional pair for the other eye, the DR and TF maps are similar to OR but include three additional pairs with different delays, the SF map is similar to OR but includes three additional pairs of On, Off cells with different Difference-Of-Gaussians CFs, and the CR map is similar to OR but includes five additional sheets of color opponent cells (BY On, RG On, RG Off, GR On, and GR Off).

As described in the indicated original source for each model, the model results for (X,Y), OR, OD, DR, and SF have been evaluated against the available animal data, and capture the main aspects of the feature value coverage and the spatial organization of the maps [49,58]. For instance, the OR and DR maps show iso-feature domains, pinwheel centers, fractures, saddle points, linear zones, and an overall ring-shaped Fourier transform, as in the animal maps [19,82]. The maps simulated together (e.g. OR and OD) also tend to intersect at right angles, such that high-gradient regions in one map avoid high-gradient regions in others [49].

These patterns primarily emerge from geometric constraints on smoothly mapping the range of values for the indicated feature, within a two-dimensional retinotopic map [49]. They are also affected by the relative amount by which each feature varies in the input dataset, how often each feature appears, and other aspects of the input image statistics [49]. For instance, orientation maps trained on natural image inputs develop a preponderance of neurons with horizontal and vertical orientation preferences, as seen in ferret maps and in natural images [13,25].

For DY, the model results [61] have not been compared systematically with the small amount of data (barely visible in the plot) from the one available experimental report on the organization for disparity preferences [42], because the model predated the experiments by several years. A preliminary analysis suggests that the model organization is comparable in that neurons for disparity tend to occur in small, local regions sensitive to horizontal disparity, but further evaluation is necessary. The results for color (CR) are preliminary due to ongoing work on this topic, but do show that color-selective neurons are found in spatially segregated blobs that include multiple color preferences, as suggested by the imaging data currently available [88]. Results for TF are also preliminary, again because they predated experimental TF maps and have not yet been characterized against the experimental results.

Overall, where it has been possible to make comparisons, the separate models have been shown to reproduce the main features of the experimental data, using a small set of assumptions. In each case, the model demonstrates how the experimentally measured map can emerge from Hebbian learning of corresponding patterns of subcortical and cortical activity. The models thus illustrate how the same basic, general-purpose adaptive mechanism will lead to very different organizations, depending on the geometrical and statistical properties of that feature.

So far, only a subset of these features have been combined into a single simulation, such as (X,Y), OR, OD, DR, and TF [15]. Future work will focus on showing how all or nearly all of these results could emerge simultaneously in the full unified GCAL model. However, note that only a few of these maps have ever been measured in the same animal, or even the same species, and thus it is not yet known whether all such maps are actually present in any single animal. The unified model can be used to understand how the maps interact, and to make predictions on the expected organization for maps not yet measured in a particular species.

### 3.2 Connection patterns

The feature maps described in the previous subsection are summaries of the properties of a set of neurons embedded in a network. To understand how these properties come about, it is necessary to look at the patterns of connectivity that underlie them. Figure 4 shows examples of such connections from an (X,Y) and OR GCAL simulation using a single pair of On, Off LGN/RGC sheets [4]. Note that this model focuses only on the emergence of orientation preferences, not on other dimensions such as spatial frequency, which
Figure 3: Simulated vs. real animal V1 maps. Imaging results for 4mm×4mm of V1 of the indicated species and from the corresponding LISSOM models of retinotopy (X,Y), orientation (OR), ocular dominance (OD), motion direction (DR), spatial frequency (SF), temporal frequency (TF), disparity (DY), and color (CR). Reprinted from references indicated; see main text for description.
Figure 4: Self-organized projections to V1 L2/3. Results from a GCAL model orientation map with separate V1 L4 and L2/3 regions allowing the emergence of complex cells; other dimensions like OD, DR, DY, SF, and CR are not included here. (a,b) Connection fields from the LGN On and Off channels to every 20th neuron in the model L4 show that the neurons develop OR preferences that cover the full range at each retinotopic location. (c) Long-range excitatory lateral connections to those neurons preferentially come from neurons with similar OR preferences. Here strong weights are colored with the OR preference of the source neuron. Strong weights occur in clumps (appearing as small dots here) corresponding to an iso-orientation domain (each approximately 0.2–0.3mm wide); the fact that most of the dots are similar in color for any given neuron shows that the connections are orientation specific. (d) Enlarged plot from (c) for a typical OR domain neuron that prefers horizontal patterns and receives connections primarily from other horizontal-prefering neurons (appearing as blobs of red or nearly red colors). (e) OR pinwheel neurons receive connections from neurons with many different OR preferences. Reprinted from ref. [4].
would require additional sheets as shown in figure 1. The afferent connections that develop reflect the feature preference of the target neuron, with elongated Gabor-shaped connection fields that respond to oriented edges in the input. Only a single orientation edge size is represented, because this simulation includes only a single RGC/LGN cell RF size. The orientation preference suggested by the afferent weight pattern strongly correlates with the measured orientation map (ref. [4]; not shown).

Because the lateral weights, like the afferent weights, are modified by Hebbian learning, they reflect the correlation patterns between V1 neurons, which are determined both by the emerging map patterns and by the input statistics. For this OR map, retinotopy and OR strongly determine the correlations, and thus the lateral connection patterns respect both the retinotopic and the orientation maps. Lateral connection patterns are thus patchy and orientation specific, as seen in tree shrews and monkeys [22, 71] (see figure 5). For neurons in orientation domains, long-range connections primarily come from neurons with similar orientation preference, because those neurons were often coactivated with this neuron during self-organization using natural images. Interestingly, a prediction is that cells near pinwheel fractures, which are less selective for orientation in the model, will have a much broader range of input connections, because their activity is correlated with neurons with a wide, non-specific range of orientation preferences. The degree of orientation selectivity of neurons near pinwheel centers has been controversial, but current evidence is in line with the model results [53]. Connectivity patterns in pinwheels have not yet been investigated experimentally, and so the model results represent predictions for future experiments.

When multiple maps are simulated in the same model, the connection patterns respect all maps at once, because there is only one set of V1 neurons and one set of lateral connections, each determined by Hebbian learning. Figure 5 shows an example from a combined (X,Y), OR, OD, DR map simulation, which reproduces the observed connection dependence on the OR map but predicts that connections will also respect the DR map (as reported by [65]), and for highly monocular neurons will respect the OD map as well. The model strongly predicts that lateral connection patterns will respect all other maps that account for a significant fraction of the response variance of the neurons, because each of those features will thus affect the correlation between neurons.

### 3.3 Orientation and phase tuning

Typical map development models focus on reproducing the map patterns, and are otherwise highly abstract (for review see refs [30, 35, 75]). These models are very useful for understanding the process of map formation in isolation, and to explain the geometric properties of maps. However, a map pattern measured in a real animal is simply a summary of one property of a complete system that processes visual information, and so a full explanation of the map pattern requires a demonstration of how the map patterns emerge from a set of neurons that process visual information in a realistic way. For such a model, the map patterns can be a way to determine if the underlying model circuit is operating like those in animals, which is a very different goal than in abstract models of map patterns in isolation. Once such a model exhibits realistic maps, it can then be tested to determine whether the feature preferences summarized in the map actually represent realistic responses to a visual feature, such as the orientation and phase of a sine grating test pattern.

For instance, neurons in a model orientation map can be tested to see if they are actually selective for the orientation of an input stimulus, and retain that selectivity as contrast is varied (i.e., have robust contrast-invariant tuning [68]). Many developmental models (e.g. those based on the elastic net or using correlation-based learning) cannot be tested with a specific bitmap input image, and so cannot directly be extended into a model visual system of the type considered here. The results from others that do allow bitmap input are not typically compared with single-unit data from animals, again because the papers focus on the map patterns. Yet without contrast-invariant tuning, the response of a neuron would be ambiguous—a strong response would could indicate either that the preferred orientation is present, or else that the input simply has very high contrast.

Figure 6 shows that GCAL model neurons, thanks to the lateral inhibition implemented at the RGC/LGN level, do retain their orientation tuning as contrast is varied [5]. Earlier models such as LISSOM [49] did not have this property and were thus valid only for a small range of contrasts. In GCAL, lateral inhibition acts like divisive normalization of the population response, giving similar patterns and levels of activity regardless of the contrast, which leads to contrast-invariant tuning [5].

In the animal V1 data, tuning for spatial phase (the specific position of an oriented sine grating) is complicated, with simple cells highly selective for both orientation and spatial phase, and complex cells selective for orientation but not spatial phase [37]. Moreover, the spatial organization is smooth for orientation [19], such that nearby neurons have similar orientation preferences, but disordered for phase, with nearby neurons having a variety of phase preferences [6, 40, 47] (though the detailed organization for spatial phase is not yet clear and consistent between studies).

Disorder in the phase map provides a simple, local way to construct complex cells—simply pool outputs from several local simple cells, each with similar orientation preferences but one of a variety of phase preferences, to construct an orientation-selective cell relatively invariant to phase (as originally proposed by Hubel and Wiesel [37]). However, it is difficult to develop a random phase map in
Figure 5: **Lateral connections across maps.** LISSOM/GCAL neurons each participate in multiple functional maps, but have only a single set of lateral connections. Connections are strongest from other neurons with similar properties, respecting each of the maps to the degree to which that map affects correlation between neurons. Maps for a combined LISSOM OR/OD/DR simulation are shown above, with the black outlines indicating the connections to the central neuron (marked with a small black outline) that remain after weak connections have been pruned. Model neurons connect to other model neurons with similar orientation preference (a) (as in tree shrew, (d)) but even more strongly respect the direction map (c). This highly monocular unit also connects strongly to the same eye (b), but the more typical binocular cells have wider connection distributions. Reprinted from refs. [15, 22] as indicated.

Figure 6: **Contrast-invariant tuning.** Unlike other developmental models such as LISSOM, GCAL shows contrast-invariant tuning, i.e., similar orientation tuning width at different contrasts, as found in animals [68]. Reprinted from ref. [5].
simulation, because existing developmental models group neurons by similarity in responses to a visual pattern, while neurons with different phase preferences will by definition respond to different visual patterns [5]. Previous models for the development of random phase preferences have relied on unrealistic mechanisms such as squaring negative activation levels [38][81], which force opposite phases to become correlated (and thus grouped together by developmental models) but have no clear physical interpretation.

Instead, the GCAL models [4][5] show how random phase maps could arise in a set of simple cells (nominally in layer 4) through a small amount of initial disorder in the mapping from the thalamus, which persists because the model includes strong long-range lateral connectivity only in the layer 2/3 complex cells rather than in layer 4 (see figure 2). The layer 2/3 cells pool from a local patch of layer 4 cells, thus becoming unselective for phase, but have strong lateral interactions that lead to well-organized maps in layer 2/3 (which is primarily what is measured in most optical imaging experiments). Feedback from layer 2/3 to layer 4 then causes layer 4 cells to develop an orientation map, without disrupting the random phase map. Figure 7 plots the results of this process, showing the resulting maps and modulation ratios for each layer. The model predicts a strong spatial organization for (weak) phase preferences in layer 2/3, and that the orientation map in layer 4 (not currently measureable with imaging techniques due to its depth beneath the cortical surface) is less well ordered than in layer 2/3. Figure 6 shows that the complex cells in layer 2/3 retain orientation tuning and contrast invariance, as expected.

3.4 Dependence on input statistics

Because development in the model is driven by input patterns (whether externally or intrinsically generated), the results depend crucially on the specific patterns used. At the extreme, models with two eyes having identical inputs do not develop ocular dominance maps, and models trained only on static images do not develop direction maps [49]. The specific map patterns also reflect the input statistics, with more neurons responding to horizontal and vertical contours because of the prevalence of such contours in natural images [14]. Relationships between the maps also reflect the input statistics—direction maps dominate the overall organization when input patterns are all moving quickly [49], and ocular dominance maps interact orthogonally with orientation maps only for some types of eye-specific input differences [39]. Finally, the lateral connection patterns depend crucially on the input statistics, with long-range orientation specific connectivity developing only for input datasets with long-range orientation-specific correlations [49]. Even so, a very large range of possible input patterns suffices to develop map patterns—orientation maps (but not realistic lateral connection patterns) can develop from random noise inputs, abstract spatially localized patterns, retinal wave model patterns, and other two-dimensional patterns [49]. Thus the emergence of maps and RFs is a robust phenomenon, but the specific patterns of response properties and neural connectivity directly reflect the input scene statistics.

3.5 Surround modulation

Given a model with realistically patchy, specific lateral connectivity and realistic single-neuron properties, as outlined above, the patterns of interaction between neurons can be compared with neurophysiological evidence for surround modulation— influences on neural responses from distant patterns in the visual field. These studies can help validate the underlying model circuit, while helping understand how the visual cortex will respond to complicated patterns such as natural images.

For instance, as the size of a patch of grating is increased, the response of a V1 neuron typically increases at first, reaches a peak, and then decreases [67, 69, 80]. Similar patterns can be observed in a GCAL-based model orientation map with complex cells and separate inhibitory and excitatory subpopulations (figure 8 from ref. [4], see connectivity in figure 2). Small patterns initially activate neurons weakly, due to low overlap with the afferent receptive fields of layer 4 cells, but the response increases with larger patterns. For large enough patterns, lateral interactions are strong and in most locations net inhibitory, causing many neurons to be suppressed (leading to a subsequent dip in response). These patterns are visualized and compared with the experimental data in figure 9. The model demonstrates that the lateral interactions are sufficient to account for typical size tuning effects, and also accounts for less commonly reported effects that result from neurons with different specific self-organized patterns of lateral connectivity. The model thus accounts both for the typical pattern of size tuning, and explains why such a diversity of patterns is observed in animals.

The effects of the self-organized lateral connections are even more evident when orientation-specific surround modulation is considered. Because the connections respect the orientation map (figures 4 and 5), interactions change depending on the relative orientation of center and surround elements (figure 10). Moreover, because these patterns also vary depending on the location of the neuron in the map, a variety of patterns of interaction are seen, just as has been reported in experimental studies. Figure 11 illustrates some of these relationships, but many more such relationships are possible—any property of the neurons that varies systematically across the cortical surface will affect the pattern of interactions, as long as it changes the correlation between neurons during the history of visual experience. These results suggest both that lateral interactions may underlie many of the observed surround modulation effects, and also that the diversity of observed effects can at least in part be traced to the diversity of lateral connection patterns,
3.6 Aftereffects

The previous sections have focused on the network organization and operation after Hebbian learning can be considered to be completed. However, the visual system is continually adapting to the visual input even during normal visual experience, resulting in phenomena such as visual aftereffects [77]. To investigate whether and how this adaptation differs from long-term self-organization, we tested LISSOM and GCAL-based models with stimuli used in visual aftereffect experiments [11, 24]. Surprisingly, the same Hebbian equations that allow neurons and maps to develop selectivity also lead to realistic aftereffects, such as for orientation and color (figure 12). In the model, we assume that connections adapt during normal visual experience just as they do in simulated long-term development, albeit with a lower learning rate appropriate for adult vision. If so, neurons that are coactive during a particular visual stimulus (such as a vertical grating) will become slightly more strongly laterally connected as they adapt to that pattern. Subsequently, the response to that pattern will be reduced, due to increased lateral excitation that leads to net (disynaptic) lateral inhibition for high contrast patterns like those in the aftereffect studies. Assuming a population decoding model such as the vector sum [11], there will be no change in the perceived orientation of the adaptation pattern, but the perceived value of a nearby orientation will be repelled away from the adapting stimulus, because the neurons activated during adaptation now inhibit each other more strongly, shifting the population response. These changes are the direct result of Hebbian learning of intracortical connections, as can be shown by disabling learning for all other connections and observing no change in the overall behavior.
Size tuning. Responses from a particular neuron vary as the input pattern size increases. For a small sine grating (radius 0.2), responses are weak due to low overlap between the pattern and the neuron’s afferent connection field, even when the position and phase are optimal for that neuron (as here). An intermediate size (0.8) leads to a peak in response for this neuron, with the maximum ratio between excitation and inhibition. Larger sizes activate a larger area of V1, but responses are sparser due to the strong lateral inhibition recruited by this salient pattern. For this particular neuron, the response happens to be suppressed for radius 1.8, but note that many other neurons are still highly active, which is one reason that surround modulation properties are highly variable. The dashed line indicates the $1.0 \times 1.0$ area of the retina that is topographically mapped to the V1 layer 2/3 sheet. Reprinted from ref. [4].

Interestingly, for distant orientations, the human data suggests an attractive effect, with a perceived orientation shifted towards the adaptation orientation [52]. The model reproduces this feature as well, and provides the novel explanation that this indirect effect is due to the divisive normalization term in the Hebbian learning equation (equation 8). Specifically, when the neurons activated during adaptation increase their mutual inhibition, the normalization term forces this increase to come at the expense of connections to other neurons not (or only weakly) activated during adaptation. Those neurons are thus disinhibited, and can respond more strongly than before, shifting the response towards the adaptation stimulus.

Similar patterns occur for the McCollough Effect [48] (figure 12). Here the adaptation stimulus coactivates neurons selective for orientation, color, or both, and again the lateral interactions between all these neurons are strengthened. Subsequent stimuli then appear different in both color and orientation, in patterns similar to the human data. Interestingly, the McCollough effect can last for months, which suggests that the modelled changes in lateral connectivity can become essentially permanent, though the effects of short-term exposure typically fade in darkness or in subsequent visual experience.

Overall, the model suggests that the same process of Hebbian learning could explain both long-term development and short-term adaptation, unifying phenomena previously considered distinct. Of course, the biophysical mechanisms may indeed be distinct, potentially operating at different time scales and being largely temporary rather than the permanent changes found early in development. Even so, the results here suggest that both early development and adult short-term adaptation may operate using similar mathematical principles. How mechanisms for long-term and short-term plasticity may interact, including possible transitions from long-term to short term plasticity during so-called “critical periods”, is an important area for future modelling and experimental studies.

4 Discussion and future work
The results reviewed above illustrate a general approach to understanding the large-scale development, organization, and function of cortical areas. The models show that a relatively small number of basic and largely uncontroversial assumptions and principles may be sufficient to explain a very wide range of experimental results from the visual cortex. Even very simple neural units, i.e., firing-rate point...
Figure 9: **Diversity in size tuning.** Plots A-F show how the neural response varies as the sine grating radius increases, for 50% (blue) and 100% (red) contrasts, for six example neurons. Those on the top row are the most common type (37% of model neurons measured), and are a good match to the phenomena reported in experimental studies in macaque (G [67]) and cat (H [69], I [80]). Those in the middle row occur less often and are less commonly reported in experimental studies, such as larger responses to low contrasts at high radii, but examples of each such pattern can be found in the experimental results shown here. The model predicts that these variations in surround tuning properties are real, not just noise or an experimental artifact, and that they derive from the many possible interactions between a diverse set of neurons in the map.

neurons, generically connected into topographic maps with initially random or isotropic weights, can form a wide range of specific feature preferences and maps via unsupervised normalized Hebbian learning of natural images and spontaneous activity patterns. The resulting maps consist of neurons with realistic visual response properties, with variability due to visual context and recent history that explains significant aspects of surround modulation and visual aftereffects. The simulator and example simulations are freely downloadable from [topographica.org](http://topographica.org) (see figure 13), allowing any interested researcher to build on this work.

The long-term goal of this project is to understand whether and how a single, generic cortical circuit and developmental process can account for the major known information-processing properties of the visual cortex. If so, such a model would be a general-purpose model of cortical computation, with potentially many applications beyond computational neuroscience. Similar models have already been used for other cortical regions, such as rodent barrel cortex [84]. Combining the existing models into a single, runnable visual system is very much a work in progress, but the results so far suggest that doing so will be both feasible and valuable.

Once a full unified model is feasible, an important test will be to determine if it can replicate both the feature-based analyses described in most of the sections above, and also findings that the visual cortex acts as a unified map of spatiotemporal energy [9]. Although the model results are compatible with the feature-based view, the actual
Figure 10: **Orientation-contrast tuning curves (OCTCs).** Surround modulation changes as the orientation of a surrounding annulus is varied relative to a center sine-grating patch (as in the sample stimulus at the bottom right). In each graph A-F, red is the orientation tuning curve (as in figure 6) for the given neuron (with just the center grating patch), blue is for surround contrast 50%, and green is for surround contrast 100%. Top row: typically (51% of model neurons tested), a collinear surround is suppressive for these contrasts, but the surround becomes less suppressive as the surround orientation is varied (as for cat [69], G and macaque [41], H). Middle row: Other patterns seen in the model include high responses at diagonals (D, 20%, as seen in ref. [69]), strongest suppression not collinear (E, as seen in ref. [41]), and facilitation for all orientations (F, 5%). The relatively rare pattern in F has not been reported in existing studies, and thus constitutes a prediction. In each case the observed variability is a consequence of the model’s Hebbian learning that leads to a diversity of patterns of lateral connectivity, rather than noise or experimental artifacts.

Building realistic models of this kind depends on having a realistic model of a subcortical pathway. The approach illustrated in figure 1 allows arbitrarily detailed implementation of subcortical populations, and allows clear specification of model properties. However, it also results in a large number of subcortical neural sheets that can be difficult to simulate, and these also introduce a significant number of new parameters. An alternative approach is to model the retina as primarily driven by random wiring, an idea which has recently been verified to a large extent [33]. In this approach, a regular hexagonal grid of photoreceptors would be set up initially, and then a wide and continuous range of retinal ganglion cell types would be created by randomly connecting photoreceptors over local areas of the grid. This process would require some parameter tuning to ensure that the results cover the range of properties seen in animals, but could automatically generate realistic RGCs to drive V1 development.

At present, all of the models reviewed contain feedforward and lateral connections, but no feedback from higher cortical areas to V1 or from V1 to the LGN, because
Figure 11: Surround population effects. Some of the observed variance in surround modulation effects can be explained based on the properties of the measured model neuron or its neighbors. These measurements are based only on a relatively small number of model neurons, because each data point requires a several-hour-long computational experiment to choose the optimal stimuli for testing, but some trends are clear in the data so far. For instance, the modulation ratio and suppression index are somewhat correlated ($r = 0.226, p < 0.07$) – i.e., simple cells appear to be suppressed a bit more strongly (left). The amount of local homogeneity is a measure of how slowly the map is changing around a given neuron. Model neurons in homogeneous regions of the map exhibit lower orientation-contrast suppression ($r = -0.362, p < 0.01$; middle) but greater overall surround suppression ($r = 0.327, p < 0.05$). Although preliminary, these predictions can be tested in animal experiments. Moreover, these analyses suggest that much of the observed variance in surround modulation properties could be due to network and map effects like these, caused by Hebbian learning, with potentially many more possible interactions for neurons embedded in multiple overlaid functional maps.

Figure 12: Aftereffects as short-term self-organization. (a) While the fully organized network is repeatedly presented patterns with the same orientation, connection strengths are updated by Hebbian learning (as during development, but at a lower learning rate). The net effect is increased inhibition, which causes the neurons that responded during adaptation to respond less afterwards. When the overall response is summarized as a “perceived value” using a vector average, the result is systematic shifts in perception, such that a previously similar orientation will now seem very different in orientation, while more distant orientations will be unchanged or go in the opposite direction. These patterns are a close match to results from humans [52], suggesting that short-term and long-term adaptation share similar rules. (b) Similar explanations apply to the McCollough effect [48], an orientation-contingent color aftereffect; here the model predicts that lateral connections between orientation and color-selective neurons cause this effect [29] (and many others in other maps, such as motion aftereffects). (a) reprinted from ref. [11] and replotting data from ref. [52]; (b) reprinted from ref. [24] and replotting data from ref. [29].
Figure 13: Topographica simulator. This Topographica session shows a user analyzing a simple three-level model with two LGN/RGC sheets and one V1 sheet. (Clockwise from top) Displays for activity patterns, gradients, Fourier transforms, CFs for one neuron, the overall network, histograms of orientation preference, orientation maps, and projections (center) are shown. The simulator allows any number of sheets to be defined, interconnected, and analyzed easily, for any simulation defined as a network of interacting two-dimensional sheets of neurons.

such feedback has not been found necessary to replicate the features surveyed. However, note that nearly all of the physiological data considered was from anesthetized animals not engaged in any visually mediated behaviors. Under those conditions, it is not surprising that feedback would have relatively little effect. Corticocortical and corticothalamic feedback is likely to be crucial to explain how these circuits operate during natural vision [70] [76], and determining the form and function of this feedback is an important aspect of developing a general-purpose cortical model.

Because they focus on long-term development, the models discussed here implement time at a relatively coarse level. Most of the models update the retinal image only a few times each second, and thus are not suitable for studying the detailed time course of neural responses. A fully detailed model would need to include spiking, which makes most of the analyses presented here more difficult and much more time consuming, but it is possible to simulate even quite fine time scales at the level of a peri-stimulus-time histogram (PSTH). I.e., rather than simulating individual spike events, one can calibrate model neurons against a PSTH from a single neuron, thus matching its average temporal response over time, or against the average population response (e.g. measured by voltage-sensitive dye (VSD) imaging). Recent work shows that it is possible to replicate quite detailed transient-onset LGN and V1 PSTHs in GCAL using the
same neural architecture as in the models discussed here, simply by adding hysteresis to control how fast neurons can become excited and by simulating at a detailed 0.5-millisecond time step [72]. Importantly, the transient responses are a natural consequence of lateral connections in the model, and do not require a more complex model of each neuron. Future work will calibrate the model responses to the full time course of population activity using VSD imaging, and investigate how self-organization is affected by using this finer time scale during long-term development.

Because GCAL models are driven by afferent activity, the type and properties of the visual input patterns assumed are important aspects of each model. Ideally, a model of visual system development in primates would be driven by color, stereo, foveated video streams replicating typical patterns of eye movements, movements of an animal in its environment, and responses to visual patterns. Collecting data of this sort is difficult, and moreover cannot capture any causal or contingent relationships between the current visual input and the current neural organization that can affect future eye and organism movements that will then change the visual input. In the long run, to account for more complex aspects of visual system development such as visual object recognition and optic flow processing, it will be necessary to implement the models as embodied, situated agents [60, 83] embedded in the real world or in realistic 3D virtual environments. Building such robotic or virtual agents will add significant additional complexity, however, so it is important first to see how much of the behavior of V1 neurons can be addressed by the present “open-loop”, non-situated approach.

As discussed throughout, the main focus of this modelling work has been on replicating experimental data using a small number of computational primitives and mechanisms, with a goal of providing a concise, concrete, and relatively simple explanation for a wide and complex range of experimental findings. A complete explanation of visual cortex development and function would go even further, demonstrating more clearly why the cortex should be built in this way, and precisely what information-processing purpose this circuit performs. For instance, realistic RFs can be obtained from “normative” models embodying the idea that the cortex is developing a set of basis functions to represent input patterns faithfully, with only a few active neurons [16, 38, 56, 62]. Maps can emerge by minimizing connection lengths in the cortex [44], and lateral connections can be modelled as decorrelating the input patterns [8, 28]. The GCAL model can be seen as a concrete, mechanistic implementation of these ideas, showing how a physically realizable local circuit could develop RFs with good coverage of the input space, via lateral interactions that also implement sparsification via decorrelation [49]. Making more explicit links between mechanistic models like GCAL and normative theories is an important goal for future work. Meanwhile, there are many aspects of cortical function not explained by current normative models. The focus of the current line of research is on first capturing those phenomena in a mechanistic model, so that researchers can then build deeper explanations for why these computations are useful for the organism.

As previously emphasized, many of the individual results found with GCAL can also be obtained using other modelling approaches, which can be complementary to the processes modeled by GCAL. For instance, it is possible to generate orientation maps without any activity-dependent plasticity, through the initial wiring pattern between the retina and the cortex [57, 63] or within the cortex itself [36]. Such an approach cannot explain subsequent experience-dependent development, whereas the Hebbian approach of GCAL can explain both the initial map and later plasticity, but it is of course possible that the initial map and plasticity occur via different mechanisms. Other models are based on abstractions of some of the mechanisms in GCAL [31, 54, 85, 90], operating similarly but at a higher level. GCAL is not meant as a competitor to such models, but as a concrete, physically realizable implementation of those ideas, forming a prototype of both the biological system and potential future artificial vision systems.

5 Conclusions

The GCAL model results suggest that it will soon be feasible to build a single model visual system that will account for a very large fraction of the visual response properties, at the firing rate level, of V1 neurons in a particular species. Such a model will help researchers make testable predictions to drive future experiments to understand cortical processing, as well as determine which properties require more complex approaches, such as feedback, attention, and detailed neural geometry and dynamics. The model suggests that cortical neurons develop to cover the typical range of variation in their thalamic inputs, within the context of a smooth, multidimensional topographic map, and that lateral connections store pairwise correlations and use this information to modulate responses to natural scenes, dynamically adapting to both long-term and short-term visual input statistics.

Because the model cortex starts without any specialization for vision, it represents a general model for any cortical region, and is also an implementation for a generic information processing device that could have important applications outside of neuroscience. By integrating and unifying a wide range of experimental results, the model should thus help advance our understanding of cortical processing and biological information processing in general.
Acknowledgements
Thanks to all of the collaborators whose modelling work is reviewed here, and to the members of the Developmental Computational Neuroscience research group, the Institute for Adaptive and Neural Computation, and the Doctoral Training Centre in Neuroinformatics, at the University of Edinburgh, for discussions and feedback on many of the models. This work was supported in part by the UK EPSRC and BBSRC Doctoral Training Centre in Neuroinformatics, under grants EP/F500385/1 and BB/F529254/1, and by the US NIMH grant RO1-MH66991. Computational resources were provided by the Edinburgh Compute and Data Facility (ECDF).

References


