INVITED PAPER Special Section on Nonlinear Dynamics —Adaptive, Learning and Neural Systems—

A Model for the Development of the Spatial Structure of Retinotopic Maps and Orientation Columns

Klaus OBERMAYER†, Helge RITTER†† *and* Klaus J. SCHULTEN†, *Nonmembers*

SUMMARY Topographic maps begin to be recognized as one of the major computational structures underlying neural computation in the brain. They provide dimension-reducing projections between feature spaces that seem to be established and maintained under the participation of selforganizing, adaptive processes. In this contribution, we investigate how well the structure of such maps can be replicated by simple adaptive processes of the kind proposed by Kohonen[15]. We will particularly address the important issue, how the dimensionality of the input space affects the spatial organization of the resulting map.

key words: *topographic map, self-organizing process, dimension-reducing projection, feature space, retinotopic location*

1. Introduction

It seems that topographic maps reflect important information processing strategies realized in the brain to match abstract feature spaces onto the spatial structure of its "parallel hardware"[22]. This makes it an intriguing question, both from the viewpoint of neural computation and from the viewpoint of basic neuroscience, to consider to what extent such structures can be understood in terms of simple pattern formation processes within the underlying substrate, or, stated in a more abstract way, can be understood in terms of rules for elementary local adaptation steps. Consequently, research on this question has formed an essential part of neural network theory from the past up to the present (see e.g. Refs. (38), (2), (1), (15), (8), (25), (20), (36)).

The main characteristic of topographic maps is a macroscopic, spatial pattern of the tuning properties of (at least a subset of) the neurons within some cortical region. It is generally accepted that for the most part these patterns of tuning properties are not established genetically, but instead evolve during ontogenesis in a self-organizing process (see e.g. Refs. (3), (13)) and show in many cases a considerable degree of adaptability even in adult life[21].

Frequently the spatial order of the tuning properties of neurons within a topographic map reflects just the spatial origin of the afferent signals, such as e.g. the location of a tactile stimulus on the skin. However, there are also maps where further stimulus properties, such as orientation in the case of visual stimuli, become expressed in a spatial pattern. Even then, the spatial pattern can be characterized by a smooth variation of stimulus features at most points in the map. Such distribution of stimulus-specificity represents a *dimension-reducing projection* from a higher-dimensional "feature space" of stimulus properties onto the two-dimensional cortical sheet[17),(8),(25]. A mathematical consequence of any such projection would be either to suppress some of the additional feature dimensions, or to exhibit distortions which appear as regions of elevated rate of change of tuning properties in the topographic map.

In the visual cortex both alternatives seem to be realized. Retinotopic location defines the "primary" feature that is mapped smoothly across visual cortex. The spatial variation of additional "secondary" features, such as orientation and ocularity is smooth only within small local domains, which are separated by boundaries where rapid changes occur. In the case of orientation, there are also "foci" where orientation selectivity is zero, i.e. in these locations the stimulus feature "orientation" is suppressed under the projection.

In the following, we will consider a specific model that is able to generate many of the features of these maps on the basis of a simple adaptive process, the so-called "self-organizing feature map" algorithm (Refs. (15), (16), "SFM" in the sequel). We will discuss two different variants of this network model, a low-dimensional "caricature" whose value lies in the fact that it is possible to derive analytical results relating the structure of the map with statistical properties of the stimuli, and a more faithful high-dimensional variant, whose simulation, however, requires the resources of a parallel computer, in this case a Connection Machine CM-2 with 32000 processors. Comparing maps generated with both models, we find that there is a high degree of similarity, allowing the gratifying conclusion that the simpler model is well justified in many situations.

2. The Low-Dimensional Model

The SFM algorithm can be considered as an adaptive procedure for the formation of a topographic

Manuscript received December 19, 1991.

† The authors are with Beckman-Institute and Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

†† The author is with the Department of Computer Science, Bielefeld University, D-4800 Bielefeld, FRG.

representation of a set of patterns (given as vectors in some "input" or "feature"-space V) on a discrete set, A, of points ("cells") endowed with a topology.

Figure 1 shows a schematic drawing of the model. The cells $\vec{r} \in A$ are arranged on a two-dimensional lattice (the network layer) to match the topology of the cortical layer containing the feature map (the cells should not be identified with single neurons, but rather with groups of neurons or with small patches of tissue, where neurons with common response properties are located). Periodic boundary conditions were chosen for the network layer as well as for the position coordinates in feature space.

To describe the receptive field properties position (x_r, y_r) of the receptive field centers in visual space, preferred orientation ϕ_r and orientation specificity q_r for a cell $\vec{r}$ we follow (34) and use a 4-dimensional feature vector

$$\vec{w}_r = (x_r, y_r, q_r \cos(2\phi_r), q_r \sin(2\phi_r)). \quad (1)$$

The dependence of these feature vectors on the cell locations $\vec{r}$ describes the spatial distribution of selectivity of cells over the cortical layer. The position of the receptive field is given by the coordinates (x_r, y_r) of its centroid, preferred orientation by the orientation ϕ_r of the receptive fields major axis and orientation specificity by its elongation.

The input to the network layer consists of localized and oriented stimuli. They are described by a feature vector also, which is of the same type as $\vec{w}_r$, i.e. its components

$$\vec{v} = (x, y, q \cos(2\phi), q \sin(2\phi)) \quad (2)$$

correspond to the stimulus properties position in the visual field (x, y), orientation ϕ of its major principal axis, and elongation q.

A set of stimuli (described by the stimulus vectors $\vec{v} \in V$), drives the model to adapt its feature vectors $\vec{w}_r$ by an iterative sequence of steps: At the beginning of each step a feature vector $\vec{v}$ is chosen at random according to the probability distribution $P(\vec{v})$. Using a distance measure $d(.,.)$ (in our simulations, $d(\vec{v}, \vec{w}) = |\vec{v} - \vec{w}|^2$), the cell $\vec{s}$, whose feature vector $\vec{w}_s$ is closest to $\vec{v}$, is determined:

$$\vec{s} = \min_r d(\vec{v}, \vec{w}_r) \quad (3)$$

and the attached feature vectors are updated according to the SFM updata rule[15],[16]:

$$\vec{w}_r(t+1) = \vec{w}_r(t) + \varepsilon(t) h(\vec{r}, \vec{s}, t)(\vec{v} - \vec{w}_r(t)). \quad (4)$$

The state of the network layer can be defined by the set $\{\vec{w}_r\}$ of all feature vectors $\vec{w}_r$. An essential element of this equation is the presence of the "kernel" $h(\vec{r}, \vec{s}, t)$ that correlates the changes of cells at neighboring positions $\vec{r}$, $\vec{s}$ and is given by:

$$h(\vec{r}, \vec{s}, t) = \exp\left(-\frac{(r_1 - s_1)^2}{\sigma_{h1}^2(t)} - \frac{(r_2 - s_2)^2}{\sigma_{h2}^2(t)}\right) \quad (5)$$

Equations (3), (5) have been shown to lead under a broad variety of conditions to spatial fields of feature vectors $\vec{w}_r$ that can be characterized by the following two conditions: (i) the variation of w_r with cell position $\vec{r}$ is as continuous as possible, and (ii) the resulting vectors $\vec{w}_r$ span the range over which the feature combinations vary in the set of input patterns[17],[25],[29]. These are two complementary requirements: (i) favors "uniformity", while (ii) demands "diversity" for the feature vectors w_r.

The quantities σ_{h1} and σ_{h2} parametrize the shape of the kernel h and, therefore, determine the range over which response properties of cells are kept correlated. The time dependence of $\varepsilon(t)$ and the width $\sigma_{h1}(t)$ and $\sigma_{h2}(t)$ were adjusted to obtain good convergence speed.

Since little is known about the statistical properties of the afferent patterns driving map formation we generated the input patterns from an unbiased, uniform probability distribution $P(\vec{v})$ given by

$$P(\vec{v}) = \begin{cases} \mathcal{N}\delta(q - q_{pat}) & x, y \in [0, d] \\ 0 & \text{else} \end{cases} \quad (6)$$

This corresponds to a uniform distribution on the 3D-surface of a cylinder in the 4-dimensional feature space of the variables x, y, $q\cos(2\phi)$ and $q\sin(2\phi)$. The goal of the map formation process is to represent this 3D-pattern manifold by the (discrete) cells of the two-dimensional network layer, obeying the constraint, that neighboring cells in the network represent neighboring regions on the pattern manifold to the extent possible (Fig. 2). For fixed number $N \times N$ of cells in the network layer and for a fixed length d, it has been shown[27],[28],[26],[24] that the final map crucially

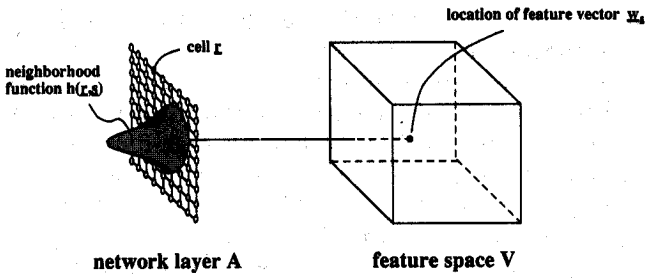

Fig. 1 The "low-dimensional" network model. The model consists of a set of cells which are arranged on a square lattice (network layer). The receptive field properties of each cell are described by a feature vector, which is an element of a four-dimensional feature space. The neighborhood function $h(\vec{r}, \vec{s})$ implements a principle of cooperativity between neighboring cells during map formation (see Eqs. (4), (5)).

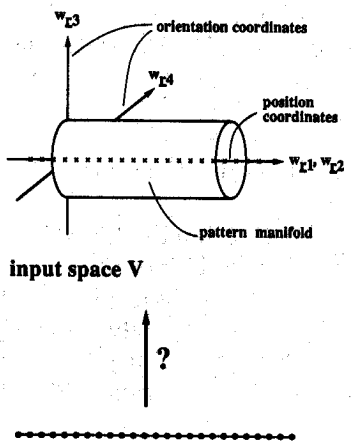

Fig. 2 Dimension reducing mapping of a cylindrical manifold in four-dimensional feature space V onto a two-dimensional network layer A. One "position"-and one network dimension was suppressed in the diagram. The location of the feature vectors $w_{\vec{r}}$ in feature space is indicated by small crosses for the regime $q_{\text{pat}} < q_{\text{thres}}$ (see text).

depends on the variance $\int v_{3,4}^2 P(\vec{v})\, dv_{3,4} = q_{pat}^2$ of the pattern distribution given by Eq. (6) along the feature dimensions describing orientation. If q_{pat} is smaller than a certain threshold q_{thres}, then for the uniform pattern distribution defined above the feature vectors of the network cells would in this case all be located on the x, y-axis in the center of the cylindrical pattern manifold. The resulting map is a topographic representation of visual space, but since $(w_{\vec{r}})_3$, $(w_{\vec{r}})_4$, and therefore $q_{\vec{r}}$ are zero, the feature "orientation" is not represented ("purely retinotopic" map).

In order to form orientation selective cells, i.e. cells $\vec{r}$ for which the third and fourth component of $\vec{w}_{\vec{r}}$, i.e. $(w_{\vec{r}})_3$ and $(w_{\vec{r}})_4$ are different from zero, q_{pat} must exceed a threshold, q_{thres}, above which the "purely retinotopic" map becomes unstable. This threshold can be calculated following the approach of Ref. (27) (details will be published elsewhere) and is given by

$$q_{\text{thres}} = \sqrt{\frac{e}{2}\frac{d}{N}}\sigma_h \qquad (7)$$

where $\sigma_h = \min(\sigma_{h1}, \sigma_{h2})$ and e is the Euler constant. Therefore the purely retinotopic map becomes unstable if the range σ_h of the neighborhood function, projected back to the feature space V, falls below q_{pat}, the standard deviation of the set of patterns along the orientation feature dimension. Let $\vec{u}_{\vec{r}} = \vec{w}_{\vec{r}} - \tilde{\vec{w}}_{\vec{r}}$ denote the deviation of the feature map from its stable state $\tilde{\vec{w}}_{\vec{r}}$. These deviations result from the stochastic nature of the adaptation process (3-5) and their size is proportional to be learning step ε. At the threshold q_{thres} a set of modes

$$\hat{\vec{u}}_{\vec{k}} = \frac{1}{N}\sum_{\vec{r}} e^{i\vec{k}\vec{r}}\, \vec{u}_{\vec{r}} \qquad (8)$$

characterized by

$$\hat{\vec{u}}_{\vec{k}} \perp \vec{k} \qquad (9)$$

and

$$\left.\begin{aligned} &|\vec{k}| = 2/\sigma_h \quad \text{if } \sigma_{h1} = \sigma_{h2} \\ &k_x = \pm 2/\sigma_{h1} \\ &k_y = 0 \end{aligned}\right\} \text{ if } \sigma_{h1} < \sigma_{h2} \qquad (10)$$

becomes unstable and for $q_{\text{pat}} > q_{\text{thres}}$ orientation selective cells form.

Figure 3 shows the final distribution of orientation preference $\phi_{\vec{r}}$ (color) and selectivity $q_{\vec{r}}$ (saturation) along the network layer for a map generated in the regime above threshold with an isotropic neighborhood function ($\sigma_{h1} = \sigma_{h2}$). The presence of only very small black areas indicated that almost all cells have become orientation specific. The cells form domains of continuously changing orientation, in which iso-orientation regions are organized as parallel slabs. The slabs start and end at vortices containing orientationally unspecific cells (dark spots). Orientation preference changes by 180° in a clockwise- or counterclockwise fashion around these foci. Neighboring domains have similar slab-orientations but, on a larger length scale, the directions of the domains are distributed isotropically. The multiple representations of a complete cycle of preferred orientation indicate, that orientation plays the role of a "secondary" feature.

Figure 4 illustrates the topographic representation of visual space generated by the SFM-algorithm. The diagram presents the locations $(x_{\vec{r}}, y_{\vec{r}})$ of receptive field centers in visual space for all cells in the network layer. Receptive field centers of neighboring cells were connected by lines. An ideal topographic projection of visual space to the network layer would give rise to a square lattice with equal mesh size in Fig. 4, since the receptive field centroids of neighboring units are equally spaced in this case. The overall preservation of the lattice topology in Fig. 4 and the absence of any major distortions demonstrate, that indeed "position" varies in a topographic fashion across the cell layer on a large length scale. On a lengh scale below the diameter of a hypercolumn, however, numerous distortions are visible. These distortions are the result of the constraint to map a more than two-dimensional feature space on a two-dimensional cortical surface, such that the response properties of the cells vary smoothly over the cortical surface. Since visubl space is mapped only once along the network layer, "position" plays the role of the primary stimulus variable.

Figure 5 shows the two-point autocorrelation functions

$$S_{ij}(\vec{s}) = \langle (w_{\vec{r}-\vec{s}})_i (w_{\vec{s}})_j \rangle \quad (i, j \in 3, 4) \qquad (11)$$

in the network layer. To good approximation, the autocorrelation function depends only on the absolute value $|\vec{s}|$ if the kernel h is isotropic. Figure 5 presents the correlation functions S_{33}, S_{44}, and S_{34} as a function of the absolute value $|\vec{s}|$ of cell distance averaged over all directions. The correlation function has a Mexican hat form: neighboring cells prefer correlated orientations, with the degree of correlation decreasing with distance. At a separation of $\vec{s}=\lambda/2$ (half the radius of a "hypercolumn") the preferred orientation between cells is anti-correlated, i.e. preferred orientations are more likely to be orthogonal. If cells are separated by a distance larger than λ the preferred orientations are uncorrelated, as one would expect from an isotropic arrangement of hypercolumns over the cortical surface. The crosscorrelation function $S_{34}(\vec{s})$ was found to be very small.

What happens, if the probability distribution $P(\vec{v})$ is biased towards a certain subset of patterns? Such a situation is artificially created in "deprivation-experiments", where an animal is exposed predominantly to patterns of a single orientation.

Figure 6 shows the distribution of orientation preference and selectivity for a map generated by a biased probability distribution $P(\vec{v})$, where 70% of all patterns had orientations ϕ from the range $[0, \pi/4]$ and the remaining 30% were uniformly distributed over the remaining orientations. As can be seen, regions of cells selective to the orientation of the predominant patterns form a network, which contains "blobs" of cells with other orientation preferences. The map still has a hierarchical structure, but the number of foci in the map is found to be smaller than in the "non-deprived" case.

If stimulus patterns are restricted to two orthogonal orientations only (i.e. only values $\phi=0$ and $\phi=\pi/2$ occur) the final distribution of orientation preference (Fig. 7) shows alternating "stripes" or "blobs", which are separated by cells unspecific for orientation (dark bands). The width of the stripes depends on the relative probabilities of occurence: the higher the probability for a pattern, the larger is the corresponding region on the map. The resulting maps very much resemble the spatial structure of an ocular dominance column system. This is not surprising, because restricting the stimulus distribution in feature space to two orthogonal orientations is equivalent to reducing the pattern manifold shown in Fig. 2 to the intersection of the (hyper-) cylinder with any (hyper-) plane, which includes the $w_{\vec{r}1,2}$-(hyper-) axis hence reducing the dimensionality of the feature space from four to three †. The new pattern manifold consists of two lines †† on opposite sides of the $w_{\vec{r}1,2}$-(hyper-) axis. If the feature, which is now coded by the distance between these planes, is re-interpreted as the degree of correlation in the activity between both eyes[11] (large values stand for small correlation in activity), and the corresponding receptive field property as the amount a particular neuron is driven by a particular eye, the same model can describe the development of retinotopic maps and ocular dominance stripes[26),(24)].

3. The High-Dimensional Model

From a modeling standpoint, the use of feature vectors in models of development is not fully satisfactory: coding and selection of features has to be "put in by hand", and the choice of the distance measure d entails some degree of arbitrariness. These shortcomings motivate the investigation of a model whose representation of input patterns is closer to the actual activity patterns themselves. To this end, we now use the activity values v_k at certain points k ("input-cells", "receptors") of an "input-layer" directly to form a pattern vector $\vec{v}=(v_1, v_2, \cdots, v_d)$ as input to the network. Of course, $\vec{v}$, and consequently also each $\vec{w}_{\vec{r}}$, is now from a space of a very high dimensionality d that equals the number of receptors contacted in the input layer. In this setting, the components $(w_{\vec{r}})_k$ can be interpreted as weights or connection strengths of afferents connecting input cells k with network cells $\vec{r}$.

The weight values are iteratively refined using a variant of the SFM algorithm described in the preceeding section. For each step, a pattern vector $\vec{v}$ is selected at random from a given ensemble described by $P(\vec{v})$ and the weights are updated according to:

$$(w_{\vec{r}})_k(t+1) = \frac{((w_{\vec{r}})_k(t)+\varepsilon(t)\,h(\vec{r},\ \vec{s},\ t)\cdot v_k)}{\sqrt{\sum_i((w_{\vec{r}})_i(t)+\varepsilon(t)\,h(\vec{r},\ \vec{s},\ t)\cdot v_i)^2}} \tag{12}$$

where the index $\vec{s}$ denotes the network cell whose input: $\vec{\sigma}_s=\vec{w}_{\vec{s}}\cdot\vec{v}$ is a global maximum. The neighborhood function $h(\vec{r},\ \vec{s},\ t)$ is again given by Eq. (5). For a more detailed interpretation of the terms in a biological context we refer the reader to Ref. (25).

Again, we have to face the problem that very little is known about the statistical properties of the activity profile that drive the formation of feature maps in the cortical layer. Therefore, we constructed sets of patterns on the basis of three general assumptions, namely, (i) that pattern activity should be locally correlated, (ii) that only features to be represented in the map are encoded in the presented patterns, and (iii) that each feature combination is generated with equal probability (under "undeprived" conditions). A "minimal" choice is a set of patterns with elliptic shape, defined

† The axis to the right in Fig. 2 represents two coordinates.

†† Actually two planes, since the $w_{\vec{r}1,2}$-axis represents a plane in the 4-dimensional feature space.

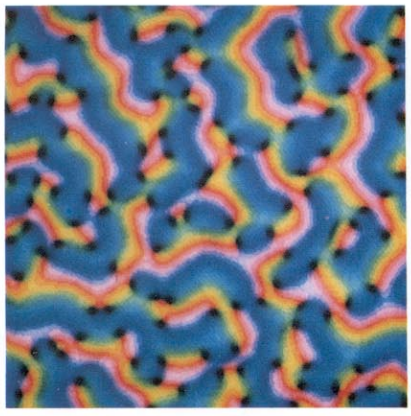

Fig. 3 Spatial organization of the features orientation preference and selectivity for a map formed with the "low-dimensional" model above threshold after 10^7 iterations using an isotropic neighborhood kernel h. The network layer contains 65,536 cells(arranged on a 256×256 square lattice with periodic boundary conditions). The parameters of the simulation were: $q_{\text{pat}}=12$, $\sigma_h=5$ and $\varepsilon=0.01$. Each image pixel corresponds to one cell in the network layer. Orientation preference is indicated by color(light blue → green → orange → purple → blue correspond to angles of 0°→ 45°→ 90°→ 135°→ 180° relative to the vertical axis). The degree $q_{\vec{r}}$ of orientation specificity is normalized to one and indicated by brightness(black: zero, bright: one).

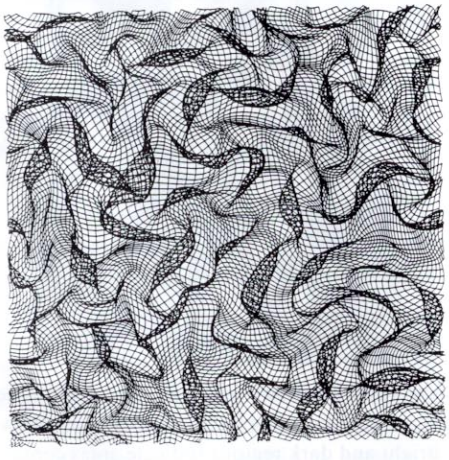

Fig. 4 "Position map" belonging to Fig. 3. The feature receptive field position"$(x_{\vec{r}}, y_{\vec{r}})$ in visual space is indicated by a dot for every second cell in the map of Fig. 3. Points$(x_{\vec{r}}, y_{\vec{r}})$, $(x_{\vec{r}}, y_{\vec{r}'})$ belonging to cells $\vec{r}$, $\vec{r}'$ that are nearest neighbors in the network layer V are connected by lines. x increases to the right, y to the top.

by the three random numbers x, y (stimulus center) and $\alpha\in[0, 180°]$(stimulus orientation)chosen from a uniform distribution. Acoordingly, cell activities v_k are then defined by:

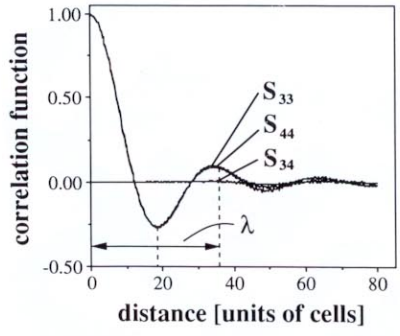

Fig. 5 Two-point correlation functions S_{33}, S_{44}, and S_{34} as a function of the absolute value $|\vec{s}|$ of cell distance for the orientation map shown in Fig. 3. Since the correlation functions are rotationally symmetric they were averaged over all directions.

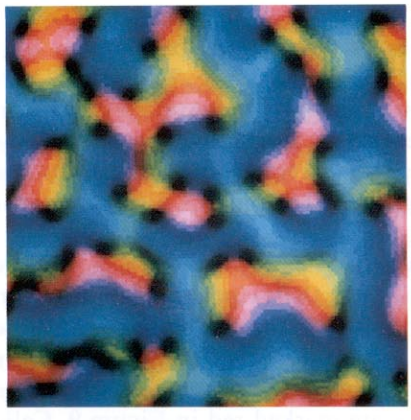

Fig. 6 "Deprivation map" for a biased pattern distribution(70% bias to one orientation). Final feature map for the "low-dimensional" model above threshold after 10^7 iterations using an isotropic neighborhood kernel. The network layer contains 16,384 cells(display of data as described in Fig. 3).

$$v_k=\exp\Big[-\frac{1}{\sigma_1^2}((x_k-x)\cos\alpha-(y_k-y)\sin\alpha)^2 -\frac{1}{\sigma_2^2}((x_k-x)\sin\alpha+(y_k-y)\cos\alpha)^2\Big] \quad (13)$$

where (x_k, y_k) denotes the spatial location of cell k in the input layer. Parameters σ_1 and σ_2 are fixed and specify the length of the major and the minor axis of the ellipsoidal intensity distribution.

Figure 8 shows the final distribution of orientation preference(color) and selectivity(saturation) along the network layer for cels initially unspecific both to stimulus orientation and position. Most of the cells have developed elongated receptive fields, whose aspect ratio approximately matches the aspect ratio of the stimuli presented to the receptor surface. These cells became selective for stimuli, whose orientation matches the orientation of their receptive fields. The

Fig. 7 "Deprivation map" for a pattern distribution restricted to two equiprobable orthogonal orientations only. Final feature map for the same network as used in Fig. 6.

high degree of orientation specificity is indicated by the bright colors in Fig. 8. The spatial structure of orientation preference is more irregular compared to maps generated by the "low-dimensional" model (Fig. 3). While in some regions the slab-like structure dominates, there are other regions displaying more patchy domains.

A gradient filter applied to the orientation values reveals (Fig. 9), that vortices with high orientation changes are often connected by bands of rapid orientation shifts, where the orientation gradient can become very large ("fractures"). These areas with rapid orientation change are marked red in Figure 9. Cells that did not become tuned to stimuli of a particular orientation kept almost circular receptive fields and consequently, respond equally well to stimuli of all orientations. They are indicated by dark color in Fig. 9. The correlation of red and dark regions in Fig. 9 thus shows that cells still unspecific for orientation are primarily found close to "fractures" where the orientation gradient is high. Interestingly, this arrangement can be understood as a direct consequence of the property of the SFM-algorithm to generate maps, where receptive field properties change smoothly along the network layer: For almost circular receptive fields, small changes in the shape of the receptive fields are sufficient to greatly change the direction of their major principal axis. Since even for such receptive fields, in the brain intracortical mechanisms are likely to greatly enhance the otherwise low orientation specificity of the associated cells[31], regions containing this type of cells may develop into the "fractures" between orientation domains. Therefore, the SFM-model demonstrates that "fractures" may result from rather smooth variations in the underlying neural circuitry.

The final size and the shape of the receptive fields depends on the size of the presented patterns as well as on the final range $\sigma_h(t_{\text{final}})$ of the neighborhood function $h(\vec{r}, \vec{s}, t_{\text{final}})$, where t_{final} is the maximum number of iterations. In the case of non-oriented patterns ($\sigma_1 = \sigma_2 = \sigma$) the size of the cell's "optimal stimulus" can be calculated analytically[23],[28]. The result shows, that the cells develop receptive fields whose area is the sum of two terms: the first term is essentially the area of a typical stimulus and the second term is essentially the area $\propto \sigma_h^2$ in the adjustment zone in the network layer, but "projected back" by inverse magnification

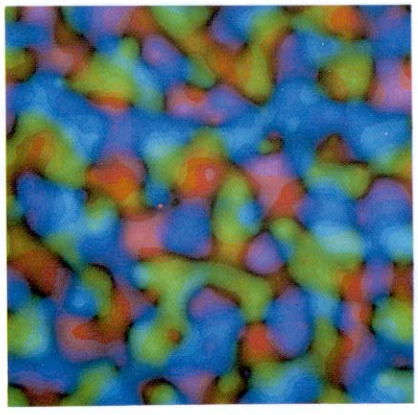

Fig. 8 Orientation map similar to Fig. 3, but obtained with the "high-dimensional" model (above threshold, 30,000 adaptive steps). The network layer contains 65,536 cells (arranged on a 256×256 square lattice with periodic boundary conditions), the input layer 900 randomly distributed "receptor cells". The initial connection strengths were chosen randomly and normalized to unity. Orientation preference of each cell is indicated by color in the same way as described in Fig. 3. The degree of orientation specificity of each cell was measured by the ratio of the variances of the receptive field along its major and minor principal axis. Its value is indicated by brightness (dark : unspecific, bright : specific).

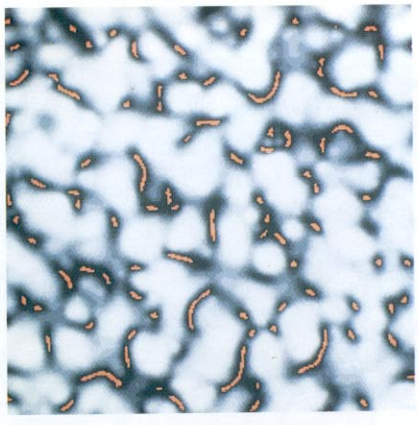

Fig. 9 Correlation between "fractures" and regions containing cells unspecific to orientation for the map shown in Fig. 8. Bright and dark regions indicate areas containing cells specific and unspecific for orientation, respectively. The regions of steepest orientation gradient are marked red. Note the excellent correlation between "fractures" and reigons containing cells unspecific for orientation.

factor M^{-1} onto the input sheet. The size of the cells' receptive fields, therefore, reflect the size of the region within which the response properties of the cells are kept correlated.

4. Discussion

The model presented above describes the formation of cortical feature maps. Biologically plausible developmental principles (activity-dependent synaptic plasticity, correlated response properties of neighboring cells, and competition between afferents) are formulated in a mathematically simple form to describe an activity-based process. This process generates maps from a given stimulus distribution, which is based on similarity between stimuli and on the probability of their occurence. The observed model dynamics does not necessarily mirror the actual sequence of events in the formation of cortical maps. The emphasis of this model (and of similar models[8]) is on the role of basic principles and on an analysis of their capacity to yield the observed organization of cortical maps.

Typical features of observed maps in the primary visual cortex of various species include (Refs. (5), (12), (14), (19), (35), (37)): hierarchical mapping of position (primary stimulus variable) and orientation (secondary variable), patch- or slab-like structure of iso-orientation domains, vortices, "fractures", topographic order and a negative correlation between the magnitude of the orientation gradient and the orientation specificity (as found e.g. in the correlation between the cytochrome oxidase blobs, which contain orientation unspecific cells and the vortices of the orientation map in the area 17 of the macaque[6],[7]). Maps exhibiting these features can be generated by the SFM-algorithm for a broad range of model parameters and independently of the initial state. The model predicts some constraints regarding the set of stimulus patterns. If the formation of the visual map employs externally or internally generated activity patterns driving a process governed by Eqs. (3)-(5) or (12), and if the configuration space of the network is not confined by some prestructuring, then the eccentricity of the patterns would have to be within a certain range, large enough to exceed the threshold required to express orientation in the map, but still sufficiently small not to destroy the observed feature hierarchy.

Since the probability distribution $P(\vec{v})$ is an important determinant of the structure of the resulting maps, it is important to compare the results of the model for probability distributions that correspond to "deprivation situations" that can be induced experimentally. In such experiments the statistical properties of the afferent activity pattern is changed by artificially changing sensory input (e.g. by rearing in darkness) or by chemically changing the firing pattern of cells in the subcortical structures. In the SFM-model, the effect of sensory deprivation was imitated by an appropriate inhomogeneous probability distribution $P(\vec{v})$ of patterns. For instance, restricting the patterns to a single orientation led to the generation of maps where deprived columns are no longer present, but if "deprived" patterns are generated with some probability, then a column system emerges, where regions corresponding to the experienced orientation are larger than the regions corresponding to "deprived" orientation. Both effects are in good qualitative agreement with experimental observations[3],[4],[32].

It was mentioned in Sect. 2 that in case $P(\vec{v})$ is restricted to two orientations, the SFM model can be re-interpreted to apply to the formation of ocular dominance columns. The stripe-width depends on the correlation structure of the patterns and increases monotonically with decreasing correlation. With proper interpretation the results shown in Fig. 6 are in qualitative agreement with experimental results from suture[18] and impulse blockade[33] experiments. In the case of "artificial strabismus"-conditions[10] (zero activity correlation between both eyes) the SFM-model wrongly predicts a reversal of feature hierarchy[26], ocular dominance now being the primary feature instead of "position". In order to explain the experimental findings one has to assume a two-stage process, where the first stage leads to a topographic prestructuring for the second stage, the input driven process. If a rough topographic prewiring is assumed, which constrains the configuration space of the network, a stripe-like pattern emerges again.

The constraint of dimension reduction leads to a trade-off between representation of orientation and the degree of retinotopy. One consequence of this trade-off is visible in the position map: the simultaneous representation of orientation above threshold causes multiple distortions in the position map. These distortions are separated by $\lambda/2$ (half the radius of a hypercolumn, see Fig. 6(b)) from each other on the average and are linked to the spatial pattern of orientation patches. At the orientation discontinuities between patches the rate of change of retinotopic location with cortical distance is, on the average, reduced compared to its change with cortical distance within orientation patches. This corresponds to a certain anticorrelation in the resolution of the cortical representation of the features retinotopic location and orientation. The resulting "modulation" of the position map should be experimentally detectable in a high resolution study of receptive field centroids.

Recently, we have complemented these more qualitative studies by a more quantitative comparison, based on evaluation and comparison of spectra and correlation functions for the SFM-maps and for experimental data. In this paper, we have reported some preliminary results in this respect. The two-dimensional Fourier spectrum of the orientation val-

ues for the isotropic neighborhood function exhibit discrete modes located on a ring around the origin in the $\vec{k}$-plane, for the anisotropic neighborhood function (Fig. 8) it consists of two localized groups of modes, which differ in the sign of $\vec{k}$. These findings are in agreement with data from monkey[24] and cat[35]. The spatial autocorrelation functions are short-ranged and have mexican hat shape (Fig. 9) similar to the orientation-autocorrelation function found in experimental data[35],[24]. Note, that there is almost no correlation between the two orientation coordinates $(w_r)_3$ and $(w_r)_4$. The finite range of the spatial autocorrelation functions reflects the finite length of the neighborhood function at the threshold q_{pat} and is in accordance with the observed random distribution of hypercolumns[5],[24]. Thus, experimental maps of orientation preference and simulation results are characterized by local correlation and global disorder and a phenomenological description of these patterns in terms of a filter process acting on spatial random noise[30] capture essential features of the system.

Considering the complexity of the brain, our models are still crude and we must suspect that we are still far from a deeper understanding of many important principles. However, research on models for the formation of topographic maps, to which the present paper hopes to make a modest contribution, has now matured to the point where we can begin to make detailed comparisons with experimental data. In view of the widespread occurrence of topographic structures in the brain, this may provide us with valuable insights which then, in the long term, may turn out fruitful also for the technological realization of artificial neurocomputing devices.

Acknowledgement

The authors would like to thank Dr. G. Blasdel for stimulating discussions. The help and support of R. Brady and R. Kufrin in all technical matters concerning the use of the Connection Machine system, and the financial support of one of the authors (K. Obermayer) by a scholarship of the Boehringer-Ingelheim Fonds is greatly acknowledged. This research has been supported by the National Science Foundation (grant number 9017051). Computer time on the Connection Machine CM-2 has been made available by the National Centers for Supercomputer Applications at Urbana-Champaign and Pittsburgh supported by the National Science Foundation.

References

(1) Amari S.: "Dynamical stability of formation of cortical maps", Dynamic Interactions in Neural Networks: Models and Data, M. A. Arbib, S. Amari eds., Springer Research Notes in Neural Computation, **1**, pp. 15-34 (1988).

(2) Amari S.: "Topographic Organisation of Nerve Fields", Bull. of Math. Biology, **42**, pp. 339-364 (1980).

(3) Blakemore C. and Van Sluyters R. C.: "Innate and Environmental Factors in the Development of the Kitten's Visual Cortex", J. Physiol., London, **248**, pp. 663-716 (1975).

(4) Blasdel G. G., Mitchell D. E., Muir D. W. and Pettigrew J. D.: "A Physiological and Behavioural Study in Cats of the Effects of Early Visual Experience with Contours of a Single Orientation", J. Physiol., London, **265**, pp. 615-636 (1977).

(5) Blasdel G. G. and Salama G.: "Voltage Sensitive Dyes Reveal a Modular Organization in Monkey Striate Cortex", Nature, **321**, pp. 579-585 (1986).

(6) Blasdel G. G.: "Differential Imaging of Ocular Dominance and Orientation Selectivity in Monkey Striate Cortex", J. Neurosci., (1992), in press.

(7) Blasdel G. G.: "Orientation Selectivity, Preference, and Continuity in Monkey Striate Cortex", J. Neurosci (1992), in press.

(8) Durbin R. and Mitchison M.: "A Dimension Reduction Framework for Understanding Cortical Maps", Nature, **343**, pp. 644-647 (1990).

(9) Erwin E., Obermayer K. and Schulten K.: "Convergence Properties of Self-organizing Maps", Proceedings of the International Conference on Artificial Neural Networks, Helsinki (1991), in press.

(10) Freeman R. D., Sclar G. and Ohzawa I.: "Cortical Binocularity is Disrupted by Strabismus more Slowly than by Monocular Deprivation", Dev. Brain Res., **3**, pp. 311-316 (1982).

(11) Goodhill G.J. and Willshaw D. J.: "Application of the Elastic Net Algorithm to the Formation of Ocular Dominance Stripes", Network, **1**, pp. 41-59 (1990).

(12) Grinvald A., Lieke E., Frostig R. D., Gilbert C. and Wiesel T. M.: "Functional Architecture of Cortex Revealed by Optical Imaging Technique", Nature, **324**, pp. 361-364 (1986).

(13) Hirsch H. V. B.: "The Role of Visual Experience in the Development of Cat Striate Cortex", Cell. Mol. Neurobiol, **5**, pp. 103-121 (1985).

(14) Hubel D. H. and Wiesel T. N.: "Sequence Regularity and Geometry of Orientation Columns in the Monkey Striate Cortex", J. Comp. Neurol., **158**, pp. 267-294 (1974).

(15) Kohonen T.: "Self-organized Formatin of Topologically Correct Feature Maps", Biol. Cybern., **43**, pp. 59-69 (1982).

(16) Kohonen T.: "Analysis of a Simple Self-organizing Process", Biol. Cybern., **44**, pp. 135-140 (1982b).

(17) Kohonen T.: "Self-Organization and Associative Memory", Springer-Verlag, New York (1983).

(18) LeVay S., Wiesel T. N. and Hubel D. H.: "The Development of Ocular Dominance Columns in Normal and Visual Deprived Monkeys", J. Comp. Neurol., **191**, pp. 1-51 (1980).

(19) Löwel S., Freeman B. and Singer W.: "Topographic Organization of the Orientation Column System in Large Flat-mounts of the Cat Visual Cortex: a 2-deoxyglucose Study", J. Comp. Neurol., **255**, pp. 401-415 (1987).

(20) von der Malsburg C.: "Self-organization of Orientation Sensitive Cells in the Striata Cortex", Kybernetik, **14**, pp. 85-100 (1973).

(21) Merzenich M. M., Kaas J. H., Wall J. T., Sur M., Nelson R. J. and Felleman D. J.: "Progession of Change Following Median Nerve Section in the Cortical Representation of the Hand in Areas 3b and 1 in Adult Owl and Squirrel Monkeys", Neurosci., **10**, pp. 639-665 (1983).

(22) Nelson M. E. and Bower J. M.: "Brain Maps and Parallel Computers", TINS, **13**, pp. 401-406(1990).

(23) Obermayer K., Ritter H. and Schulten K.: "A Neural Network Model for the Formation of Topographic Maps in the CNS: Development of Receptive Fields", Proceedings of the IJCNN 1990, Vol. II, IEEE Computer Society Press, pp. 423-429.

(24) Obermayer K., Blasdel G. G. and Schulten K.: "A Neural Network Model for the Formation of the Spatial Structure of Retinotopic Maps, Orientation- and Ocular Dominance Columns", Artificial Neural Networks I, Eds. Kohonen T. et al., North Holland, pp. 505-511(1991).

(25) Obermayer K., Ritter H. and Schulten K.: "A Principle for the Formation of the Spatial Structure of Cortical Feature Maps", Proc. Natl. Acad. Sci. USA, **87**, pp. 8345-8349(1990).

(26) Obermayer K., Ritter H. and Schulten K.: "Development and Spatial Structure of Cortical Feature Maps: A Model Study", Advances in Neural Information Processing Systems 3, Eds. Touretzky D. S., Lippman R., Morgan Kaufmann Publishers, pp. 11-17(1991).

(27) Ritter H. and Schulten K.: "Convergence Properties of Kohonen's Topology Conserving Maps: Fluctuations, Stability and Dimension Selection", Biol. Cybern., **60**, pp. 59-71 (1989).

(28) Ritter H., Obermayer K., Schulten K. and Rubner J.: "Self organizing Maps and Adaptive Filters", Physics of Neural Networks, Eds. Domani E., van Hemmen J. L., Schulten K., Springer-Verlag, New York, pp. 281-306 (1991).

(29) Ritter H., Martinetz T. and Schulten K.: "Neural Computation and Self-Organizing Maps", Addison-Wesley, Reading Mass.(1991).

(30) Rojer A. S. and Schwarz E. L.: "Cat and Monkey Cortical Columnar Patterns Modeled by Bbandpass-filtered 2d White Noise", Biol. Cybern., **62**, pp. 381-391 (1990).

(31) Silito A. M.: "Cerebral Cortex", Eds. Peters A. and Jones E. G., Plenum Press, New York, **2**, pp. 91-117(1984).

(32) Singer W., Freeman B. and Rauschecker J.: "Restriction of Visual Experience to a Single Orientation Affects the Organization of Orientation Columns in Cat Visual Cortex", Exp. Brain Res., **41**, pp. 119-215(1981).

(33) Stryker M. P. and Harris W.: "Binocular Impulse Blockade Prevents the Formation of Ocular Dominance Columns in Cat Visual Cortex", J. Neurosci., **6**, pp. 2117-2133(1986).

(34) Swindale N. V.: "A Model for the Formation of Orientation Columns", Proc. R. Soc. Lond., **B215**, pp. 211-230 (1982).

(35) Swindale N. V., Matsubara J. A. and Cynader M. S. : "Surface Organization of Orientation and Direction Selectivity in Cat Area 18", J. Neurosci., **7**, pp. 1414-1427 (1987).

(36) Takeuchi A. and Amari S.: "Topographic Organization of Nerve Fields", Biol. Cybern., **35**, pp. 63-72(1979).

(37) Ts'o D. Y., Frostig R. D., Lieke E. and Grinvald A.: "Functional Organization of Primate Visual Cortex Revealed by High Resolution Optical Imaging", Science, **249**, pp. 417-420(1990).

(38) Willshaw D. J. and von der Malsburg C.: "How Patterned Neural Connections can be set up by Self-organization", Proc. R. Soc. Lond. B, **194**, pp. 431-445 (1976).

Klaus Obermayer was born in 1961. He studied physics and biology at the University of Stuttgart, Federal Republic of Germany and received his Diplom degree in physics from the University of Stuttgart in 1987. He is a fellow of the Boehringer-Ingelheim Foundation and is currently with the Department of Physics and with the Beckman Institute at the University of Illinois at Urbana-Champaign, USA. Since 1988 he has been engaged in research in the field of neural networks where he has worked on the mathematical analysis of adaptive neural network architectures, neural network implementations on massively parallel computers and on pattern formation and self-organization during the development of the brain. He tries to combine a "physical" approach to self-organization, where processes are studied in simple idealized systems, with a thorough analysis and a detailed comparison with experimental data. He has published articles on molecular switching elements, quantum computers, and neural networks.

Helge Ritter was born in 1958. He studied Physics and Mathematics at the Universities of Bayreuth, Heidelberg and Munich, Fed. Rep. Germany, and received a Ph.D. degree in Physics at the Technical University of Munich in 1988. Since 1985 he has been engaged in research in the field of neural networks. His main interests are the study and application of biologically motivated strategies of massively parallel computation, in particular the analysis of neural network architectures for adaptive signal processing and for the coordination of robot movements. In 1989 he spent two months at the Laboratory of Computer and Information Science at Helsinki University of Technology. Subsequently he moved to the Beckman Institute for Advanced Science and Technology at the University of Illinois at Urbana-Champaign. Since 1990 he is assistant professor at the Department of Information Science, Bielefeld University, where he is carrying out research on neural network based robot vision and control strategies.

Klaus J. Schulten was born in 1947. He received his Diplom degree in physics from the University of Münster, Federal Republic of Germany, in 1969, and the Ph.D. degree in chemical physics from Harvard University in 1974. In 1974 he joined the Max-Planck-Institute for Biophysical Chemistry in Göttingen and in 1980 became Professor of Theoretical Physics at the Technical University of Munich. In 1988 he moved to the University of Illinois at Urbana-Champaign, USA, where he is Professor of Physics, Chemistry, and Biophysics, and a member of the new Beckman Institute. His research areas are many-particle physics, statistical mechanics, molecular biophysics, and computational neural science. In the latter area he studies currently the problems of motion synthesis in robots, data compression through adaptive filters, and pattern formation during the ontogenetic development of the central nervous system.