Learning Innate Face Preferences

James A. Bednar and Risto Miikkulainen

Department of Computer Sciences
The University of Texas at Austin
Austin, TX 78712
{jbednar,risto}@cs.utexas.edu

Neural Computation 15(7): 1525–1557, 2003.

Abstract

Newborn humans preferentially orient to face-like patterns at birth, but months of experience with faces is required for full face processing abilities to develop. Several models have been proposed for how the interaction of genetic and evironmental influences can explain this data. These models generally assume that the brain areas responsible for newborn orienting responses are not capable of learning and are physically separate from those that later learn from real faces. However, it has been difficult to reconcile these models with recent discoveries of face learning in newborns and young infants. We propose a general mechanism by which genetically specified and environment-driven preferences can coexist in the same visual areas. In particular, newborn face orienting may be the result of prenatal exposure of a learning system to internally generated input patterns, such as those found in PGO waves during REM sleep. Simulating this process with the HLISSOM biological model of the visual system, we demonstrate that the combination of learning and internal patterns is an efficient way to specify and develop circuitry for face perception. This prenatal learning can account for the newborn preferences for schematic and photographic images of faces, providing a computational explanation for how genetic influences interact with experience to construct a complex adaptive system.

1 Introduction

A number of studies have found that newborn humans will preferentially turn their eyes or head towards facelike stimuli within minutes or hours after birth (Goren et al., 1975; Johnson et al., 1991; Johnson and Morton, 1991; Mondloch et al., 1999; Simion et al., 1998b; Valenza et al., 1996). These results suggest that some rudimentary face-selective visual circuitry develops prenatally. However, full face processing abilities appear to require months or years of experience with faces (as reviewed in de Haan, 2001; Gauthier and Nelson, 2001; Pascalis and Slater, 2001; Tovée, 1998).

As will be described in the next section, several theoretical and computational models have been proposed to explain how face proessing develops through a combination of genetic and environmental influences. Most of these models assume that the brain areas responsible for newborn orienting behavior do not learn and are physically separate from those that later learn from real faces (e.g. Johnson and Morton, 1991). However, recent studies show that newborns are capable of learning faces in the first few hours and days after birth (Bushnell et al., 1989; Field et al., 1984; Walton et al., 1997; Walton and Bower, 1993). It has been difficult to reconcile the existing interactional models with early learning of faces.

In this paper we propose a novel explanation for how genetic and environmental influences can interact in a way that allows the visual system to learn faces at all ages. The main hypothesis is that the same brain regions and learning mechanisms are active throughout development, both before and after birth. In both cases the learning mechanisms are extracting and encoding regularities in their input patterns, but in prenatal development the input patterns are internally generated and genetically controlled, whereas in postnatal development the input patterns come from the visual environment (Constantine-Paton et al., 1990; Jouvet, 1999; Marks et al., 1995; Roffwarg et al., 1966; Shatz, 1990, 1996). This way, instead of precisely specifying the organization of the brain, the genome can encode a developmental process that is based on presenting genetically determined patterns to a learning mechanism. After birth, the learning mechanism can then seamlessly integrate environmental and genetic information into the same cortical hardware. Separating the genetic specification and visual processing circuitry in this manner also lets each evolve and develop independently. A simple, genetically fixed pattern generating region can then serve to direct the development of an arbitrarily complex visual system. This explanation can help resolve a number of the confusing results from studies of infant face perception.

We test this idea computationally using a self-organizing model of cortical development (see Erwin et al., 1995; Swindale, 1996 for reviews of this class of models). Computational models allow the assumptions and implications of a conceptual model to be evaluated rigorously, and are particularly useful for generating testable predictions. In previous experiments using the RF-LISSOM model (Receptive-Field Laterally Interconnected Synergetically Self-Organizing Map; Miikkulainen, Bednar, Choe, and Sirosh, 1997; Sirosh and Miikkulainen, 1994; Sirosh, Miikkulainen, and Bednar, 1996) we have shown that patterns of orientation preferences and lateral connections in the primary visual cortex can form from spontaneously generated activity (Bednar and Miikkulainen, 1998). We have also shown in preliminary work how a simplified model of face processing can account for much of the behavioral data on newborn face preferences (Bednar and Miikkulainen, 2000).

In this paper we present HLISSOM, a significant extension to RF-LISSOM that allows processing of real photographic images by modeling the retina, LGN, and higher-level visual areas in addition to the primary visual cortex. The same model allows preferences for low-level features such as orientation or spatial frequency, and high-level features such as faces. to be compared, in order to account for the different categories of preferences found in experiments with newborns. Very few self-organizing models have been tested with real images, and none to our knowledge have previously modeled the self-organization of both the primary visual cortex and higher regions, nor have any simulated the large retinal and cortical area needed to process the large stimuli tested with newborns. Using HLISSOM, we demonstrate how patterns found during REM sleep could explain a newborn's orienting responses towards faces, within a general-purpose adaptive system for learning both low-level and high-level features of images. This explanation will be called the pattern-generation model of newborn face preferences.

In the next section, we will review other models of the early development of face perception as well as experimental evidence for the internally generated brain activity upon which the pattern-generation model is based. Next, we introduce the HLISSOM architecture and describe the experimental setup for testing face detection computationally. We then show how this model can account for newborn face preferences using both schematic and photographic images, and suggest future computational and psychophysical experiments to help understand the development of face processing.

2 Related work

There are three main previous explanations for how a newborn can show an initial face preference and later develop full face-processing abilities: (1) the *Linear Systems Model* (LSM), (2) *Sensory models* (of which we consider the Acerra et al. (2002) computational model and the *Top-heavy* conceptual model separately), and (3) *Multiple systems models*. These hypotheses will be reviewed below, and we will argue that the pattern generation model provides a simpler, more effective explanation.

2.1 Linear Systems Model

The Linear Systems Model (Banks and Salapatek, 1981; Kleiner, 1993) is a straightforward and effective way of explaining a wide variety of newborn pattern preferences, and could easily be implemented as a computational model. Due to its generality and simplicity, it constitutes a baseline model against which others can be compared. The LSM is based solely on the newborn's measured contrast sensitivity function (CSF). For a given spatial frequency, the value of the CSF will be high if the early visual pathways respond strongly to that size of pattern, and lower otherwise. The newborn CSF is limited by the immature state of the eye and the early visual pathways, which makes lower frequencies more visible than fine detail.

The assumption of the LSM is that newborns simply pay attention to those patterns that when convolved with the CSF give the largest response. Low-contrast patterns and patterns with only very fine detail are only faintly visible, if at all, to newborns (Banks and Salapatek, 1981). Con-

versely, faces might be preferred simply because they have strong spatial-frequency components in the ranges that are most visible to newborns.

However, studies have found that the LSM fails to account for the responses to facelike stimuli. For instance, some of the facelike patterns preferred by newborns have a lower amplitude spectrum in the visible range (and thus lower expected LSM response) than patterns that are less preferred (Johnson and Morton, 1991). The LSM also predicts that the newborn will respond equally well to a schematic face regardless of its orientation, because the orientation does not affect the spatial frequency or the contrast. Instead, newborns prefer schematic facelike stimuli oriented right-side-up. Such a preference is found even when the inverted stimulus is a better match to the CSF (Valenza et al., 1996). Thus the CSF alone does not explain face preferences, and a more complicated model is required.

2.2 Acerra et al. sensory model

The LSM is only a high-level abstraction of the properties of the early visual system. *Sensory models* extend the LSM to include additional constraints and circuitry, but without adding face-selective visual regions or systems. Acerra et al. (2002) recently developed such a computational model that can account for some of the face preferences found in the Valenza et al. (1996) study. Their model consists of a fixed Gabor-filter–based model of V1, plus a higher level sheet of neurons with modifiable connections. They model two conditions separately: newborn face preferences, and postnatal development by 4 months. The newborn model includes only V1, because they assume that the higher level sheet is not yet functional at birth.

Acerra et al. showed that the newborn model responds slightly more strongly to the upright schematic face pattern used by Valenza et al. (1996) than to the inverted one, because of differences in the responses of a subset of the V1 neurons. This surprising result replicates the newborn face preferences from Valenza et al. The difference in model responses is because Valenza et al. actually inverted only the internal facial features, not the entire pattern. In the upright case, the spacing is more regular between the internal features and the face outline (compare figure 5*d* with 5*g*, row Retina, on page 38). Making the spacing more regular between alternating white and black blobs will increase the response of a Gabor filter whose RF lobes match the blob size. As a result, the total response of all filters will be slightly higher for the facelike (upright) pattern than to the non-facelike (inverted) pattern.

However, the Acerra et al. model was not tested with patterns from other studies of newborn face preferences, such as Johnson et al. (1991). The facelike stimuli published in Johnson et al. (1991) do not have a regular spacing between the internal features and the outline, and so we do not expect that the Acerra et al. model will replicate preferences for these patterns. Moreover, Johnson et al. used a white paddle against a light-colored ceiling, and so their face outlines would have a much lower contrast than the black-background patterns used in Valenza et al. (1996). Thus although border effects may contribute to the face preferences found by Valenza et al., they are unlikely to be explain those measured by Johnson et al.

The Acerra et al. newborn model also does not explain newborn learning of faces, because their V1 model is fixed and the higher level area is assumed not to be functional at birth. Also importantly, the Acerra et al. newborn model was not tested with real images of faces, where the

spacing of the internal features from the face outline varies widely depending on the way the hair falls. Because of these differences, we do not expect that the Acerra et al. model will show a significantly higher response overall to photographs of real faces than to other similar images. The pattern-generation model will make the opposite prediction, and will explain how newborns can learn faces.

2.3 Top-heavy sensory model

Simion et al. (2001) have also presented a sensory model of newborn preferences, but unlike Acerra et al. (2002) it is a conceptual model only. Simion et al. observed that nearly all of the facelike schematic patterns that have been tested with newborns are "top-heavy," i.e. they have a boundary that contains denser patterns in the upper than the lower half. Simion et al. (2001) also ran behavioral experiments showing that newborns prefer several top-heavy (but not facelike) schematic patterns to similar but inverted patterns. Based on these results, they proposed that newborns prefer top-heavy patterns in general, and thus prefer facelike schematic patterns as a special case.

This hypothesis is compatible with most of the experimental data so far collected in newborns. However, facelike patterns have not yet been compared directly with other top-heavy patterns in newborn studies. Thus it is not yet known whether newborns would prefer a facelike pattern to a similarly top-heavy but not facelike pattern. Future experimental tests with newborns can resolve this issue.

To be tested computationally, the top-heavy hypothesis would need to be made more explicit, with a specific mechanism for locating object boundaries and the relative locations of patterns within them. We expect that such a test would find only a small preference (if any) for photographs of real faces, compared to many other common stimuli. For faces with beards, wide smiles, or wide-open mouths, many control patterns would be preferred over the face image, because such face images will no longer be top-heavy. Because the bulk of the current evidence suggests that instead the newborn preferences are more selective for faces, most of the HLISSOM simulations will use training patterns that result in strongly face-selective visual responses. Section 6.1 will summarize results from training patterns that are less facelike and thus more comparable to the top-heavy model.

2.4 Multiple-systems models

The most widely known conceptual model for newborn face preferences and later learning was proposed by Johnson and Morton (1991). Apart from our tests using the HLISSOM architecture, it has not yet been evaluated computationally. Johnson and Morton suggested that infant face preferences are mediated by two hypothetical visual processing systems that they dubbed CONSPEC and CONLERN.

CONSPEC is a non-plastic system controlling orienting to face-like patterns, assumed to be located in the subcortical superior-colliculus–pulvinar pathway. Johnson and Morton proposed that a CONSPEC responding to three blobs in a triangular configuration, one each for the eyes and one for the nose/mouth region, would account for the newborn face preferences (for examples

see figure 3*a*, page 36 and the Retina row of 4*c*, page 37).

CONLERN is a separate plastic cortical system, presumably the face-processing areas that have been found in adults and in infant monkeys (Kanwisher et al., 1997; Rodman, 1994). The CONLERN system would assume control only after about 6 weeks of age, and would account for the face preferences seen in older infants. Eventually, as it learns from real faces, CONLERN would gradually stop responding to schematic faces, which would explain why face preferences can no longer be measured with static schematic patterns by five months (Johnson and Morton, 1991).

The CONSPEC/CONLERN model is plausible, given that the superior colliculus is fairly mature in newborn monkeys and does seem to be controlling attention and other functions (Wallace et al., 1997). Moreover, some neurons in the adult superior colliculus/pulvinar pathway appear to be selective for faces (Morris et al., 1999), although such neurons have not yet been found in young animals. The model also helps explain why infants after one month of age show a reduced interest in faces in the periphery, which could occur as attentional control shifts to the not-quite-mature cortical system (Johnson et al., 1991; Johnson and Morton, 1991).

However, subsequent studies showed that even newborns are capable of learning individual faces (Slater, 1993; Slater and Kirby, 1998). Thus if there are two visual processing systems, either both are plastic or both are functioning at birth, and thus there is no *a priori* reason why a single face-selective visual system would be insufficient. On the other hand, de Schonen et al. (1998) argue for *three* visual processing systems: a subcortical one responsible for facial feature preferences at birth, another one responsible for newborn learning (of objects and head/hair outlines; Slater, 1993), and a cortical system responsible for older infant and adult learning of facial features. And Simion et al. (1998a) proposed that face selectivity instead relies on multiple visual processing systems *within* the cortex, maturing first in the dorsal stream but later supplanted by the ventral stream (which is where most of the face-selective visual regions have been found in adult humans).

In contrast to the increasing complexity of these explanations, we propose that a single general-purpose, plastic visual processing system is sufficient, if that system is first exposed to internally generated facelike patterns of neural activity.

2.5 Internally generated activity

Spontaneous neural activity has recently been documented in many cortical and subcortical areas as they develop, including the visual cortex, the retina, the auditory system, and the spinal cord (Feller et al., 1996; Leinekugel et al., 2002; Lippe, 1994; Wong et al., 1993; Yuste et al., 1995; reviewed in O'Donovan, 1999; Wong, 1999). Such activity has been shown to be responsible for the segregation of the LGN into eye-specific layers before birth, indicating that internally generated activity is crucial for normal development (Shatz, 1990, 1996).

We will focus on one common type of spontaneous activity: ponto-geniculo-occipital (PGO) waves generated during rapid-eye-movement (REM) sleep. Developing embryos spend a large percentage of their time in what appears to be REM sleep, which suggests that it plays a major role in development (Marks et al., 1995; Roffwarg et al., 1966). During and just before REM

sleep, PGO waves originate in the pons of the brain stem and travel to the LGN, visual cortex, and a variety of subcortical areas (see Callaway et al., 1987 for a review). PGO waves are strongly correlated with eye movements and with vivid visual imagery in dreams, suggesting that they activate the visual system as if they were visual inputs (Marks et al., 1995). PGO waves have been found to elicit different distributions of activity in different species (Datta, 1997), and interrupting them has been shown to increase the influence of the environment on development (Marks et al., 1995; Shaffery et al., 1999).

All of these characteristics suggest that PGO waves may be providing species-specific training patterns for development (Jouvet, 1998). However, due to limitations in experimental imaging equipment and techniques, it has not yet been possible to measure the two-dimensional shape of the activity resulting from the PGO waves; only the temporal behavior is known (Rector et al., 1997). How closely neonatal REM sleep and PGO waves are related to those of adults is also controversial (Dugovic and Turek, 2001; Jouvet, 1999). In any case, it is clear that there is substantial brainstem-generated activity in the visual areas of newborns during sleep. This paper predicts that if these activity patterns have a simple spatial configuration of three active areas (or other similar patterns), they can account for the measured face-detection performance of human newborns.

The three-dot training patterns are similar to the three-dot preferences proposed by Johnson and Morton (1991). However, in that model the patterns are implemented as a hard-wired subcortical visual area responsible for orienting to faces. In the pattern-generation model, the areas receiving visual input learn from their input patterns, and only the pattern generator is assumed to be hard-wired.

Both of these possible mechanisms for newborn orienting responses would require about the same amount of genetic specification, but the crucial difference is that in the pattern generation approach the visual processing system can be arbitrarily complex. In contrast, a hard-coded visual processing system like CONSPEC is limited in complexity by what can specifically be encoded in the genome. For this reason, a system implemented like CONSPEC is a plausible model for subcortically mediated orienting at birth, but would become less plausible if imaging techniques later show that newborn face preferences are mediated by a hierarchy of cortical areas as in the adult.

Specifying only the training patterns also makes sense from an evolutionary perspective, because that way the patterns and the visual processing hardware can evolve independently. Separating these capabilities thus allows the visual system to become arbitrarily complex, while maintaining the same genetically specified function. Finally, assuming that only the pattern generator is hardwired explains how infants of all ages can learn faces.

3 HLISSOM model

We will investigate the pattern-generation hypothesis using the HLISSOM model of the primate visual system, focusing on how sub-cortically generated patterns can drive the development of cortical areas. The HLISSOM architecture is shown in figure 1.The model consists of a hierarchy of two-dimensional sheets of neural units modeling different areas of the nervous system: two sheets of input units (the retinal photoreceptors and the PGO generator), two sheets of LGN units

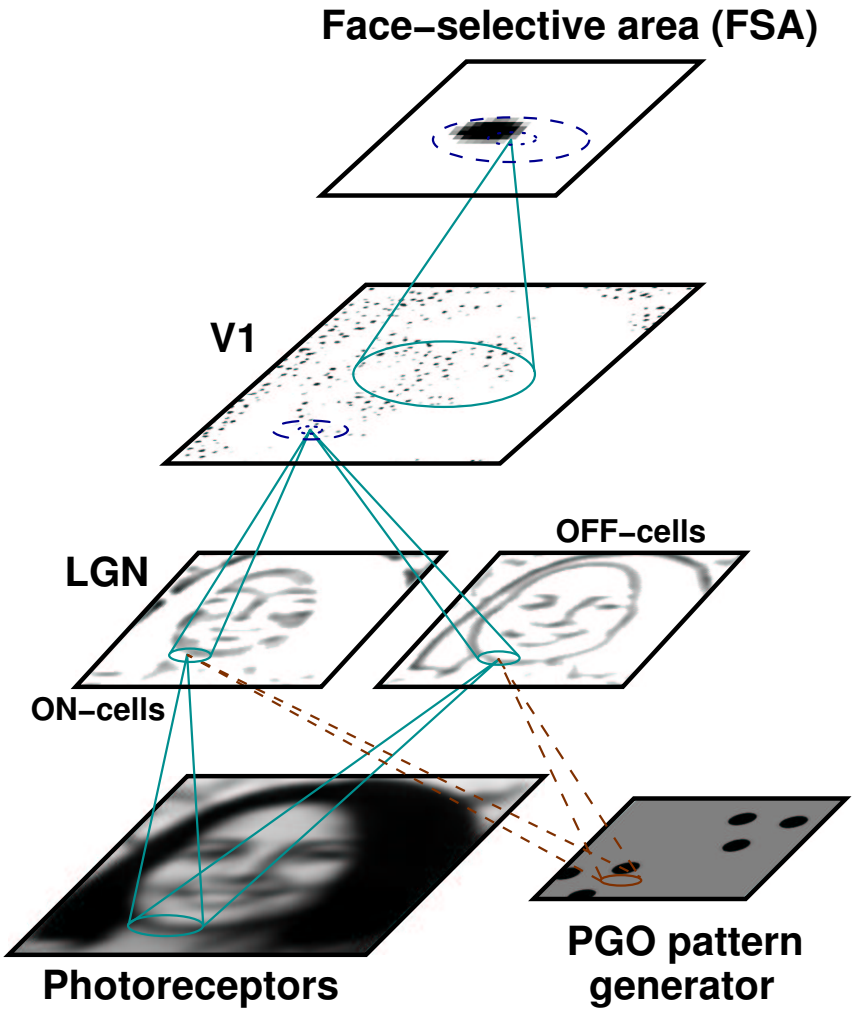


Figure 1: **Schematic diagram of the HLISSOM model.** Each sheet of units in the model visual pathway is shown with a sample activation pattern and the connections to one example unit. Grayscale visual inputs are presented on the photoreceptors, and the resulting activity propagates through afferent connections to each of the higher levels. Internally generated PGO input propagates similarly to visual input, but travels through the ponto-geniculate (PG) pathway to occipital cortex (i.e., V1), rather than the retino-geniculate pathway. As in humans and animals, activity may be generated either in the PGO area or the photoreceptors, but not both at once. In the cortical levels (V1 and FSA), activity is focused by lateral connections, which are initially excitatory between nearby neurons (dotted circles) and inhibitory between more distant neurons (dashed circles). The final patterns of lateral and afferent connections in the cortical areas develop through an unsupervised self-organizing process. After self-organization is complete, each stage in the hierarchy represents a different level of abstraction, transforming the raw image input into a more biologically relevant representation. The LGN responds best to edges and lines, suppressing areas with no information. The V1 response is further selective for the orientation of each contour; the patchiness is due to geometrical constraints on representing all orientations on a two-dimensional surface. The FSA represents the highest level of abstraction — the response of an FSA neuron signals the presence of a face at the corresponding location on the retina, which is information that an organism can use directly to control behaviors like orienting.

(ON-center and OFF-center), and two sheets of cortical units ("neurons"): the primary visual cortex (V1), and a higher-level area from the ventral processing stream with large receptive fields (here called the face-selective area, FSA, as described below). Because the focus is on the two-dimensional organization of each cortical sheet, a neuron in V1 or the FSA corresponds to a vertical column of cells through the six anatomical layers of that area of the cortex.

The FSA represents the first region in the ventral processing pathway above V1 that has receptive fields (RFs) spanning approximately 45° of visual arc, i.e. large enough to span a human face at close range. Areas V4v and LO are likely FSA candidates based on adult patterns of connectivity, but the infant connectivity patterns are not known (Rolls, 1990). We use the generic term "face-selective area" rather than V4v or LO to emphasize that the model results depend only on having receptive fields large enough to allow face-selective responses, not on the region's precise location or architecture.

Within each cortical sheet (V1 or the FSA), neurons receive afferent connections from receptive fields in the form of broad overlapping circular patches of units in the previous sheet(s) in the hierarchy. Neurons in the cortical sheets also have reciprocal excitatory and inhibitory lateral connections to neurons within the same sheet. Lateral excitatory connections are short range, connecting each neuron with itself and its close neighbors in a circular radius. Lateral inhibitory connections extend in a larger radius, but also include connections to the neuron itself and to its neighbors.

This pattern of lateral connections models the lateral interactions measured experimentally for high-contrast input patterns. Because the Hebbian learning used in this model is proportional to the activity levels of the neurons, the higher contrast inputs have more influence on adaptation. Optical imaging and electrophysiological studies have shown that at these contrasts long-range interactions in the cortex are inhibitory, even though individual lateral connections are primarily excitatory (Hirsch and Gilbert, 1991; Weliky et al., 1995). Computational studies have also shown that long-range interactions between high-contrast inputs must be inhibitory for proper self-organization to occur (Sirosh and Miikkulainen, 1994). For simplicity, the model uses only inhibitory interactions, but additional long-range excitation can be included when studying phenomena such as perceptual grouping (Choe, 2001; Choe and Miikkulainen, 2000).

Previous models have explained how the connections in the LGN could develop from internally generated activity in the retina (Eglen, 1997; Haith, 1998). Because the HLISSOM model focuses on learning at the cortical level, all connection strengths to such subcortical neurons are set to fixed values. The strengths were chosen to approximate the receptive fields that have been measured in adult LGN cells, using a standard Difference-of-Gaussians model (Rodieck, 1965). In this model, the center of each receptive field is first mapped to the location in the input sheet corresponding to the location of the LGN unit, and the weight to a receptor is then a fixed function of the distance from that center. For an ON-center cell, the receptor weights are calculated from the difference of two normalized Gaussians with radii σ_c (center) and σ_s (surround). The weights for an OFF-center cell are the negative of the ON-center weights; i.e., they are the surround minus the center instead of the center minus the surround. The parameter σ_c determines the peak spatial frequency to which the LGN responds, and σ_s is generally set to a fixed multiple of σ_c.

The connections to neurons in the cortical sheets are initially unselective (either random, for afferent weights, or Gaussian-shaped, for lateral weights), and their strengths are subsequently

modified through an unsupervised learning process. The learning process is driven by input patterns. Patterns may be presented on the photoreceptor sheet, representing external visual input, or they may be generated internally, reaching V1 through the PGO pathway. The PGO-generating region of the pons and its pathway to the LGN have not yet been mapped out in detail, but since the activity that results from the PGO waves is similar to that from visual input (Marks et al., 1995), for simplicity we assume that the retino-geniculate and PGO pathways are similar. Therefore, the PGO pattern input is modeled with an area like the photoreceptor sheet, connecting to the LGN in the same way.

At each training step, the activities of all units are initialized to zero, and a grayscale pattern is drawn on the input sheet (either the PGO generator or the photoreceptors; see figures 2*a*,*i* and 3*a* for examples). The cells in the ON and OFF channels of the LGN compute their responses as a scalar product of their fixed weight vector and the activity of units in their receptive fields (as in figures 2*b*,*j* and 3*b*). The response η_{ab} of ON or OFF-center cell (a, b) is calculated from the weighted sum of the retinal activations in its receptive field as

$$\eta_{ab} = \sigma\left(\gamma_{\mathrm{A}} \sum_{xy} \xi_{xy} \mu_{ab,xy}\right), \tag{1}$$

where σ is a piecewise linear approximation of the sigmoid activation function, ξ_{xy} is the activation of retina cell (x, y), and $\mu_{ab,xy}$ is the corresponding afferent weight. An LGN neuron will respond when the input pattern is a better match to the central portion of the RF than to the surrounding portion; the response will be larger with higher contrast, subject to the minimum and maximum values enforced by the sigmoid. The constant afferent scaling factor γ_{A} is set so that the LGN outputs approach 1.0 in the highest-contrast regions of typical input patterns.

The cortical activation is similar to the LGN, but extended to support self-organization and to include lateral interactions. The total activation is computed from both afferent and lateral responses, which will be discussed separately here. First, the afferent response ζ_{ij} of V1 neuron (i, j) is calculated like the retinal sum in equation 1, with an additional divisive (shunting) normalization:

$$\zeta_{ij} = \frac{\gamma_{\mathrm{A}} \sum_{\rho ab} \xi_{\rho ab} \mu_{ij,\rho ab}}{1 + \gamma_{\mathrm{N}} \sum_{\rho ab} \xi_{\rho ab}}, \tag{2}$$

where $\xi_{\rho ab}$ is the activation of unit (a, b) in sheet ρ (either the ON channel or the OFF channel and $\mu_{ij,\rho ab}$ is the corresponding afferent weight. The normalization strength γ_{N} starts at zero and is gradually increased over training as neurons become more selective, and the scaling factor γ_{A} is set to compensate so that the afferent response ζ will continue to have values in the range 0 to 1.0 for typical input patterns. The FSA computes its afferent response just as V1 does, except that sheet ρ in equation 2 is the settled response (described below) of V1, instead of the response of an ON or OFF channel of the LGN.

Note that although only a single ON and OFF channel were used here, afferent inputs from additional channels with different peak spatial frequencies (i.e. different σ_c) can be included in equation 2 in the same way. Simulating a single channel allows V1 to be smaller and makes its output easier to interpret. This is a valid approximation because the response of each channel is broad, so even a single channel responds to a wide range of input patterns.

The initial response of a cortical neuron, whether in V1 or the FSA, is computed from the afferent response only:

$$\eta_{ij}(0) = \sigma\left(\zeta_{ij}\right), \tag{3}$$

where σ is a piecewise linear approximation of the sigmoid activation function. After the initial response, the cortical activity evolves over a very short time scale through lateral interaction. At each subsequent time step, the neuron combines the afferent response ζ with lateral excitation and inhibition:

$$\eta_{ij}(t) = \sigma\left(\zeta_{ij} + \gamma_{\mathrm{E}} \sum_{kl} E_{ij,kl}\eta_{kl}(t-1) - \gamma_{\mathrm{I}} \sum_{kl} I_{ij,kl}\eta_{kl}(t-1)\right), \tag{4}$$

where $E_{ij,kl}$ is the excitatory lateral connection weight on the connection from neuron (k,l) to neuron (i,j), $I_{ij,kl}$ is the inhibitory connection weight, and $\eta_{kl}(t-1)$ is the activity of neuron (k,l) during the previous time step. The scaling factors γ_{E} and γ_{I} determine the relative strengths of excitatory and inhibitory lateral interactions.

While the cortical response in V1 and the FSA is settling, the afferent response remains constant. The cortical activity pattern in both areas starts out diffuse, but within a few iterations of equation 4, converges into a small number of stable focused patches of activity, or activity bubbles (as in figures 2*c*,*d*,*k* and 3*e*,*f*). After the activity has settled, the connection weights of each neuron are modified. Afferent and lateral weights in both areas adapt according to the same mechanism: the Hebb rule, normalized so that the sum of the weights is constant:

$$w_{ij,mn}(t+\delta t) = \frac{w_{ij,mn}(t) + \alpha\eta_{ij}X_{mn}}{\sum_{mn}\left[w_{ij,mn}(t) + \alpha\eta_{ij}X_{mn}\right]}, \tag{5}$$

where η_{ij} stands for the activity of neuron (i,j) in the final activity bubble, $w_{ij,mn}$ is the afferent or lateral connection weight (μ, E or I), α is the learning rate for each type of connection (α_A for afferent weights, α_E for excitatory, and α_I for inhibitory) and X_{mn} is the presynaptic activity (ξ for afferent, η for lateral). The larger the product of the pre- and post-synaptic activity $\eta_{ij}X_{mn}$, the larger the weight change. At long distances, very few neurons in the cortical regions have correlated activity and therefore most long-range connections eventually become weak. The weak connections are eliminated periodically, resulting in patchy lateral connectivity similar to that observed in the visual cortex.

4 Experiments

To model the presentation of large stimuli at close range (as in typical experiments with newborns), V1 parameters were generated using the scaling equations from Kelkar et al. (2001) to model a large V1 area (approximately 1600 mm^2 total) at a relatively low sampling density per mm^2 (approximately 50 neurons/mm^2). The total area was chosen to be just large enough for the corresponding photoreceptor area to cover the required range, and the density was chosen to be the minimum that would still provide a V1 organization that matched animal maps. The input pattern scale was then set so that the spatial frequencies for which V1 response was maximal would match the frequency range to which newborns are most sensitive, as cited by Valenza et al. (1996). As mentioned above, the response is broad, and the model also responds to other spatial frequencies.

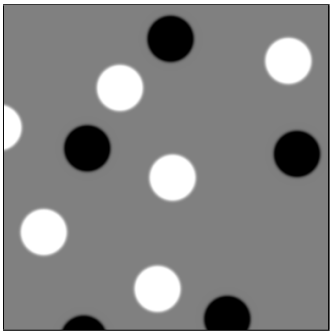
Figure 2: **Training the V1 orientation map (color figure; on next page).** The input consisted of randomly located circular patterns (*a*) approximating spontaneous neural activity waves. The ON and OFF LGN channels compute their responses based on this input; in (*b*), the ON response minus the OFF response is shown, with medium gray indicating regions with no response. The V1 neurons compute their responses based on the output of the ON and OFF channels. Initially V1 neurons have random weights in their receptive fields, and thus all neurons with RFs overlapping the active areas of the LGN respond (*c*). After training, the afferent weights of each neuron show a clear preference for a particular orientation (*e*); the plot shows the weights to the ON-center LGN sheet minus those to the OFF-center sheet, drawn on the full LGN sheet. The orientation maps (*f*,*g*) summarize the development of such preferences. Each neuron is color-coded according to its preferred orientation (using the color key beside *d*); brighter colors indicate a stronger selectivity. Initially the neurons are not selective for orientation (*f*), but after training, V1 develops an orientation map similar to those measured in experimental animals (*g*; Blasdel, 1992). The map covers a large area of V1; previous work corresponded to only the central 36×36 portion (*h*) of the 288×288 map in (*g*). For the same input (*a*), the self-organized network response is patchy (*d*), because only those neurons whose preferred orientation matches the orientations of local features in the image respond. In (*d*) each neuron is colored with the orientation it prefers, and the brightness of the color indicates how active that neuron is. The trained orientation map works well with natural images; each V1 neuron responds to the orientation of lines and edges in the corresponding region of the image (*i*-*k*). Color versions of these figures are available online at `http://www.cs.utexas.edu/users/nn/`.

The FSA parameters were generated using the scaling equations from Kelkar et al. (2001) to model a small region of units that have receptive fields large enough to span the central portion of a face. The final FSA activation threshold (lower bound on the sigmoid function in equations 3 and 4) was set to a high value so that the presence of a blob of activity in the FSA would be a criterion for the presence of a face at the corresponding location in the retina. This high threshold allows the FSA output to be interpreted unambiguously as a behavioral response, without the need for feedback from an additional higher level set of units modeling newborn attention and fixation processes.

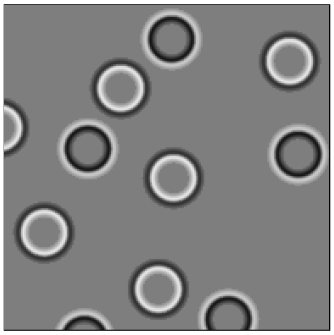
The model consisted of 438×438 photoreceptors, 220×220 PGO generator units, 204×204 ON-center LGN units, 204×204 OFF-center LGN units, 288×288 V1 units, and 36×36 FSA units, for a total of 408,000 distinct units. Area FSA was mapped to the central 160×160 region of V1 such that even the units near the borders of the FSA had a complete set of afferent connections on V1, with no RF extending over the outside edges of V1. V1 was similarly mapped to the central 192×192 region of the LGN channels, and the LGN channels to the central 384×384 region of the photoreceptors and the central 192×192 region of the PGO generators. In the figures below, only the area mapped directly to V1 is shown, to ensure that plots for different sheets have the same size scale. The model requires 300 megabytes of physical memory to represent the 80 million connections in the two cortical sheets. The remaining parameters are listed in appendix A.

4.1 Training V1

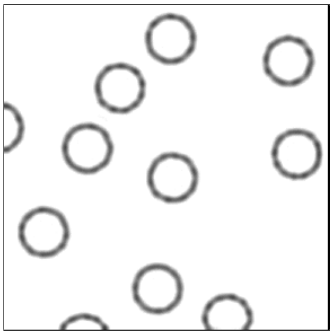
For simplicity, V1 and the FSA were self-organized in separate training phases. V1 was self-organized for $10,000$ iterations, each with approximately 11 randomly located circular discs 25 PGO units wide (figure 2*a*).The background activity level was 0.5, and the brightness of each disc



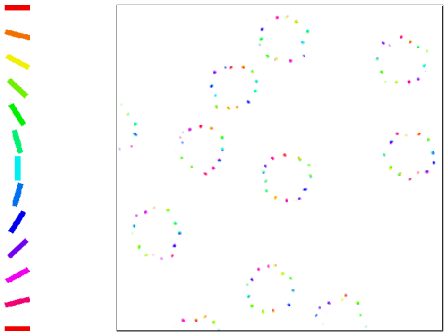
(*a*) Generated pattern



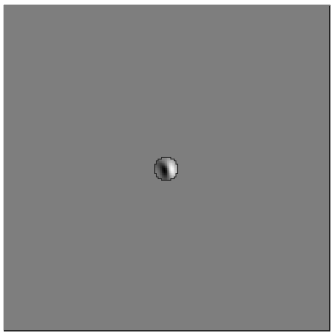
(*b*) LGN activity before training



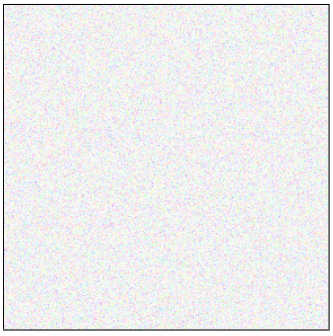
(*c*) V1 activity before training



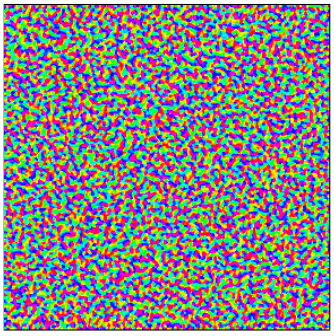
(*d*) V1 activity after training



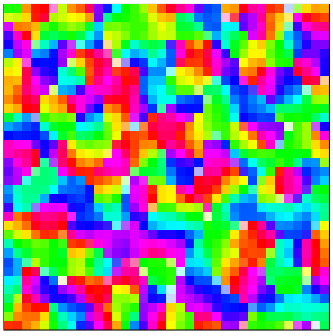
(*e*) Afferent weights of one V1 neuron



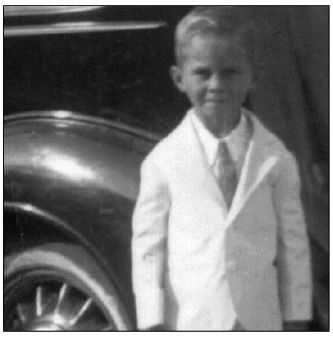
(*f*) Initial V1 orientation map



(*g*) Final V1 orientation map

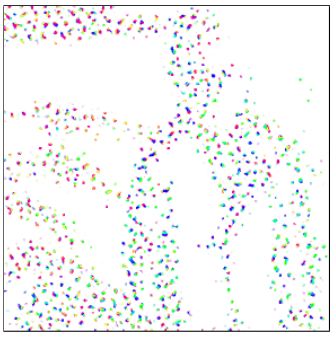


(*h*) Detail of map (*g*)



(*i*) Sample visual image

(*j*) LGN response



(*k*) V1 response

relative to this surround (either +0.3 or -0.3) was chosen randomly. The borders of each disc were smoothed into the background level following a Gaussian of half-width $\sigma = 1.5$. After $10,000$ such iterations (10 hours on a 600MHz Pentium III workstation), the orientation map shown in figure 2*g*-*h* emerged, consisting of neurons that respond strongly to oriented edges and thick lines. The map organization is similar to maps measured in experimental animals (Blasdel, 1992).

We selected these V1 training patterns as well-defined approximations of internally generated spatial patterns, such as spontaneous retinal waves. Imaging data shows that these patterns have spatially coherent blobs of activity (Feller et al., 1996), unlike the uniform random noise assumed in earlier models of spontaneous activity (e.g. Linsker, 1986). Each pattern represents the state of one wave at a given instant, and together the patterns result in a map of static input properties. It is possible to train the model with moving patterns as well, and the resulting map would represent directional selectivity in addition to orientation selectivity (Bednar and Miikkulainen, 2003). However, including only orientation processing is appropriate in this case because all patterns tested with newborns have either been static, or else moving at approximately the same rate as control patterns.

Miller (1994) had previously argued that because retinal waves are much larger than the V1 receptive fields and not strongly oriented, they cannot account for how orientation selectivity develops. The results here demonstrate the opposite conclusion, and show that patterns do not need to be oriented as long as they are smooth and large relative to the RF size of the orientation map. To our knowledge, these simulations are the first to show that realistic orientation maps can develop from large, unoriented (but spatially coherent) stimuli like the retinal waves.

4.2 Training the FSA

After V1 training, the V1 weights were fixed to their self-organized values, and the FSA was allowed to activate and learn from the V1 responses. The FSA was self-organized for $10,000$ iterations using two triples of circular dots, each arranged in a triangular facelike configuration (figure 3*a*). Each dot was 20 PGO units wide, and was 0.3 units darker than the surround, which itself was 0.5 on a scale of 0 to 1.0. Each triple was placed at a random location each iteration, at least 118 PGO units away from the center of the other one (to avoid overlap). The angle of each triple was drawn randomly from a narrow ($\sigma = \pi/36$ radians) normal distribution around vertical.

Because the face preferences found in newborns have all been for faces of approximately the same size (life-sized at a distance of around 20 cm), only a single training pattern size was used. As a result, the model will only be able to detect facelike patterns at one particular size scale. If response to multiple face sizes (i.e., distances) is desired, the spatial scale of the training patterns can be varied during self-organization (Sirosh and Miikkulainen, 1996). However, the FSA in such a simulation would need to be much larger to represent the different sizes, and the resulting patchy FSA responses would require more complex methods of analysis.

Using these three-dot patterns, the FSA was self-organized for $10,000$ iterations (13 hours on a 600MHz Pentium III workstation). The result was the face-selective map shown in figure 3*h*, consisting of an array of neurons that respond most strongly to patterns similar to the training patterns. Despite the overall similarities between neurons, the individual weight patterns are unique because each neuron targets specific orientation blobs in V1. Such complicated patterns would be

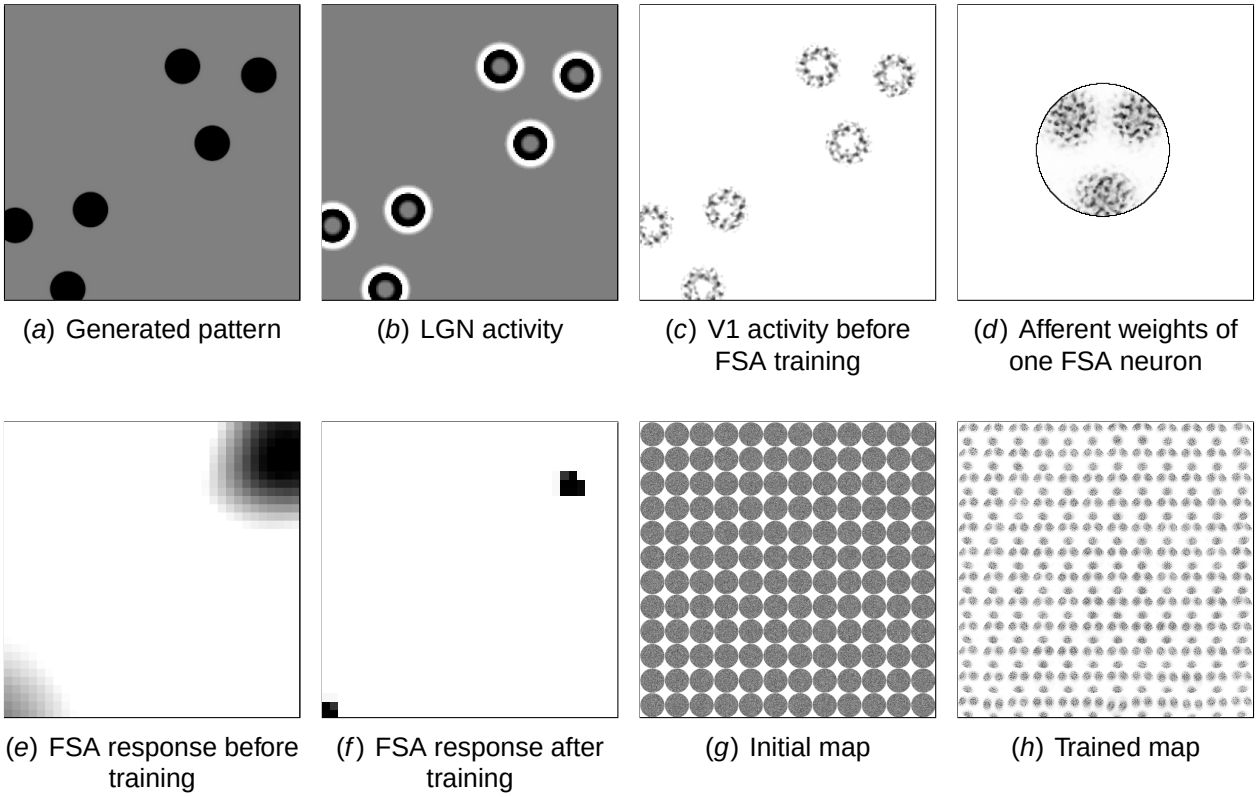


Figure 3: **Training the FSA face map.** (*a*) For training, the input consisted of simple three-dot configurations with random nearly vertical orientations presented at random locations in the PGO layer. These patterns were chosen based on the experiments of Johnson and Morton (1991). (*b*-*c*) The LGN and V1 sheets compute their responses based on this input. Initially, FSA neurons will respond to any activity in their receptive fields (*e*), but after training only neurons with closely matching RFs respond (*f*). Through self-organization, the FSA neurons develop RFs (*d*) that are selective for a range of V1 activity patterns like those resulting from the three-dot stimuli. Importantly, the RFs are patchy because the weights target specific orientation blobs in V1. This complex match between the FSA and the local self-organized pattern in V1 would be difficult to ensure without training on internally generated patterns. Plots (*g*) and (*h*) show the afferent weights for every third neuron in the FSA. All neurons develop roughly similar weight profiles, differing primarily by the position of their preferred stimuli on the retina and by the specific orientation blobs targeted in V1. The largest differences between RFs are along the outside border, where the neurons are less selective for three dot patterns. Overall, the FSA develops into a face detection map, signaling the location of facelike stimuli.

difficult to specify and hard-wire directly, but they arise naturally from internally generated activity.

As FSA training completed, the lower threshold of the sigmoid was increased to ensure that only patterns that are a strong match to an FSA neuron's weights will activate it. This way, the presence of FSA activity can be interpreted unambiguously as an indication that there is a face at the corresponding location in the retina, which will be important for measuring the face detection ability of the network.

The cortical maps were then tested on natural images and with the same schematic stimuli on which human newborns have been tested (Goren et al., 1975; Johnson and Morton, 1991; Simion

et al., 1998b; Valenza et al., 1996). For all of the tests, the same set of model parameter settings described in appendix A were used. Schematic images were scaled to a brightness range (difference between the darkest and lightest pixels in the image) of 1.0. Natural images were scaled to a brightness range of 2.5, so that facial features in images with faces would have a contrast comparable to that of the schematic images.

If both types of images were instead scaled to the same brightness range, the model would prefer the schematic faces over real ones. Using different scales is appropriate for these experiments because we are only comparing responses within groups of similar images, and we assume that the infant has become adapted to images of this type. For simplicity, the model does not specifically include the mechanisms in the eye, LGN, and V1 that are responsible for such contrast adaptation.

4.3 Predicting behavioral responses

The HLISSOM model provides detailed neural responses for each neural region. These responses constitute predictions for future electrophysiological and imaging measurements in animals and humans. However, the data currently available for infant face perception is behavioral. It consists of newborn attention preferences measured from visual tracking distance and looking time. Thus, validating the model on this data will require predicting a behavioral response based on the simulated neural responses.

As a general principle, we hypothesize that newborns pay attention to the stimulus whose overall neural response most clearly differs from those to typical stimuli. This idea can be quantified as

$$a(t) = \frac{F(t)}{\overline{F}} + \frac{V(t)}{\overline{V}}, + \frac{L(t)}{\overline{L}}, \tag{6}$$

where $a(t)$ is the attention level at time t, $X(t)$ is the total activity in region X at that time, $\overline{X}$ is the average (or median) activity of region X over recent history, and F, V, and L represent the FSA, V1, and LGN regions, respectively. Because most stimuli activate the LGN and V1 but not the FSA, when a pattern evokes activity in the FSA newborns would attend to it more strongly. Yet stimuli evoking only V1 activity could still be preferred over facelike patterns if their V1 activity is much higher than typical.

Unfortunately, it is difficult to use such a formula to compare to newborn experiments, because the presentation order of stimuli in those experiments is usually not known. As a result, the average or median value of patterns in recent history is not available. Furthermore, the numerical preference values computed in this way will differ depending on the specific set of patterns chosen, and thus will be different for each study.

Instead, for the simulations below we will use a categorical approach inspired by Cohen (1998) that should yield similar results. Specifically, we will assume that when two stimuli both activate the model FSA, the one with the higher total FSA activation will be preferred. Similarly, with two stimuli activating only V1, the higher total V1 activation will be preferred. When one stimulus activates only V1 and another activates both V1 and the FSA, we will assume that the pattern that produces FSA activity will be preferred, unless the V1 activity is much larger than for typical patterns. Using these guidelines, the computed model preferences can be validated against the newborn's looking preferences, to determine if the model shows the same behavior as the newborn.

5 Results

This section presents the model's response to schematic patterns and real images after it has completed training on the internally generated patterns. The response to each schematic pattern is compared to behavioral results from infants, and the responses to real images constitute predictions for future experiments.

5.1 Schematic patterns

Figures 4 and 5 show that the model responses match the measured stimulus preference of newborns remarkably well, with the same relative ranking in each case where infants have shown a significant preference between schematic patterns. These figures and the rankings are the main result from this simulation. We will next analyze each of the categories of preference rankings, to show what parts of the model underlie the pattern preferences.

Most non-facelike patterns activate only V1, and thus the preferences between those patterns are based only on the V1 activity values (figure 4*a*,*e*-*i*; 5*f*-*i*). Patterns with numerous high-contrast edges have greater V1 response, which explains why newborns would prefer them. These preferences are in accord with the simple LSM model, because they are based only on the early visual processing (section 2.1).

Facelike schematic patterns activate the FSA, whether they are realistic or simply patterns of three dots (figure 4*b*-*d*; 5*a*-*d*). The different FSA activation levels reflect the level of V1 response to each pattern, not the precise shape of the face pattern. The responses explain why newborns would prefer patterns like the three-square-blob face over the oval-blob face, which has a lower edge length. The preferences between these patterns are also compatible with the LSM, because in each case the V1 responses in the model match newborn preferences.

The comparisons between facelike and non-facelike patterns show how the predictions of the HLISSOM pattern generation model differ from those of the LSM. In the HLISSOM simulations, the patterns that activate the FSA are preferred over those activating only V1, except when the V1 response is highly anomalous. Most V1 responses are very similar, and so the patterns with FSA activity (figure 4*b*-*d*; 5*a*-*d*) should be preferred over most of the other patterns, as found in infants. However, the checkerboard pattern (4*a*) has nearly three times as much V1 activity as any other pattern with a similar background. Thus the HLISSOM results explain why the checkerboard would be preferred over other patterns, even ones that activate the FSA.

Because the RFs in the model are only a rough match to the schematic patterns, the FSA can have spurious responses to patterns that to adults do not look like faces. For instance, the inverted three-dot pattern in Figure 5*e* activated the FSA. In this case part of the square outline filled in for the missing third dot of an upright pattern. With high enough pattern contrast, two dots can also be sufficient to give a response. Because inverted patterns have two dots in the correct location, these results suggest that inverted patterns are not an ideal control to use in experimental studies (Bednar, 2002).

Interestingly, the model also showed a clear preference in one case where no significant preference was found in newborns (Simion et al., 1998a): for the upright 3-blob pattern with no face

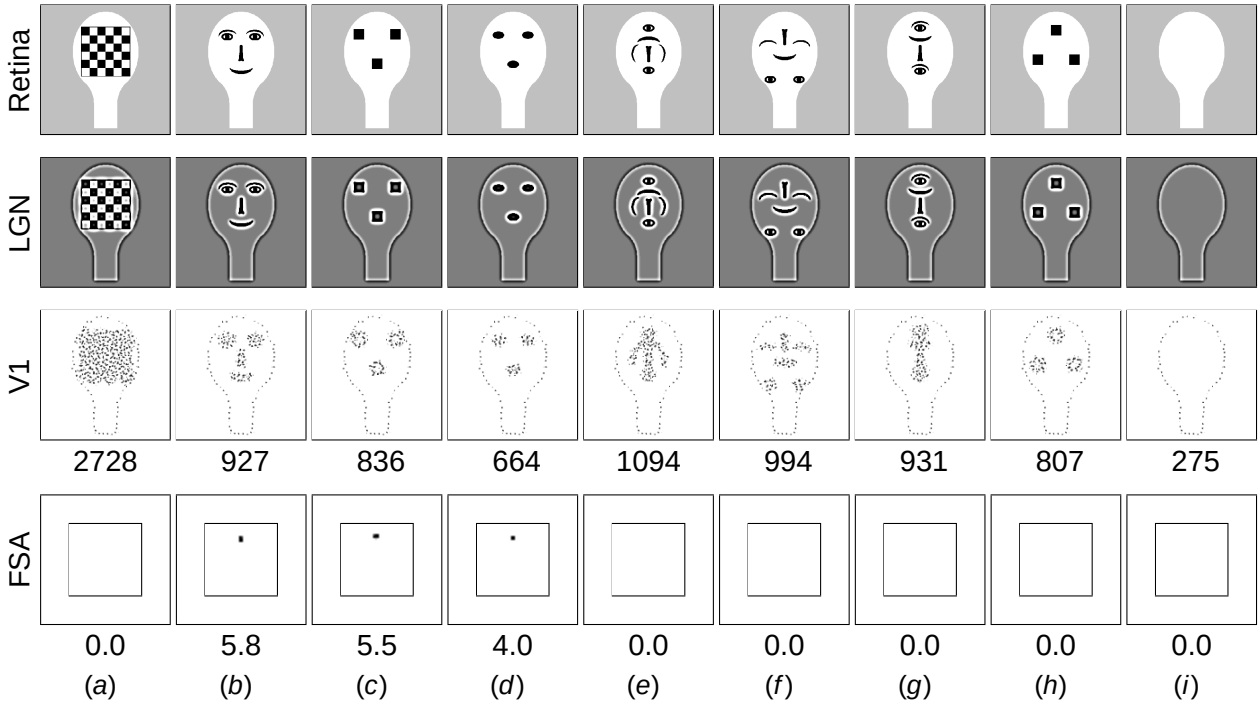


Figure 4: **Human newborn and model response to Goren et al.'s (1975) and Johnson et al.'s (1991) schematic images.** The Retina row along the top shows a set of patterns as they are drawn on the photoreceptor sheet. These patterns have been presented to newborn human infants on head-shaped paddles moving at a short distance (about 20cm) from the newborn's eyes, against a light-colored ceiling. In the experimental studies, the newborn's preference was determined by measuring the average distance his or her eyes or head tracked each pattern, compared to other patterns. Below, $x>y$ indicates that image x was preferred over image y under those conditions. Goren et al. (1975) measured infants between 3 and 27 minutes after birth. They found that *b*>*f*>*i* and *b*>*e*>*i*. Similarly, Johnson et al. (1991), in one experiment measuring within one hour of birth, found *b*>*e*>*i*. In another, measuring at an average of 43 minutes, they found *b*>*e*, and *b*>*h*. Finally, Johnson and Morton (1991), measuring newborns an average of 21 hours old, found that *a*>(*b*,*c*,*d*), *c*>*d*, and *b*>*d*. The HLISSOM model has the same preference for each of these patterns, as shown in the images above. The second row shows the model LGN activations (ON minus OFF) resulting from the patterns in the top row. The third row shows the V1 activations, with the numerical sum of the activities shown underneath. If one unit were fully active, the sum would be 1.0; higher values indicate that more units are active. The bottom row shows the settled response of the FSA. Activity at a given location in the FSA corresponds to a facelike stimulus at that location on the retina, and the sum of this activity is shown underneath. The images are sorted left to right according to the preferences of the model. The strongest V1 response by nearly a factor of three is to the checkerboard pattern (*a*), which explains why the newborn would prefer that pattern over the others. The facelike patterns (*b*-*d*) are preferred over patterns (*e*-*i*) because of activation in the FSA. The details of the facelike patterns do not significantly affect the results — all of the facelike patterns (*b*)–(*d*) lead to FSA activation, generally in proportion to their V1 activation levels. The remaining patterns are ranked by their V1 activity alone, because they do not activate the FSA. In all conditions tested, the HLISSOM model shows behavior remarkably similar to that of the newborns, and provides a detailed computational explanation for why these behaviors occur.

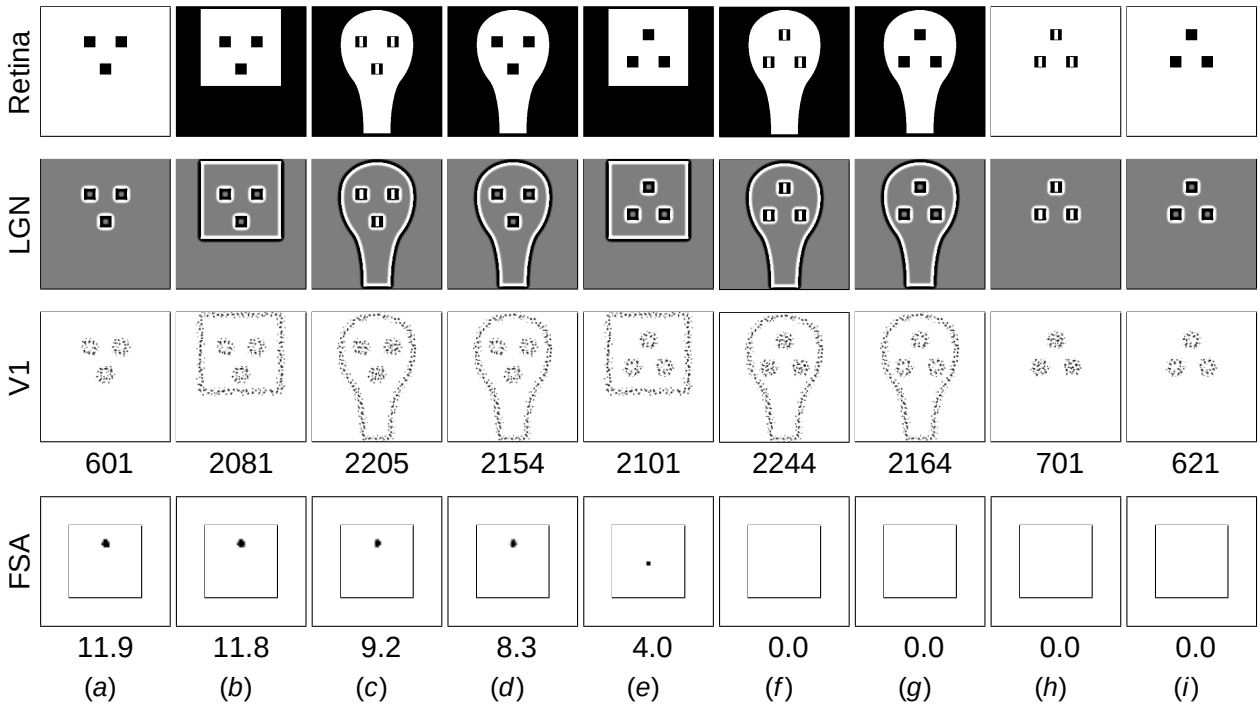


Figure 5: **Response to schematic images from Valenza et al. (1996) and Simion et al. (1998a).** Valenza et al. measured preference between static, projected versions of pairs of the schematic images in the top row, using newborns ranging from 24 to 155 hours after birth. They found the following preferences: *d*>*f*, *d*>*g*, *f*>*g*, and *h*>*i*. Simion et al. similarly found a preference for *d*>*g* and *b*>*e*. The LGN, V1, and FSA responses of the model to these images are displayed here as in figure 4, and are again sorted by the model's preference. In all cases where the newborn showed a preference, the model preference matched it. For instance, the model FSA responds to the facelike pattern (*d*) but not to the inverted version (*g*). Patterns that closely match the newborn's preferred spatial frequencies (*f*,*h*) caused a greater V1 response than their corresponding lower-frequency patterns (*g*,*i*). Some non-facelike patterns with high-contrast borders can cause spurious FSA activation (*e*), because part of the border completes a three-dot pattern. Such spurious responses did not affect the predicted preferences, because they are smaller than the genuine responses.

outline (figure 5*a*), over the similar but inverted pattern (5*i*). The V1 responses to both patterns are similar, but only the upright pattern has an FSA response, and thus the model predicts that the upright pattern would be preferred.

This potentially conflicting result may be due to postnatal learning rather than capabilities at birth. As will be summarized in section 6.2, postnatal learning of face outlines can have a strong effect on FSA responses. The newborns in the Simion et al. study were already 1–6 days old, and Pascalis et al. (1995) has shown that newborns within this age range have already learned some of the features of their mother's face outline. Thus the pattern-generation model predicts that if younger newborns are tested, they will show a preference for upright patterns even without a border.

Alternatively, the border may satisfy some minimum requirement on size or complexity for patterns to attract a newborn's interest. For instance, the total V1 activity may need to be above a certain threshold for the newborn to pay any attention to a pattern. If so, the HLISSOM procedure

for deriving a behavioral response from the model response (section 4.3) would need to be modified to include this constraint.

Overall, these results provide strong computational support for the speculation of Johnson and Morton (1991) that the newborn could simply be responding to a three-dot facelike configuration, rather than performing sophisticated face detection. Internally generated patterns provide an account for how such "innate" machinery can be constructed during prenatal development, within a system that can also learn postnatally from visual experience.

5.2 Real face images

Researchers testing newborns with schematic patterns often assume that the responses to schematics are representative of responses to real faces. However, no experiment has yet tested that assumption by comparing real faces to similar but non-facelike controls. Similarly, no previous computational model of newborn face preferences has been tested with real images. HLISSOM makes such tests practical, which is an important way to determine the significance of the behavioral data. The examples in figure 6 show that the prenatally trained HLISSOM model works remarkably well as a face detector for natural images.

The face detection performance of the map was tested quantitatively using two image databases: a set of 150 images of 15 adult males without glasses, photographed at the same distance against blank backgrounds (Achermann, 1995), and a set of 58 non-face images of various natural scenes (National Park Service, 1995). The face image set contained two views of each person facing forwards, upwards, downwards, left, and right; figure 6*a* shows an example of a frontal view. Each natural scene was presented at 6 different size scales, for a total of 348 non-face presentations. Overall, the results indicated very high face detection performance: the FSA responded to 91% (137/150) of the face images, but to only 4.3% (15/348) of the natural scenes.

Because the two sets of real images were not closely matched in terms of lighting, backgrounds, and distances, it is important to consider the actual response patterns to be sure that the differences in the overall totals are genuine. The FSA responded with activation in the location corresponding to the center of the face in 88% (132/150) of the face images. At the same time, the FSA had spurious responses in 27% (40/150) of the face images, i.e. responses in locations other than the center of the face. Nearly half of the spurious responses were from less-selective neurons along the outer border of the model FSA (see figure 3*h*); these can be ignored. Most of the remaining spurious responses resulted from a genuine V1 eye or mouth response plus V1 responses to the hair or jaw outlines. Such responses would actually serve to direct attention to the general region of the face, although they would not pinpoint the precise center. For the natural scenes, most of the responses were from the less-selective neurons along the border of the FSA, and those responses can again be ignored. The remainder were in image regions that coincidentally had a triangular arrangement of three contour-rich areas, surrounded by smooth shading.

In summary, the FSA responds to most top-lit human faces of about the right size, signaling their location in the visual field. It does not respond to most other stimuli, except when they contain accidental three-dot patterns. The model predicts that human newborns will have a similar pattern of responses in the face-selective regions of the visual cortex.

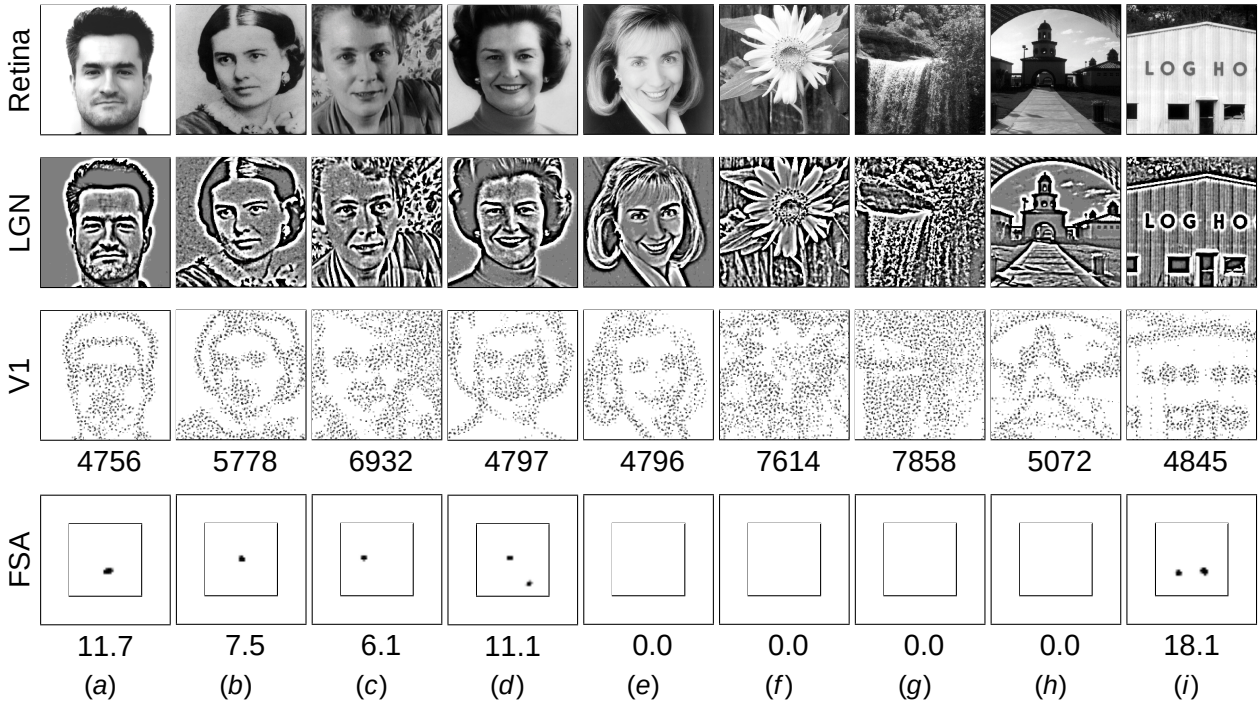


Figure 6: **Model response to natural images.** The top row shows a sample set of photographic images. The LGN, V1, and FSA responses of the model to these images are displayed as in figures 4 and 5. The FSA is indeed activated at the correct location for most top-lit faces of the correct size and orientation (e.g. *a*–*d*). Image (*a*) is an example of one frontal view from the the Achermann (1995) database. For this database, 88% of the faces resulted in FSA activity in the correct location. Just as important, the network is not activated for most natural scenes (*f*–*g*) and man-made objects (*h*). For example, the FSA responded to only 4.3% of 348 presentations of landscapes and other natural scenes from the National Park Service (1995). Besides human faces, the FSA responds to patterns causing a V1 response similar to that of a three-dot arrangement of contours (*d*,*i*), including related patterns such as dog and monkey faces (not shown). Response is low to images where hair or glasses obscures the borders of the eyes, nose, or mouth, and to front-lit downward-looking faces, which have low V1 responses from nose and mouth contours (*e*). The model predicts that newborns would show similar responses if tested. Credits: (*a*) copyright 1995 Bernard Achermann, (*b*–*e*) public domain; (*f*–*i*) copyright 1999–2001, James A. Bednar.

6 Discussion and future work

The HLISSOM simulations show that internally generated patterns and self-organization can together account for newborn face detection. The novel contributions of the model are: (1) providing a rigorous, computational test of the Johnson and Morton (1991) three-dot conceptual model for newborn schematic face preferences, (2) demonstrating that preferences for three-dot patterns result in significant preference for faces in natural images, (3) proposing a specific, plausible way that a system for face detection could be constructed in the developing brain, (4) demonstrating how a single system can both exhibit genetically encoded preferences in the newborn and later learn from real faces, (5) demonstrating how a hierarchy of brain areas could self-organize and function, at the level of columnar maps and their connections, (6) showing how the reported newborn pattern preferences can result from specific feature preferences at the different levels of the hierarchy, and

(7) showing that the effects of internally generated patterns should be considered when explaining genetically influenced behavior. The key idea is that pattern generation can combine the strengths of both genetic pre-specification and input-driven learning into a system more powerful than the sum of its parts. The following subsections discuss results with other training patterns and postnatal learning, relate the HLISSOM results to those from other models, and discuss future research directions.

6.1 Effect of the training pattern shape

How crucial is the specific pattern of three dots for the model, given that other patterns might also provide face selectivity? For instance, Viola and Jones (2001) found that a horizontal bar or a pair of eye spots can detect faces well in engineering applications. In separate work, we tested the effectiveness of a variety of training pattern shapes at detecting faces in real images and for replicating the psychological data (Bednar, 2002). These simulations used nine different training patterns, including dots, bars, top-heavy triangles, and circles.

For all training patterns, test faces that matched the overall training pattern size gave higher responses than did other natural images. Thus simply matching the size of typical faces gives some degree of face selectivity in real images. However, different training patterns led to large differences in responses to the schematic patterns. Of all training patterns tested, the three-dot pattern provided the most face selectivity in overhead lighting and the best match to the experimental data, and thus it is a good choice for pattern-generation simulations. It leads to strong predictions that can be matched with experimental data and contrasted with other models.

All of these tests used upright training patterns, because the Simion et al. (1998a) and Simion et al. (1998b) studies suggest that the orientation of face patterns is important even in the first few days of life. In the areas that generate training patterns, such a bias might be due to anisotropic lateral connectivity, which would cause spontaneous patterns in one part of the visual field to suppress those below them. On the other hand, tests with the youngest infants (less than one day) have not yet found orientation biases (Johnson et al., 1991). Thus the experimental data is also consistent with a model that assumes unoriented patterns prenatally, followed by rapid learning from upright faces. Such a model would be more complicated to simulate and describe than the one presented here, but could use a similar architecture and learning rules.

Importantly, not all possible training patterns result in face preferences. For example, if only a single low-level feature like eye size is matched between the training and testing images, the model shows no face preference. Together, the results with different patterns show that the pattern shape is important, but that the shape does not need to be strictly controlled to provide face preferences.

6.2 Postnatal learning

The work reported in this paper focuses on understanding the origin of newborn face preferences, but a crucial motivation for the model is that the same components can learn from real faces and not just prenatal patterns. To understand postnatal learning computationally, we have trained a map with internally generated patterns, then trained it on images of real faces in front of natural image

backgrounds (Bednar, 2002; Bednar and Miikkulainen, 2002).

Preliminary results from these simulations suggest that (1) prenatal training leads to faster and more robust face learning after birth, compared to a system exposed only to environmental stimuli, (2) the decline in infants' response to schematic patterns in the periphery after one month (section 2.4) may simply result from learning real faces, not from a shift to a separate system, (3) learning of real faces may explain why young infants prefer their mother's face, as well as why the preference disappears when the face outline is masked, and (4) delayed maturation of the fovea may explain why facial preferences appear in central vision later than in peripheral vision. These results are consistent with the postnatal disappearance of newborn face preferences that in part inspired the Johnson and Morton (1991) model, while also providing a simple account for why infants of all ages appear to be able to learn faces.

6.3 Comparison to other models

Predictions from the HLISSOM pattern-generation model differ strongly from those of other recent models. These predictions can be tested by future experiments.

One easily tested prediction is that newborn face preferences should not depend on the precise shape of the face outline. The Acerra et al. (2002) model (section 2.2) makes the opposite prediction, because in that model preferences arise from precise spacing differences between the external border and the internal facial features. Results from the HLISSOM simulations also suggest that newborns will have a strong preference for real faces (e.g. in photographs), while we argue that the Acerra et al. model predicts only a weak preference for real faces, if any.

Experimental evidence to date cannot yet decide between the predictions of these two models. For instance, Simion et al. (1998a) did not find a significant schematic face preference in newborns 1–6 days old without a contour surrounding the internal features, which is consistent with the Acerra et al. model. However, the same study concluded that the shape of the contour "did not seem to affect the preference" for the patterns, which would not be consistent with the Acerra et al. model. As discussed earlier, newborns may not require any external contour, as the pattern generation model predicts, until they have had postnatal experience with faces. Future experiments with younger newborns should compare model and newborn preferences between schematic patterns with a variety of border shapes and spacings. These experiments will either show that the border shape is crucial, as predicted by Acerra et al., or that it is unimportant, as predicted by the pattern-generation model.

The predictions of the pattern-generation model also differ from those of the Simion et al. (2001) "top-heavy" model (reviewed in section 2.3). The top-heavy model predicts that any face-sized border that encloses objects denser at the top than the bottom will be preferred over similar schematic patterns. The pattern-generation model predicts instead that a pattern with three dots in the typical symmetrical arrangement would be preferred over the same pattern with both eye dots pushed to one side, despite both patterns being equally top-heavy. These two models represent very different explanations of the existing data, and thus testing such patterns should offer clear support for one model over the other.

On the other hand, many of the predictions of the fully trained pattern-generation model im-

plemented in HLISSOM are similar to those of the CONSPEC model proposed by Johnson and Morton (1991). In fact, the reduced HLISSOM face preference network in section 6.1 (which does not include V1) can be seen as the first CONSPEC system to be implemented computationally, along with a concrete proposal for how such a system could be constructed during prenatal development (Bednar and Miikkulainen, 2000). The primary functional difference between the trained HLISSOM network and CONSPEC is that in HLISSOM only neurons located in cortical visual areas respond selectively to faces in the visual field.

Whether newborn face detection is mediated cortically or subcortically has been debated extensively, yet no clear consensus has emerged from behavioral studies (Simion et al., 1998a). If future brain imaging studies do discover face-selective visual neurons in subcortical areas in newborns, HLISSOM will need to be modified to include such areas. Yet the key principles would remain the same, because subcortical regions also appear to be organized from internally generated activity (Wong, 1999). Thus experimental tests of the pattern-generation model vs. CONSPEC should focus on how the initial system is constructed, and not where it is located.

6.4 Future directions

The pattern-generation hypothesis can be verified directly by imaging the actual activity patterns produced during REM sleep. Very recent advances in imaging hardware have made limited measurements of this type possible (Rector et al., 1997). It may soon be feasible to test the assumptions and predictions of the pattern generation model in developing animals.

Why would evolution have favored a pattern-generation approach over fixed hard-wiring or strictly general learning? One important reason is that the genome can encode the desired outcome, i.e. three-blob responses, independently of the actual architecture of the face processing hardware. As shown by Dailey and Cottrell (1999), even when the underlying cortical architecture is specified only in domain-general terms, specific regions can still become primarily devoted to face processing. The divide-and-conquer strategy of pattern generation would allow such an architecture to evolve and develop independently of the function, which could potentially enable the rapid evolution and development of more complex adaptive systems.

In terms of information processing, pattern generation combined with self-organization may represent a general way to solve difficult problems like face detection and recognition. Rather than meticulously specifying the final, desired individual, the specification need only encode a process for constructing an individual through interaction with its environment. The result can seamlessly combine the full complexity of the environment with *a priori* information about the desired function of the system. In the future, this approach could be used for engineering complex artificial systems for real-world tasks, e.g. handwriting recognition, speech recognition, and language processing.

7 Conclusion

The pattern-generator model of newborn face preferences can explain how newborns can show a face preference at birth yet remain able to learn faces at all ages. The predictions of the model can be tested experimentally, in order to better understand the development of face perception. The simulations and future human experiments will also help explain how environmental and genetic influences can together construct complex, adaptive systems.

A Parameter values

Most of the model parameters were calculated using the scaling equations from Kelkar et al. (2001). Of those below, only the γ parameters and the sigmoid thresholds were set empirically for this study. For the LGN and V1, the values for these free parameters were set so that the output from each neural sheet would be in a useful range for the following sheet over the course of training. For the highest level (FSA), the free parameters were set so that the response would be nearly binary, and therefore an unambiguous criterion for the presence of a facelike pattern on the input. Apart from such threshold parameters, small parameter variations produce roughly equivalent results. The specific parameter values for each subcortical and cortical region follow.

For the LGN, σ_c was 0.4 and σ_s was 1.6. The afferent scale γ_{A} was 10.6, and the upper and lower thresholds of the sigmoid were 0.14 and 1.0, respectively.

For V1, the initial lateral excitation radius was 3.6 and was gradually decreased to 1.5. The lateral inhibitory radius of each neuron was 8, and inhibitory connections whose strength was below 0.01 were pruned away at $10,000$ iterations. The lateral inhibitory connections were initialized to a Gaussian profile with $\sigma = 17$, and the lateral excitatory connections to a Gaussian with $\sigma = 2.8$, with no connections outside the nominal circular radius. Initially, the divisive normalization strength γ_{N} was zero, the afferent scale γ_{A} was 1.0, and the lateral excitation γ_{E} and inhibition strength γ_{I} were both 0.9. The γ values were constant during V1 training and were gradually increased over the course of FSA training to $\gamma_{\mathrm{N}} = 4$, $\gamma_{\mathrm{A}} = 3.25$, $\gamma_{\mathrm{E}} = 1.2$, and $\gamma_{\mathrm{I}} = 1.4$. The learning rate α_{A} was gradually decreased from 0.0035 to 0.00075, α_{E} from 0.059 to 0.0029, and α_{I} was a constant 0.00088. The lower and upper thresholds of the sigmoid were increased from 0.08 to 0.5 and from 0.63 to 0.86, respectively over the course of V1 training; the lower threshold was then decreased to 0.22 over the course of FSA training. The number of iterations for which the lateral connections were allowed to settle at each training iteration was initially 9, and was increased to 13.

For the FSA, the initial lateral excitation radius was 6.3 and was gradually decreased to 1.5. The lateral inhibitory radius of each neuron was 15.8, and inhibitory connections whose strength was below 0.0027 were pruned away at $20,000$ iterations. The lateral inhibitory connections were initialized to a Gaussian profile with $\sigma = 33$, and the lateral excitatory connections to a Gaussian with $\sigma = 4.9$. The lateral excitation γ_{E} was 0.9 and the inhibition strength γ_{I} was 0.9. Initially, the divisive normalization strength γ_{N} was zero, the afferent scale γ_{A} was 3.0, and the lateral excitation γ_{E} and inhibition strength γ_{I} were both 0.9. These values were gradually changed over the course of FSA training to $\gamma_{\mathrm{N}} = 9$, $\gamma_{\mathrm{A}} = 10.6$, $\gamma_{\mathrm{E}} = 0.4$, and $\gamma_{\mathrm{I}} = 0.6$. The learning rate α_{A} was gradually decreased from 0.0001 to 0.000022, α_{E} from 0.025 to 0.013 and α_{I} was a constant 0.003. The lower and upper thresholds of the sigmoid were increased from 0.1 to 0.81 and from 0.65 to 0.88, respectively over the course of training. The number of iterations for which the lateral connections were allowed to settle at each training iteration was initially 9, and was increased to 13.

Acknowledgments

This research was supported in part by the National Science Foundation under grant #IIS-9811478 and by the National Institutes of Mental Health under Human Brain Project grant 1R01-MH66991. We thank Cornelius Weber, Lisa Kaczmarczyk, an anonymous reviewer, and especially Mark H. Johnson for critical comments on earlier drafts. Software, demonstrations, color figures, and related publications are available at `http://www.cs.utexas.edu/users/nn/`.

References

Acerra, F., Burnod, Y., and de Schonen, S. (2002). Modelling aspects of face processing in early infancy. *Developmental Science*, 5(1):98–117.

Achermann, B. (1995). Full-faces database. Copyright 1995, University of Bern, all rights reserved. http://iamwww.unibe.ch/ fkiwww/Personen/achermann.html.

Banks, M. S., and Salapatek, P. (1981). Infant pattern vision: A new approach based on the contrast sensitivity function. *Journal of Experimental Child Psychology*, 31(1):1–45.

Bednar, J. A. (2002). *Learning to See: Genetic and Environmental Influences on Visual Development*. PhD thesis, Department of Computer Sciences, The University of Texas at Austin. Technical Report AI-TR-02-294.

Bednar, J. A., and Miikkulainen, R. (1998). Pattern-generator-driven development in self-organizing models. In Bower, J. M., editor, *Computational Neuroscience: Trends in Research, 1998*, 317–323. New York: Plenum.

Bednar, J. A., and Miikkulainen, R. (2000). Self-organization of innate face preferences: Could genetics be expressed through learning?. In *Proceedings of the 17th National Conference on Artificial Intelligence*, 117–122. Cambridge, MA: MIT Press.

Bednar, J. A., and Miikkulainen, R. (2002). Neonatal learning of faces: Interactions between genetic and environmental inputs. In Gray, W., and Schunn, C., editors, *Proceedings of the 24th Annual Conference of the Cognitive Science Society*, 107–112. Mahwah, NJ: Erlbaum.

Bednar, J. A., and Miikkulainen, R. (2003). Self-organization of spatiotemporal receptive fields and laterally connected direction and orientation maps. *Neurocomputing*. In press.

Blasdel, G. G. (1992). Orientation selectivity, preference, and continuity in monkey striate cortex. *Journal of Neuroscience*, 12:3139–3161.

Bushnell, I. W. R., Sai, F., and Mullin, J. T. (1989). Neonatal recognition of the mother's face. *The British Journal of Developmental Psychology*, 7:3–15.

Callaway, C. W., Lydic, R., Baghdoyan, H. A., and Hobson, J. A. (1987). Pontogeniculooccipital waves: Spontaneous visual system activity during rapid eye movement sleep. *Cellular and Molecular Neurobiology*, 7(2):105–49.

Choe, Y. (2001). *Perceptual Grouping in a Self-Organizing Map of Spiking Neurons*. PhD thesis, Department of Computer Sciences, The University of Texas at Austin, Austin, TX. Technical Report AI01-292.

Choe, Y., and Miikkulainen, R. (2000). Contour integration and segmentation with self-organized lateral connections. Technical Report AI2000-286, Department of Computer Sciences, The University of Texas at Austin.

Cohen, L. B. (1998). An information-processing approach to infant perception and cognition. In (Simion and Butterworth, 1998), 277–300.

Constantine-Paton, M., Cline, H. T., and Debski, E. (1990). Patterned activity, synaptic convergence, and the NMDA receptor in developing visual pathways. *Annual Review of Neuroscience*, 13:129–154.

Dailey, M. N., and Cottrell, G. W. (1999). Organization of face and object recognition in modular neural network models. *Neural Networks*, 12(7):1053–1074.

Datta, S. (1997). Cellular basis of pontine ponto-geniculo-occipital wave generation and modulation. *Cellular and Molecular Neurobiology*, 17(3):341–365.

de Boysson-Bardies, B., editor (1993). *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*. Dordrecht; Boston: Kluwer.

de Haan, M. (2001). The neuropsychology of face processing during infancy and childhood. In Nelson, C. A., and Luciana, M., editors, *Handbook of Developmental Cognitive Neuroscience*. MIT Press.

de Schonen, S., Mancini, J., and Leigeois, F. (1998). About functional cortical specialization: The development of face recognition. In (Simion and Butterworth, 1998), 103–120.

Dugovic, C., and Turek, F. W. (2001). Similar genetic mechanisms may underlie sleep-wake states in neonatal and adult rats. *Neuroreport*, 12(14):3085–3089.

Eglen, S. (1997). *Modeling the Development of the Retinogeniculate Pathway*. PhD thesis, University of Sussex at Brighton, Brighton, UK. Technical Report CSRP 467.

Erwin, E., Obermayer, K., and Schulten, K. (1995). Models of orientation and ocular dominance columns in the visual cortex: A critical comparison. *Neural Computation*, 7(3):425–468.

Feller, M. B., Wellis, D. P., Stellwagen, D., Werblin, F. S., and Shatz, C. J. (1996). Requirement for cholinergic synaptic transmission in the propagation of spontaneous retinal waves. *Science*, 272:1182–1187.

Field, T. M., Cohen, D., Garcia, R., and Greenberg, R. (1984). Mother–stranger face discrimination by the newborn. *Infant Behavior and Development*, 7:19–25.

Gauthier, I., and Nelson, C. A. (2001). The development of face expertise. *Current Opinion in Neurobiology*, 11(2):219–224.

Goren, C. C., Sarty, M., and Wu, P. Y. (1975). Visual following and pattern discrimination of face-like stimuli by newborn infants. *Pediatrics*, 56(4):544–549.

Haith, G. L. (1998). *Modeling Activity-Dependent Development in the Retinogeniculate Projection*. PhD thesis, Department of Psychology, Stanford University, Stanford, CA.

Hirsch, J. A., and Gilbert, C. D. (1991). Synaptic physiology of horizontal connections in the cat's visual cortex. *Journal of Neuroscience*, 11:1800–1809.

Johnson, M. H., Dziurawiec, S., Ellis, H., and Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, 40:1–19.

Johnson, M. H., and Morton, J. (1991). *Biology and Cognitive Development: The Case of Face*

Recognition. Oxford, UK; New York: Blackwell.

Jouvet, M. (1998). Paradoxical sleep as a programming system. *Journal of Sleep Research*, 7(Suppl 1):1–5.

Jouvet, M. (1999). *The Paradox of Sleep: The Story of Dreaming*. Cambridge, MA: MIT Press.

Kanwisher, N., McDermott, J., and Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17(11):4302–4311.

Kelkar, A., Bednar, J. A., and Miikkulainen, R. (2001). Modeling large cortical networks with growing self-organizing maps. Technical report, Department of Computer Sciences, The University of Texas at Austin. Technical Report AI-00-285.

Kleiner, K. (1993). Specific vs. non-specific face-recognition device. In (de Boysson-Bardies, 1993), 103–108.

Leinekugel, X., Khazipov, R., Cannon, R., Hirase, H., Ben-Ari, Y., and Buzsáki, G. (2002). Correlated bursts of activity in the neonatal hippocampus in vivo. *Science*, 296(5575):2049–2052.

Linsker, R. (1986). From basic network principles to neural architecture: Emergence of orientation columns. *Proceedings of the National Academy of Sciences, USA*, 83:8779–8783.

Lippe, W. R. (1994). Rhythmic spontaneous activity in the developing avian auditory system. *Journal of Neuroscience*, 14(3):1486–1495.

Marks, G. A., Shaffery, J. P., Oksenberg, A., Speciale, S. G., and Roffwarg, H. P. (1995). A functional role for REM sleep in brain maturation. *Behavioural Brain Research*, 69:1–11.

Miikkulainen, R., Bednar, J. A., Choe, Y., and Sirosh, J. (1997). Self-organization, plasticity, and low-level visual phenomena in a laterally connected map model of the primary visual cortex. In Goldstone, R. L., Schyns, P. G., and Medin, D. L., editors, *Perceptual Learning*, vol. 36 of *Psychology of Learning and Motivation*, 257–308. San Diego, CA: Academic Press.

Miller, K. D. (1994). A model for the development of simple cell receptive fields and the ordered arrangement of orientation columns through activity-dependent competition between ON- and OFF-center inputs. *Journal of Neuroscience*, 14:409–441.

Mondloch, C. J., Lewis, T. L., Budreau, D. R., Maurer, D., Dannemiller, J. L., Stephens, B. R., and Kleiner-Gathercoal, K. A. (1999). Face perception during early infancy. *Psychological Science*, 10(5):419–422.

Morris, J. S., Ohman, A., and Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. *Proceedings of the National Academy of Sciences, USA*, 96(4):1680–1685.

National Park Service (1995). Image database. Http://www.freestockphotos.com/NPS.

O'Donovan, M. J. (1999). The origin of spontaneous activity in developing networks of the vertebrate nervous system. *Current Opinion in Neurobiology*, 9:94–104.

Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., and Fabre-Grenet, M. (1995). Mother's face recognition by neonates: A replication and an extension. *Infant Behavior and Development*, 18:79–85.

Pascalis, O., and Slater, A. (2001). The development of face processing in infancy and early childhood: Current perspectives. *Infant and Child Development*, 10(1/2):1–2.

Rector, D. M., Poe, G. R., Redgrave, P., and Harper, R. M. (1997). A miniature CCD video camera for high-sensitivity light measurements in freely behaving animals. *Journal of Neuroscience Methods*, 78(1-2):85–91.

Rodieck, R. W. (1965). Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Research*, 5(11):583–601.

Rodman, H. R. (1994). Development of inferior temporal cortex in the monkey. *Cerebral Cortex*, 4(5):484–98.

Roffwarg, H. P., Muzio, J. N., and Dement, W. C. (1966). Ontogenetic development of the human sleep-dream cycle. *Science*, 152:604–619.

Rolls, E. T. (1990). The representation of information in the temporal lobe visual cortical areas of macaques. In Eckmiller, R., editor, *Advanced Neural Computers*, 69–78. New York: Elsevier.

Shaffery, J. P., Roffwarg, H. P., Speciale, S. G., and Marks, G. A. (1999). Ponto-geniculo-occipital-wave suppression amplifies lateral geniculate nucleus cell-size changes in monocularly deprived kittens. *Brain Research. Developmental Brain Research*, 114(1):109–119.

Shatz, C. J. (1990). Impulse activity and the patterning of connections during CNS development. *Neuron*, 5:745–756.

Shatz, C. J. (1996). Emergence of order in visual system development. *Proceedings of the National Academy of Sciences, USA*, 93:602–608.

Simion, F., and Butterworth, G., editors (1998). *The Development of Sensory, Motor and Cognitive Capacities in Early Infancy: From Perception to Cognition*. East Sussex, UK: Psychology Press.

Simion, F., Macchi Cassia, V., Turati, C., and Valenza, E. (2001). The origins of face perception: Specific versus non-specific mechanisms. *Infant and Child Development*, 10(1/2):59–66.

Simion, F., Valenza, E., and Umiltà, C. (1998a). Mechanisms underlying face preference at birth. In (Simion and Butterworth, 1998), 87–102.

Simion, F., Valenza, E., Umiltà, C., and Dalla Barba, B. (1998b). Preferential orienting to faces in newborns: A temporal-nasal asymmetry. *Journal of Experimental Psychology: Human Perception and Performance*, 24(5):1399–1405.

Sirosh, J., and Miikkulainen, R. (1994). Cooperative self-organization of afferent and lateral connections in cortical maps. *Biological Cybernetics*, 71:66–78.

Sirosh, J., and Miikkulainen, R. (1996). Self-organization and functional role of lateral connections and multisize receptive fields in the primary visual cortex. *Neural Processing Letters*, 3:39–

48.

Sirosh, J., Miikkulainen, R., and Bednar, J. A. (1996). Self-organization of orientation maps, lateral connections, and dynamic receptive fields in the primary visual cortex. In Sirosh, J., Miikkulainen, R., and Choe, Y., editors, *Lateral Interactions in the Cortex: Structure and Function*. Austin, TX: The UTCS Neural Networks Research Group. Electronic book, ISBN 0-9647060-0-8, http://www.cs.utexas.edu/users/nn/web-pubs/htmlbook96.

Slater, A. (1993). Visual perceptual abilities at birth: Implications for face perception. In (de Boysson-Bardies, 1993), 125–134.

Slater, A., and Kirby, R. (1998). Innate and learned perceptual abilities in the newborn infant. *Experimental Brain Research*, 123(1-2):90–94.

Swindale, N. V. (1996). The development of topography in the visual cortex: A review of models. *Network – Computation in Neural Systems*, 7:161–247.

Tovée, M. J. (1998). Face processing: Getting by with a little help from its friends. *Current Biology*, 8:R317–R320.

Valenza, E., Simion, F., Cassia, V. M., and Umiltà, C. (1996). Face preference at birth. *Journal of Experimental Psychology: Human Perception and Performance*, 22(4):892–903.

Viola, P., and Jones, M. (2001). Robust real-time object detection. In *Second International Workshop on Statistical and Computational Theories of Vision – Modeling, Learning, Computing, and Sampling*.

Wallace, M. T., McHaffie, J. G., and Stein, B. E. (1997). Visual response properties and visuotopic representation in the newborn monkey superior colliculus. *Journal of Neurophysiology*, 78(5):2732–2741.

Walton, G. E., Armstrong, E. S., and Bower, T. G. R. (1997). Faces as forms in the world of the newborn. *Infant Behavior and Development*, 20(4):537–543.

Walton, G. E., and Bower, T. G. R. (1993). Newborns form "prototypes" in less than 1 minute. *Psychological Science*, 4:203–205.

Weliky, M., Kandler, K., Fitzpatrick, D., and Katz, L. C. (1995). Patterns of excitation and inhibition evoked by horizontal connections in visual cortex share a common relationship to orientation columns. *Neuron*, 15:541–552.

Wong, R. O. L. (1999). Retinal waves and visual system development. *Annual Review of Neuroscience*, 22:29–47.

Wong, R. O. L., Meister, M., and Shatz, C. J. (1993). Transient period of correlated bursting activity during development of the mammalian retina. *Neuron*, 11(5):923–938.

Yuste, R., Nelson, D. A., Rubin, W. W., and Katz, L. C. (1995). Neuronal domains in developing neocortex: Mechanisms of coactivation. *Neuron*, 14(1):7–17.