An Incremental Hebbian Learning Model of the Primary Visual Cortex with Lateral Plasticity and Real Input Patterns

Thomas Burger and Elmar W. Lang

Institut für Biophysik und physikalische Biochemie, Universität Regensburg, D-93040 Regensburg, Germany

Z. Naturforsch. **54c,** 128–140 (1999); received August 3/September 25, 1998

Neural Network, Hebbian Learning, Visual Cortex, Lateral Plasticity, Self-Organization

We present a simplified binocular neural network model of the primary visual cortex with separate ON/OFF-pathways and modifiable afferent as well as intracortical synaptic couplings. Random as well as natural image stimuli drive the weight adaptation which follows Hebbian learning rules stabilized with constant norm and constant sum constraints. The simulations consider the development of orientation and ocular dominance maps under different conditions concerning stimulus patterns and lateral couplings. With random input patterns realistic orientation maps with $\pm 1/2$-vortices mostly develop and plastic lateral couplings self-organize into mexican hat type structures on average. Using natural greyscale images as input patterns, realistic orientation maps develop as well and the lateral coupling profiles of the cortical neurons represent the two point correlations of the input image used.

1. Introduction

The primary visual cortex is one of the most intensely studied areas of the mammalian cerebral cortex and its functional architecture is known in considerable detail now (Hubel and Wiesel, 1977). The early stages of visual processing are devoted to extracting relatively simple features like oriented contrast contours (lines and bars) from the visual input. One type of feature detectors frequently encountered in the visual cortex of cats and monkeys are orientation selective simple cells. Their localized receptive fields are subdivided into few elongated subfields of alternating ON or OFF response to small light spot stimuli (Jones and Palmer, 1987). Several investigations using optical imaging techniques showed these simple cells to be organized into piecewise continuous orientation preference maps along the cortical surface containing $\pm 1/2$-vortices, where orientation preferences change by $\pm 180°$ along a closed path around the vortex center (Blasdel, 1992; Bonhoeffer and Grinvald, 1991). Also there are cortical regions, so-called iso-orientation domains, where all neurons have similar orientation preferences, linear zones, where they change smoothly, and fractures, which connect singularities and mark rapid changes of orientation preferences.

Another feature of cortical neurons is their being driven preferably by input from one eye, known as ocular dominance. These largely monocular neurons form ocular dominance stripes in monkey cortex, or patches in cat cortex (LeVay *et al.*, 1975, 1978). Both map structures, orientation preference and ocular dominance, are superimposed with the pinwheel singularities generally located in the centers of ocular dominance domains (Erwin *et al.*, 1995; Obermayer and Blasdel, 1993, 1997).

Although these cortical structures are well characterized, their activity-dependent self-organization during prenatal and early postnatal development is still not well understood yet (Hubel and Wiesel, 1962, 1968; LeVay *et al.*, 1978; Singer, 1988). The role of neural activity in the formation of ocular dominance has been well characterized (Stryker and Harris, 1986), but its role for the emergence and structure of orientation specificity maps is still in disput (Weliky and Katz, 1997; Crair *et al.*, 1998). Ocular dominance and orientation preference maps seem to be present at the time of eye opening. Yet most neurons are still weakly orientation-selective at this stage and visual experience, hence structured stimulus pat-

Reprint requests to Prof. Lang.
Fax: 0049–941/943–2479.
E-mail: elmar.lang@biologie.uni-regensburg.de

0939–5075/99/0100–0128 $ 06.00

terns, seems necessary for receptive field and orientation map structure refinement to mature adult levels.

In the past several computational models have been deviced to investigate the mechanisms controlling the development of these cortical feature maps (Swindale, 1996).

One group of models, the so-called SOM-(Self-Organizing-Map-)models based on Kohonen's competitive learning algorithm, produced orientation and ocular dominance maps comparing favourably with results of neurobiological experiments (Kohonen, 1987; Obermayer *et al.*, 1992; Erwin *et al.*, 1995). But frequently assumptions of SOM-models like normalized feature vectors as input are not or only hardly justifiable neurobiologically.

Another group of models are correlation-based Hebbian learning (CBL) models (von der Malsburg, 1973; Linsker, 1986; Miller *et al.*, 1989; Miller, 1990, 1994; Kammen and Yuille *et al.*, 1988; Yuille, 1989; Stetter *et al.*, 1993, 1994, 1995). Following ideas put forward by von der Malsburg, Linsker presented a linear multilayer feed forward neural network using a Hebbian learning rule and uncorrelated white noise as input (Linsker, 1986). Only neurons from within a restricted area of one layer were connected to single postsynaptic neurons of the subsequent layer. In Linkser's model stable orientation selective cells developed. A thorough account of Linsker's model has been given by MacKay and Miller (MacKay and Miller, 1990). Kammen and Yuille (1988) introduced a learning rule, which performed gradient descent on an energy function and led to orientation selective neurons as a result of a spontaneous symmetry breaking process. Yuille further implemented a modified type of Oja's learning rule (Oja, 1982; Yuille *et al.*, 1989) where cells are performing principal component analysis to their input stimuli. Stetter *et al.* (1993) extended these deterministic network models to stochastic models with randomly scattered size, eccentricity and position of receptive fields and presented a systematic investigation of the parameter space in these models.

Combining the ideas of von der Malsburg (1973) and Linsker (1986) Miller (Miller *et al.*, 1989, 1990, 1994) deviced correlation-based Hebbian learning models of ocular dominance and orientation colum formation. Correlations determine receptive field structure, intracortical interactions determine projective field structure and both determine the cortical map. Ocular dominance columns then form through a correlation-based competition between left-eye and right-eye inputs, while orientation columns form through a competition between ON-center and OFF-center inputs. These models are still not completely adequate to account for certain details of the cortical map structure. Also there is no parameter regime where both cortical maps could develop simultaneously.

Stetter *et al.* (1994), following the seminal paper of Miller (Miller *et al.*, 1989, 1990) on ocular dominance column formation, presented a two-layer neural network model for the formation of cortical orientation columns with congruent ON- and OFF-ganglion cell center-surround contrast filters, localized and strongly overlapping receptive fields of cortical cells and spatially oscillating intracortical feedback connections. Receptive field formation and cortical map self-organization were driven by input activity correlation functions resulting from random input stimuli convolved with strongly overlapping center-surround input filters and controlled by a Hebbian learning rule with multiplicative constraints. Within the parameter regime of the most stable bilobed – (0,1) in the nomenclature of Linsker (1986) – receptive field structures the model yielded orientation selective cortical simple cells with structured receptive fields and realistic orientation maps containing ± 1-vortices, however, as a result of overlap maximization and phase locking of neighboring bilobed receptive fields. Experimentally observed $\pm 1/2$-pinwheel vortices could be obtained only within a parameter regime with highly labil easily deformable (1,0) receptive fields. Within a binocular version of the model ocular dominance patches did only emerge, if there was an unbalance of ON- and OFF-ganglion cell responses (Stetter *et al.*, 1995).

Besides these "linear" Hebbian learning models (in fact many of them introduce non-linearities indirectly), non-linear models like the BCM model (Bienenstock *et al.*, 1982) were proposed also. They differ from the linear models in the information they extract from the input in that they also rely on higher order input correlations than just second order correlations. The BCM model has been shown to reproduce rather successfully many characteristics of cortical neuronal responses to

random and natural image visual stimuli (Shouval *et al.*, 1997). However, it has so far been concerned with single neuron properties only.

Almost all previous network models used non-modifiable lateral couplings in cortex. But recent experimental findings demonstrated a rather substantial plasticity of lateral connections in the primary visual cortex of monkeys and cats (Callaway and Katz, 1990; Gilbert and Wiesel, 1989; Gilbert *et al.*, 1990; Katz *et al.*, 1989; Katz and Callaway, 1992). Exploration of CBL-models with modifiable lateral couplings is thus essential to an understanding of how such plastic lateral interactions may affect the development and structures of cortical receptive fields and cortical feature maps. Sirosh and Miikkulainen investigated lateral plasticity within their LISSOM-model, which is an extension of competitive SOM-models (Sirosh and Miikkulainen, 1994, 1997). Receptive field profiles and cortical feature maps have not been examined so far, however. Furthermore the difficulties concerning a neurobiological justification of the model assumptions remain. A correlation-based Hebbian learning model with lateral plasticity has not been described in the literature so far.

As visual stimuli random input patterns were used in the majority of these developmental neural network models. Also idealized correlation functions of input activities were substituted for a direct training with real input patterns, whether random or structured. As cortical maps seem to emerge prenatally but mature to adult levels postnatally only, the respective role of structured versus random input patterns is still obscure and needs to be clarified. A two-phase learning model was recently deviced discussing the disparity tuning of cortical cells (Berns *et al.*, 1993).

In this study we investigate a neural network model with incremental Hebbian learning of modifiable afferent and lateral synaptic couplings. Theoretical input correlation functions are replaced by real input stimulus patterns during training. Both random input patterns and natural images are examined as visual stimuli.

2. The Model

2.1. The network architecture

In figure 1 the network architecture is presented. The binocular model possesses two retinal input layers, the neurons of which project their output from within localized receptive fields to separate cell layers with ON- and OFF-center-surround contrast filters. These will somewhat loosely be termed LGN cells in the following. The afferent activity of these preprocessing units converge onto neurons of the output layer which may correspond to cortical simple cells. Their localized receptive fields are centered on retinotopically corresponding sites of the underlying ON/OFF-layers. The binocular model is thus given a fixed topography.

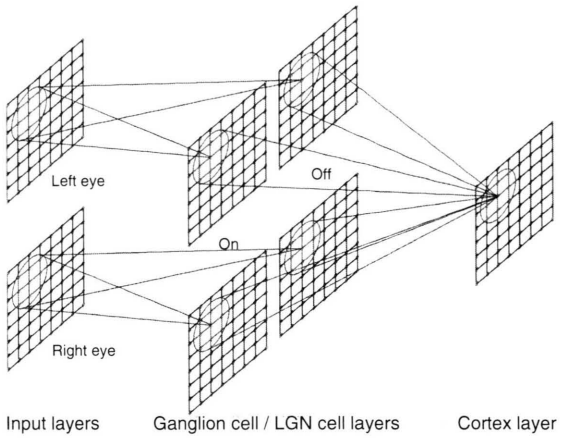

Fig. 1. Architecture of the neural network. There are three levels consisting of input layers, ganglion cell / LGN cell layers, and a cortex layer. The input cells project their outputs to the ON- and OFF-LGN cells, which are connected with cortical neurons.

The activities of the input layer neurons are confined to non-negative values within the interval [0.0,1.0]. Two types of input patterns are considered. Patterned input is not essential for map formation (Crair *et al.*, 1998). Hence, simulating a prenatal developmental stage, uncorrelated input activities correspond to equally distributed random numbers fluctuating about a mean activity 0.5 and embody spontaneous uncorrelated ganglion cell activities found by several investigators (Braitenberg and Schütz, 1991). Correlated retinal ganglion cell activity waves (Meister *et al.*, 1991; Wong *et al.*, 1993, 1996) have recently been found also. They are probably too large in their characteristic wave length to be able to drive the development of orientation specificity in simple cells (Miller, 1994). Also Weliky and Katz (1997) recently showed that their artificially correlated activity patterns did not interfere with these sponta-

neous retinal waves in blocking normal development of orientation specificity. Therefore these patterned spontaneous retinal waves have not been considered as input stimuli in this study. Because of the retinotopic network architecture correlated activities of second layer neurons result due to overlapping contrast filter profiles. Map refinement and persistence of feature selectivity need visual experience. Hence, simulating early postnatal developmental stages, natural greyscale images as highly correlated stimulus patterns were presented to the input layers of the network.

The connections between input- and LGN-cells are fixed. Their spatial organization forms the LGN receptive field profiles which resemble mexican-hat-like contours with a center-surround antagonism (Jones and Palmer, 1987). These DOG profiles are normalized to yield a vanishing response to constant illumination. Furthermore there are four separate layers of ON/OFF contrast filters, where an OFF filter is simply the inverse of an ON filter. All ON- or OFF-LGN contrast filters are identical. A determining parameter of their mexican hat profile is the total center radius, which defines the border between center and surround response and, together with the cortical receptive field size, determines the type of RF-structure obtained (Stetter *et al.*, 1993).

The afferent connections converging onto cortical neurons are plastic. Their non-negative weights, corresponding to exclusively excitatory afferents, are denoted as $w^{\alpha,\beta}(\vec{r}_y,\vec{r}_x)$, where $\alpha \in$ {left, right}, $\beta \in$ {on, off}. $\vec{r}_y$ and $\vec{r}_x$ describe the locations of cells within the cortex layer and the LGN layers, respectively. The adaptation of these synaptic weights to the ensemble of input patterns during pre- and postnatal development is controlled by learning rules to be specified later. Their spatial structure determines the receptive field organization of the cortical cells, though the latter would have to be convolved with the mexican hat profile of the underlying contrast filters (Stetter, 1993).

Within the cortex layer also plastic lateral interactions exist within a finite radius around each neuron. Both inhibitory and excitatory lateral synaptic couplings are considered. The variable $v_E(\vec{r}_y,\vec{r}_{y'})$ denotes the excitatory synaptic coupling strength between a postsynaptic cortical neuron at $\vec{r}_y$ and a presynaptic cortical neuron at $\vec{r}_{y'}$ and v_I denotes a corresponding inhibitory connection. These coupling strengths are both non-negative and cannot change sign during adaptation.

The excitatory and inhibitory lateral couplings exist within circular regions with radii r_{LE} and r_{LI}, respectively, around any given cortical neuron. Also $r_{LE} \leq r_{LI}$ will be assumed throughout. Furthermore the variable v_{EI} is introduced allowing a more convenient evaluation of the lateral coupling structure:

$$v_{EI}(\vec{r}_y,\vec{r}_{y'}) = \begin{cases} v_E(\vec{r}_y,\vec{r}_{y'}) - v_I(\vec{r}_y,\vec{r}_{y'}) & \text{for } |\vec{r}_{y'}-\vec{r}_y| \leq r_{LE}, \\ -v_I(\vec{r}_y,\vec{r}_{y'}) & \text{for } r_{LE} < |\vec{r}_{y'}-\vec{r}_y| \leq r_{LI}. \end{cases} \quad (1)$$

It accounts for the dominance of one type of lateral coupling over the other in the close neighborhood of any given neuron and for a general inhibition from other cortical neurons further apart.

Including these effective lateral connections, the postsynaptic activity of a cortical neuron is given by

$$y(\vec{r}_y, t) = \left\{ \sum_{\alpha,\beta} \left[\sum_{\vec{r}_x} \left(A(\vec{r}_x-\vec{r}_y) w^{\alpha,\beta}(\vec{r}_y,\vec{r}_x,t) x^{\alpha,\beta}(\vec{r}_x,t) \right) \right] \right\} + \\ + \sum_{\vec{r}_{y'}} (v_{EI}(\vec{r}_y,\vec{r}_{y'},t) y(\vec{r}_{y'}-\vec{r}_y,t-1)), \quad (2)$$

where $x^{\alpha,\beta}(\vec{r}_x,t)$ denotes the LGN cell activities resulting from a convolution of input activity values and contrast filters. The arbor function A models the overlap of afferent axonal arborazation with the dendritic tree of the cortical target neuron. The finite size of the afferent arbor is characterized by a Fermi-function with an arbor radius r_{A0}

$$A(\vec{r}_y,\vec{r}_x) = \begin{cases} \dfrac{1}{1+e^{\frac{|\vec{r}_x-\vec{r}_y|-r_{A0}}{0.3}}} & \text{for } |\vec{r}_x-\vec{r}_y| \leq r_{A0}, \\ 0 & \text{for } |\vec{r}_x-\vec{r}_y| > r_{A0}. \end{cases} \quad (3)$$

This function is radially symmetric and depends only on $|\vec{r}_x - \vec{r}_y|$. Though cortical receptive field size varies with eccentricity, the cortex layer of our network model represents only a small part of the cortex. Hence a constant arbor radius is certainly a valid assumption. The first term in equation (2) represents the weighted sum of afferent inputs from the LGN and the second term represents the modulation of cortical activity through lateral input from other cortical neurons.

All neurons are assumed to operate in the linear regime of their activation function, i.e. the post-

synaptic activity of a neuron between its threshold value and its saturation value rises linearly.

2.2. The learning rules

For the afferent synaptic weights w a Hebbian learning rule with a multiplicative constraint (Yuille decay) is used (Yuille *et al.*, 1989; Stetter *et al.*, 1993) ensuring convergence to a stable fixpoint.

$$w^{\alpha,\beta}(\vec{r}_y,\vec{r}_x,\, t+1) = w^{\alpha,\beta}(\vec{r}_y,\vec{r}_x,\, t) + \Delta w^{\alpha,\beta}(\vec{r}_y,\vec{r}_x,\, t \rightarrow t+1) \quad (4)$$

and

$$\Delta w^{\alpha,\beta}(\vec{r}_y,\vec{r}_x,\, t \rightarrow t+1) = \eta_{heb}\, y(\vec{r}_y,t)\, x^{\alpha,\beta}(\vec{r}_x,t) - \\ - \eta_{dec}\left[\sum_{\vec{r}_x,\alpha,\beta} (w^{\alpha,\beta}(\vec{r}_y,\vec{r}_x,t))^2\right] w^{\alpha,\beta}(\vec{r}_y,\vec{r}_x,t), \quad (5)$$

where $t = 0,1,2,\ldots$ denotes sequential update cycles and η_{heb} (Hebbian term) and η_{dec} (decay term) represent appropriately chosen learning rates. The first term in equation (5) represents Hebb's postulat, that synchronous pre- and postsynaptic activities strengthen their respective synapse. The second term constrains the growth of synaptic weights and represents some form of forgetting during the learning process. This learning rule has been shown to perform PCA on its inputs and, with properly chosen parameters, yields well segregated receptive field structures. Even with the inclusion of lateral couplings the learning dynamics have been shown to converge to stable fixed points resembling the principal components of the related correlation matrix C of LGN activities (Stetter *et al.*, 1994). Hence, the resulting cortical receptive field structure is still determined mainly by the ratio of the total center radius of the LGN contrast filter to the arbor radius of the cortical cell.

The intracortical synaptic weights are also modified according to a Hebb rule and are normalized to unity after each update cycle (constant sum (CS) constraint). Hence they are given by

$$v_X(\vec{r}_y,\vec{r}_{y'},\, t+1) = \Sigma_X \cdot \frac{v_X(\vec{r}_y,\vec{r}_{y'},t) + \mu_X \cdot \eta_{lat,heb}\, y(\vec{r}_y,t)\, y(\vec{r}_{y'},t)}{\Sigma_{\vec{r}_{y'}} [v_X(\vec{r}_y,\vec{r}_{y'},t) + \mu_X\, \eta_{lat,heb}\, y(\vec{r}_y,t) y(\vec{r}_{y'},t)]}, \quad (6)$$

where $X \in \{E,I\}$. $\eta_{lat,heb}$ is the learning rate for the Hebbian modification of the lateral synaptic weights. The factors Σ_E and Σ_I denote the sum of excitatory or inhibitory intracortical synaptic coupling strengths of any cortical neuron, respectively:

$$\Sigma_X = \sum_{\vec{r}_{y'}} (v_X(\vec{r}_y,\vec{r}_{y'})). \quad (7)$$

These sums are independently kept constant at every cortical neuron during a simulation. They introduce a competition between the synapses for the total amount of coupling strengths.

3. Results and Discussion

3.1. Random input patterns

The activity-dependent prenatal self-organization of receptive field structures and cortical feature maps is simulated by training the network with uncorrelated random input patterns. After every update cycle new random input patterns are presented. The results obtained with modifiable lateral couplings will be contrasted with results obtained with fixed lateral interactions. In CBL-models the ratio of the total center radius of the input filters to the arbor radius of the cortical cells is the sole determining parameter concerning the segregation of afferents into ON/OFF-subfields. It has been chosen in the simulations to yield a structured receptive field of any cortical neuron corresponding to a bilobed (0,1) profile. This is the most stable eigenmode in CBL-models with multiplicative constraints and has been found to be most robust against any perturbations from noise or intracortical couplings which modulate the output activity (Stetter *et al.*, 1993, 1994). Accordingly with random pattern stimuli, vanishing between-eye correlations and no intracortical interaction the receptive fields resulting from incremental Hebbian learning models very closely resemble the eigenmodes of the covariance matrix resulting from CBL-models, though with less perfect symmetry.

3.1.1. Non-plastic lateral synaptic weights

Intracortical interactions control the self-organization process of the orientation map and determine its layout. Non-plastic lateral interactions are implemented as spatially oscillating mexican-hat-like functions and cortical activities are iterated five times between sequential updates to stabilize the resulting activity pattern. Lateral couplings of moderate strengths hardly disturb the afferent weight segregation into pure bilobed (0,1) cortical receptive field structures and hardly alter the orientation tuning curves either. Still these lateral

couplings suffice to promote the emergence of well organized feature maps. Their structure begins to appear in the simulations only after completion of the receptive field organization into ON/OFF-subfields. The scale of the resulting map structures, as given by the dominant wave vector of the corresponding Fourier transform, is determined largely by the range of the lateral interaction (Miller, 1990, 1994; Stetter *et al.*, 1994).

The mature orientation maps are characterized by extended iso-orientation domains, linear zones and saddle points and exhibit pinwheel singularities with $\pm 1/2$- and ± 1-vorticity as well (see Fig. 2). Neurons in the vortex centers are only weakly tuned for orientation if at all. Note that CBL-models using multiplicative constraints and theoretical correlation functions of input activities instead of using explicit stimulus patterns yield in the same parameter regime ± 1 vortices exclusively due to the 360° symmetrie of the bilobed receptive fields and the overlap maximization mechanism driving the map formation process (Stetter *et al.*, 1994). Only in a parameter regime yielding unperturbed (1,0) fields which became strongly deformed and phase shifted by the action of the lateral coupling were $\pm 1/2$-vortices observable (Stetter *et al.*, 1995). This suggests that even direct training with explicit stimulus patterns introduces sufficient perturbation into the learning process to suppress the overlapp maximization mechanism by strong fluctuations induced with incremental learning modes. Note also that CBL-models with subtractive constraints and theoretical input activity correlation functions always yield $\pm 1/2$ vortices (Miller, 1994). The reason for this behaviour might be, that a subtractive constraint does not stabilize the principal component, but traps the synapse functions at hard synaptic boundaries (Miller and MacKay, 1994).

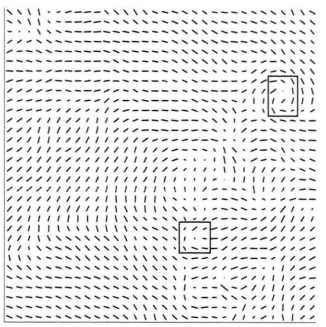

Fig. 2. Binocular orientation maps obtained using non-plastic lateral synaptic weights. Each line in the map belongs to a cortical neuron. The line orientation denotes the direction of spatial frequency vector of that sinus wave, which causes the maximum response, and the line length denotes the extent of selectivity (for details see Stetter *et al.* 1994). Note, that the vector direction is perpendicular to the optimal stimulus orientation, a vertical line thus corresponds to 0° preference angle, whereby angles are measured clock-wise. Some vortices are marked.

The distribution of orientation preferences of the cortical neurons forming the orientation map shows some structure indicating an over-representation of certain orientations. Periodic boundary conditions could alleviate this asymmetry somewhat, though not completely. As both retinae see different random patterns corresponding to a vanishing contralateral correlation of input activities, the two monocular orientation maps and the corresponding binocular map are not congruent, but show the same characteristic structures. This is in accord with the fact that receptive fields of newborn kittens are misaligned by as much as 30° (Pettigrew, 1974). But it seems to be in disagreement with recent findings of almost identical monocular orientation maps in young kittens raised according to the reverse suturing protocol (Gödecke and Bonhoeffer, 1996). According to recent results by Crair *et al.* (1998) this may be due to a delayed development of ipsilateral versus contralateral afferents with the latter providing sort of a template orientation map. This feature has not been implemented in any learning model so far. It is further possible that the optical imaging technique used by Gödecke and Bonhoeffer does not possess a high enough spatial resolution which would allow the detection a moderate receptive field misalignment of a cortical cell.

Most cortical neurons showed some eye preference in the simulations even though CBL-models afford unstructered, hence pure ON or OFF receptive fields for ocular dominance to develop (Miller *et al.*, 1989). The distribution of eye preferences segregated into clustered domains with pure binocular neurons located at the domain boundaries. The resulting ocular dominance map (Fig. 3(a)) had a patchy appearance as found in cats. The ODI histogram peaks around ± 0.6, hence most neurons prefer input from one eye over the other. Most pinwheel vortices of the ori-

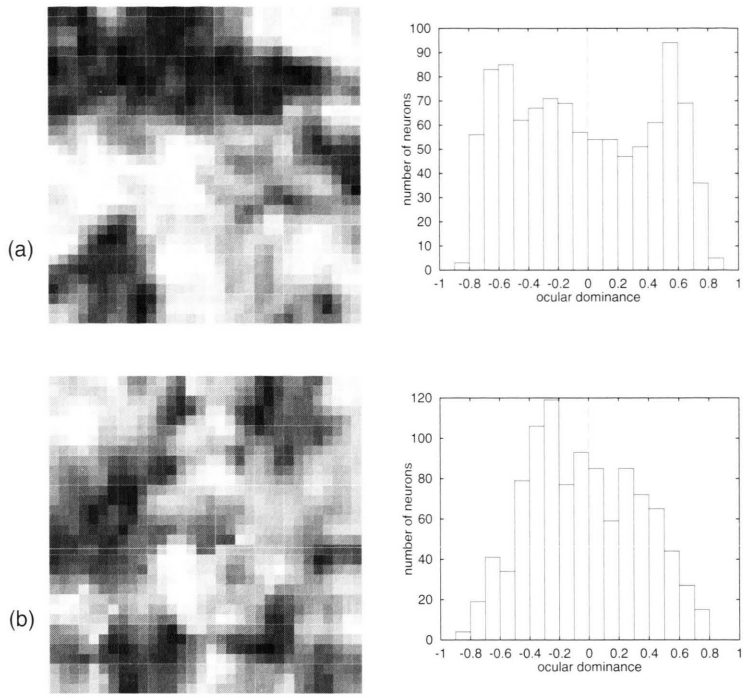

Fig. 3. Ocular dominance maps and related eye preference histograms obtained using either non-plastic (part (a)) or plastic (part (b)) lateral synaptic weights. The pixel grey value denotes the neuron's ocular dominance value (for details see Stetter *et al.* 1995), whereby white (OD = −1) corresponds to the strongest left monocular and black (OD = +1) to the strongest right monocular dominance value. Totally binocular neurons thus correspond to an intermediate grey value (OD = 0).

entation map were located at the borders of the ocular dominance domains, however, whereas optical imaging studies of feature maps in adult mammals show vortices in the center of ocular dominance patches (Obermayer and Blasdel, 1993). However, no corresponding investigations of newborn animals yet exist. Furthermore the patterns of activity of the cortical neurons in response to these random visual stimuli formed well localized stripe-like blobs.

3.1.2. Plastic lateral synaptic weights

Next plastic lateral synaptic weights have been considered. At first the radius of excitatory lateral couplings has been chosen roughly three times shorter than the range of inhibitory interactions, i.e. ($r_{LE} = 2.7$, $r_{LI} = 8.0$). The radii chosen correspond to the center radius and the maximal range, respectively, of the fixed mexican-hat-like interaction function discussed in the last section. All coupling strengths have been initialized with random values according to a constant sum constraint $\Sigma_E = \Sigma_I$ thus balancing the total coupling strength of both types of interaction on any cortical neuron. Hebbian learning has been performed with the CS rule (equation (6)) applied to the intracortical couplings and the Yuille rule (equation (5)) to the afferent couplings.

The dynamics of the weak lateral interaction strengths did not interfere with the weight dynamics of the afferent couplings. Hence the latter again segregated into only slightly disturbed bilobed receptive field profiles resulting in almost unaltered orientation tuning curves. When averaged over an ensemble of random stimulus patterns a mexican hat type lateral coupling structure resulted (Fig. 4). Note, that the range of lateral couplings never exceeded the range of a hypercolumn. Though specific long-range couplings connecting similarly tuned neurons in adjacent hypercolumns were not considered.

The orientation maps (Fig. 5) obtained do not differ much from those resulting from fixed mexican-hat-like lateral couplings except that the number of ± 1/2-vortices dominated over few ± 1-vortices still observed occasionally in the simulations. This seems in accord with the further degradation of any high symmetry structures like pure bilobed receptive fields or rotationally invariant mexican-hat like lateral coupling structures. Again the distribution of orientation preference angles

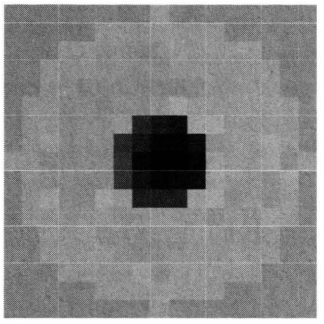

Fig. 4. Lateral coupling structures of a cortical cell resulting from simulations with plastic lateral interactions ($r_{LE} < r_{LI}$) averaged over many input patterns. White denotes the strongest inhibitory, black the strongest excitatory formal synaptic weights v_{EI}. A vanishing lateral synaptic weight is denoted by a intermediate grey value. The central pixel grey value denotes the self-coupling of the neuron considered.

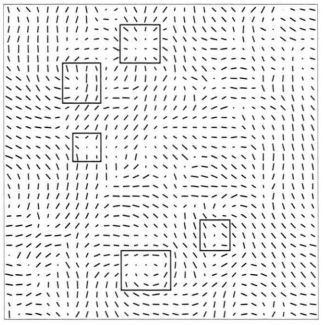

Fig. 5. Binocular orientation map obtained using plastic lateral synaptic weights ($r_{LE} < r_{LI}$).

consistently showed some structure which, however, fluctuated due to the incessant presentation of random stimulus activity patterns. This structure certainly results from a too small map size in the simulations.

Clustered ocular dominance domains (Fig. 3(b)) formed with smaller domain size than before yielding a larger number of boundary areas and a concomitant larger number of binocular neurons in the maps. Pinwheel vortices again were located at ocular dominance domain boundaries mostly.

Additional simulations have been performed with unbalanced total excitatory and inhibitory coupling strengths. The simulations showed no qualitative changes of the resulting structures but instead a reduced or enhanced rate of maturation of the feature maps was observed in case of $\Sigma_E < \Sigma_I$ and $\Sigma_E > \Sigma_I$ respectively.

In order not to bias a mexican hat type lateral interaction structure, simulations with equal excitatory and inhibitory lateral coupling radii ($r_{LE} = r_{LI} = 8.0$) were performed, too. Thereby the inhibitory synaptic weights must be adapted by an anti-Hebbian rule, otherwise $r_{LE} = r_{LI}$ in conjunction with $\Sigma_E = \Sigma_I$ causes all formal synaptic weights v_{EI} to vanish. When averaged over many input patterns a decrease of synaptic weights with increasing distance was observed. Lateral synaptic couplings of the same kind formed diffuse clusters with most excitatory connections being concentrated to the immediate vicinity of the cell (Fig. 6). Overall, however, the range of excitatory and inhibitory couplings remained largely unchanged. Considering the output activity distribution, no blob formation could be seen. Instead blob-like cortical activity patterns disassembled into tiny patches of high activity. Hence Hebbian learning could not enforce excitatory lateral couplings to concentrate to the center of the cell's coupling areas. On average, no mexican-hat-like coupling

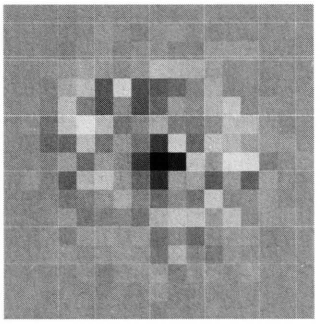

Fig. 6. Lateral coupling structure of a cortical cell arising with plastic lateral interactions (r_{LE}=r_{LI}) and averaged over many input patterns.

structure thus appeared. Lateral interactions therefore did not act as an ordering force and highly distorted feature maps resulted with many pinwheel vortices and very small iso-orientation domains and ocular dominance patches.

3.2. Structured input patterns

These simulations intended to investigate a postnatal development of synaptic couplings driven by structured stimulus patterns. From a greyscale picture of size (256×256) pixels showing a natural scene with thicket and foliage (Fig. 7(a)) smaller images matching the size of the input layers of the model network were extracted at random. Each stimulus pattern was presented to both input layers of the network to account for correlations resulting from fixation during binocular vision. A parameter Δ_x was further introduced to effect a horizontal shift of both input patterns between the two input layers. $\Delta_x = 0$ means maximal correlation of input activities at retinotopically corresponding sites within both input layers and $\Delta_x > 0$ takes account of the fact that the two eyes of mammals are not looking at exactly the same scene. In addition this parameter allows to simulate misalignment of both monocular receptive fields as occurs in case of strabismic vision.

With the parameters of the center-surround input filters chosen as before segregation of ON/OFF-afferents between rather than within cortical receptive fields occured. This result corroborates the finding that the principal component of natural images often corresponds to an unstructured receptive field (Liu and Shouval, 1994; Shouval and Liu, 1996; Hancock *et al.*, 1992) signaling the presence of many low frequency structures of constant activity in the images.

Modifiable lateral couplings with inhibitory couplings being of longer range than excitatory couplings and a strong contralateral correlation of monocular stimuli resulted in complementary arrangements of monocular ON/OFF-patches. Hence, no orderly orientation maps emerged. Also no clear mexican hat type lateral coupling structure appeared even after averaging over many input patterns. The emergent averaged lateral coupling profiles were elliptically deformed and did not resemble a mexican hat type structure (Fig. 8). Realistic binocular orientation maps could be restored only with both types of lateral couplings being of equal range ($r_{LE} = r_{LI} = 8.0$). Monocular non-segregated as well as binocular bilobed receptive field profiles developed. The related binocular orientation map exhibited $\pm 1/2$ vortices exclusively, which, furthermore, occurred in pairs of opposing polarity (Fig. 9). The lateral connections were reciprocal and all self-couplings were excitatory. The individual lateral coupling structures of any two cortical neurons were identical, if they were connected excitatorily, and were complementary, if they were connected inhibitorily. Further these corresponding lateral coupling structures were shifted according to the different topographic locations of both cortical neurons. The elliptic fading out structure of the lateral synaptic weights corresponds directly with the central area of the two point correlation matrix of the contrast filtered input image patches (Fig. 7(b)) averaged over the whole image. Therefore the coupling profile represents exactly the two point correlations of the input image used.

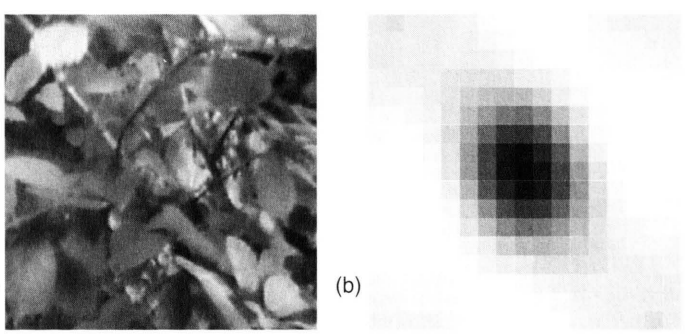

Fig. 7. Structured input image. (a) natural scene input image, (b) two point correlation matrix of the contrast filtered input image averaged over the whole image.

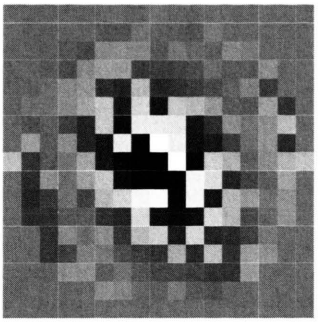

Fig. 8. Lateral coupling structure of the cortical cell at (24,16) in figure 9 obtained using plastic lateral interactions (r_{LE}=r_{LI}) and structured input images and averaged over many input patterns. The retinotopic position of a cell in the cortical map is given by a tupel (line, column), whereby (0,0) denotes the upper left neuron in a map.

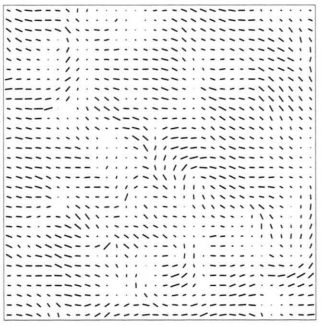

Fig. 9. Binocular orientation map resulting with plastic lateral synaptic weights (r_{LE}=r_{LI}) and structured input images.

3.3. Sequential combination of random and structured input pattern phases

So far either random input patterns or natural image stimuli have been used exclusively. Now two phase simulations are described, which considered a transition from prenatal to postnatal activity-dependent developmental processes. The simulations started with random input patterns. After corresponding structures have stabilized structured input patterns (natural images) were presented. Plastic lateral synaptic weights were implemented with r_{LE} = 2.7, r_{LI} = 8.0 and $\Sigma_E = \Sigma_I$. During the postnatal developmental phase a horizontal shift Δ_x = 10 of both retinal images was employed.

Two phase simulations with natural image stimuli presented during postnatal development following a prenatal phase with random input patterns showed most monocular receptive field structures to switch immediately to non-segregated profiles whereas binocular bilobed structures survived and self-organized into realistic orientation maps with ± 1/2-vortices exclusively. Also average lateral coupling structures v_{EI} resembled spatially oscillating mexican-hat like contours though less well defined than in case of random input patterns only. In summary most structures which evolved during prenatal development became erased by structures emerging from postnatal development driven by structured stimuli. This is certainly a consequence of the linearity of the learning model and the cortical activation function which prevents synaptic couplings from saturating during prenatal development. These results suggest that non-linear models will be necessary to preserve structural features which formed during the prenatal phase and would then be refined and modulated only during postnatal maturation.

4. Conclusions

We have investigated in this study the simultaneous and activity-dependent development of modifiable afferent and intracortical synaptic couplings into receptive field structures and cortical feature maps within the realm of incremental Hebbian learning models. The latter differ from correlation-based Hebbian learning models in that real stimuli are presented to the input layer instead of using theoretical ensemble averaged correlation functions inserted into the learning rule as introduced by von der Malsburg (1973) and thoroughly studied by Miller (1990, 1994) and Stetter *et al.* (1994, 1995). The model architecture was characterized by separate monocular ON/OFF-pathways converging onto any binocular cortical neuron. Random as well as natural image stimuli presented to both retinae have been considered to account for pre- and postnatal phases of development. Also varying degrees of contra-

lateral correlations have been considered to model misalignment of monocular receptive fields and strabismus.

Correlation-based Hebbian learning of afferent weights with multiplicative constraints extracts the principal components of the input pattern. These eigenmodes represent the receptive fields of the corresponding neurons. Which principal mode results is solely determined by the ratio of the center radius of the input contrast filters to the arbor radius of the orientation filters. Due to the inherent rotational symmetry these eigenmodes are invariant under rotation, too. A rotationally symmetric and spatially oscillating lateral coupling guiding the self-organization of these receptive fields into cortical maps then can only lead to orientation singularities with ± 1 vorticity because of overlap maximization of neighboring receptive fields.

Incremental Hebbian learning of afferent weights with real input patterns can introduce sufficient fluctuations into the map formation process for $\pm 1/2$ vortices to appear in case of fixed spatially oscillating couplings. Modifiable lateral couplings adjusted according to a Hebb-like rule with constant sum constraints further degraded any symmetry and introduced additional fluctuations into the weight dynamics to increase the number of $\pm 1/2$ vortices even further. When averaged over an ensemble of random input patterns with rotationally invariant power spectra a mexican-hat like lateral coupling structure again developed if the range of excitatory couplings was chosen smaller than the range of inhibitory couplings and both couplings were randomly initialized. With natural images as stimulus patterns, the power spectra of which showed directions of prefered correlations, such a coupling structure could be obtained only if the range of excitatory and inhibitory couplings were chosen equally wide from the beginning. The lateral coupling structure then reflected the two-point statistics of the input image, hence showed some elliptic fading out in agreement with the non-rotationally symmetric power spectrum. The additional degradation of any inherent symmetry lead to $\pm 1/2$ vortices exclusively which furthermore occured in pairs mostly.

In summary, incremental Hebbian learning with multiplicative constraints can lead to realistic orientation maps with correct pinwheel vorticities whereas correlation-based Hebbian learning cannot. Also, ensemble averaged, mexican-hat like lateral coupling structures result from both random as well as natural image visual stimuli. Cortical activity distributions thereby play a decisive role. Distinctive activity blobs must form for the excitatory lateral couplings to shrink to the center of any cortical neurons lateral coupling field – a necessary prerequisite for any spatially oscillating lateral coupling structure to emerge.

Though, in contrast to CBL-models, ocular dominance patches developed simultaneously to the formation of orientation maps, a major deficiency of the model stems from the fact, that regions of high feature variation are strongly correlated locally in both maps. Hence orientation vortices are mostly located at the borders of ocular dominance patches in obvious contradiction to experimental feature maps obtained with optical imaging techniques. It is currently not obvious how to remedy this deficiency.

Acknowledgements

Financial support of this work by the Friedrich-Ebert-Stiftung, Bonn, Germany, and the Deutsche Forschungsgemeinschaft, Germany, is greatfully acknowledged.

Berns G. S., Dayan P. and Sejnowski T. J. (1993), A correlational model for the development of disparity selectivity in visual cortex that depends on prenatal and postnatal phases. Proc. Natl. Sci. USA **90**, 8277–8281.

Bienenstock E. L., Cooper L. N. and Munro P. W. (1982), Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. J. Neurosci. **2**, 32–48.

Blasdel G. G. (1992), Differential imaging of ocular dominance and orientation selectivity in monkey striate cortex. J. Neurosci. **12**, 3115–3138.

Bonhoeffer T. and Grinvald A. (1991), Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. Nature **353**, 429–431.

Braitenberg V. and Schütz A. (1991), Anatomy of the Cortex: Statistics and Geometry. Springer New York.

Callaway E. M. and Katz L. C. (1990), Emergence and refinement of clustered horizontal connections in cat striate cortex. J. Neurosci. **10**, 1134–1153.

Crair M. C., Deda C. G. and Stryker M. P. (1998), The role of visual experience in the development of columns in cat visual cortex. Science **279**, 566–570.

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**, 425–468.

Gilbert C. D., Hirsch J. A. and Wiesel T. N. (1990), Lateral interactions in visual cortex. Cold Spring Harbor Symposia on Quantitative Biology **55**, 663–677.

Gilbert C. D. and Wiesel T. N. (1989), Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. J. Neurosci. **9**, 2432–2442.

Gödecke I. and Bonhoeffer T. (1996), Development of identical orientation maps for two eyes without common visual experience. Nature **379**, 251–254.

Hancock P. J. B., Baddeley R. J. and Smith L. S. (1992), The principal components of natural images. Network **3**, 61–70.

Hubel D. H. and Wiesel T. N. (1962), Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. **160**, 106–154.

Hubel D. H. and Wiesel T. N. (1968), Receptive fields and functional architecture of monkey striate cortex. J. Physiol. **195**, 215–243.

Hubel D. H. and Wiesel T. N. (1977), Functional architecture of macaque monkey visual cortex. Proc. R. Soc. Lond. B **198**, 1–59.

Jones J. P. and Palmer L. A. (1987), The two-dimensional spatial structure of simple receptive fields in cat striate cortex. J. Neurophysiol. **58**, 1187–1211.

Kammen D. M. and Yuille A. L. (1988), Spontaneous symmetry-breaking energy functions and the emergence of orientation selective cortical cells. Biol. Cybern. **59**, 23–31.

Katz L. C. and Callaway E. M. (1992), Development of local circuits in mammalian visual cortex. Ann. Rev. Neurosci. **15**, 31–56.

Katz L. C., Gilbert C. D. and Wiesel T. N. (1989), Local circuits and ocular dominance columns in monkey striate cortex. J. Neurosci. **9**, 1389–1399.

Kohonen T. (1987), Self-Organization and Associative Memory. Springer New York.

LeVay S., Hubel D. H. and Wiesel T. N. (1975), The pattern of ocular dominance columns in macaque visual cortex revealed by a reduced silver stain. J. Comp. Neurol. **159**, 559–576.

LeVay S., Stryker M. P. and Shatz C. J. (1978), Ocular dominance columns and their development in layer IV of the cat's visual cortex. J. Comp. Neurol. **179**, 223–44.

Linsker R. (1986), From basic networks principles to neural architecture (series). Proc. Natl. Sci. USA} **83**, 7508–12, 8390–4, 8779–8783.

Linsker R. (1990), Designing a sensory processing system: what can be learned from principal component analysis?. Proc. of the Int. Joint Conf. on Neural Networks. IJCNN, Washington (DC), USA.

Liu Y. and Shouval H. (1994), Localized principal components of natural images – an analytic solution. Network **5**, 317–325.

MacKay D. J. C. and Miller K. D. (1990), Analysis of Linsker's simulations of Hebbian rules. Neural Computation **2**, 173–187.

Meister M., Wong R. O. L., Baylor D. A. and Shatz C. J. (1991), Synchronous bursts of action potentials in ganglion cells of the developing mammalian retina. Science **17**, 939–943.

Miller K. D., Keller J. B. and Striker M. P. (1989), Ocular dominance column development: analysis and simulation. Science **245**, 605–615.

Miller K. D. (1990), Correlation-based models of neural development. In: Neuroscience and Connectionist Theory (ed Gluck M A and Rumelhart D E) Hillsdale, NJ: Lawrence Erlbaum Associates, pp 267–353.

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. J. Neurosci. **14**, 409–441.

Miller K. D. and MacKay D. J. C. (1994), The role of constraints in Hebbian learning. Neural Computation **6**, 100–126.

Obermayer K. and Blasdel G. G. (1993), Geometry of orientation and ocular dominance columns in monkey striate cortex. J. Neurosci. **13**, 4114–4129.

Obermayer K. and Blasdel G. G. (1997), Singularities in primate orientation maps. Neural Computation **9**, 555–575.

Obermayer K., Blasdel G. G. and Schulten K. (1992), Statistical-mechanical analysis of self-organization and pattern formation during the development of visual maps. Phys. Rev. A **45**, 7568–7589.

Oja E. (1982), A simplified neuron model as a principal component analyzer. J. Math. Biol. **15**, 267–273.

Pettigrew J. D. (1974), The effect of visual experience on the development of stimulus specificity by kitten cortical neurons. J. Physiol. **237**, 49–74.

Shouval H. and Liu Y. (1996), Principal component neurons in a realistic visual environment. Network **7**, 501–515.

Shouval H., Intrator N. and Cooper L. N. (1997), BCM network develops orientation selectivity and ocular dominance in natural scene environment. Vision Research **37**, 3339–3342.

Singer W. (1988), Activity-dependent self-organization of the mammalian visual cortex. In: Models of the Visual Cortex (ed Rose D and Dobson V G). John Wiley & Sons, Chichester, Great Britain, pp 123–136.

Sirosh J. and Miikkulainen R. (1994), Cooperative self-organization of afferent and lateral connections in cortical maps. Biol. Cybern. **71**, 65–78.

Sirosh J. and Miikkulainen R. (1997), Topographic receptive fields and patterned lateral interaction in a self-organization model of the primary visual cortex. Neural Computation **9**, 577–594.

Stetter M., Lang E. W. and Müller A. (1993), Emergence of orientation selective simple cells simulated in deterministic and stochastic neural networks. Biol. Cybern. **68**, 465–476.

Stetter M., Müller A. and Lang E. W. (1994), Neural network model for the coordinated formation of orientation preference and orientation selectivity maps. Phys. Rev. E **50**, 4167–4181.

Stetter M., Kussinger M., Schels A., Seeger E. and Lang E. W. (1995), Self-organization of cortical receptive fields and columnar structures in a Hebb-trained neural network. Lecture Notes in Computer Science **930**, 37–44.

Stryker M. P. and Harris W. A. (1986), Binocular impulse blockade prevents formation of ocular dominance columns in the cat's visual cortex. J. Neurosci. **6**, 2117–2133.

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

von der Malsburg C. (1973), Self-organization of orientation sensitivity cells in the striate cortex. Kybernetik **14**, 85–100.

Weliky M. and Katz L. C. (1997), Disruption of orientation tuning in visual cortex by artificially correlated neuronal activity. Nature **386**, 680–685.

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**, 923–938.

Wong R. O. L. and Oakley D. M. (1996), Changing patterns of spontaneous bursting activity of on and off retinal ganglion cells during development. Neuron **16**, 1087–1095.

Yuille A. L., Kammen D. M. and Cohen D. S. (1989), Quadrature and the development of orientation selective cortical cells by Hebb rules. Biol. Cybern. **61**, 183–194.