Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 8345–8349, November 1990
Neurobiology

A principle for the formation of the spatial structure of cortical feature maps

(adaptive map/neural network/orientation columns/self-organization/striate cortex)

K. OBERMAYER, H. RITTER, AND K. SCHULTEN

Beckman Institute and Department of Physics, University of Illinois at Urbana–Champaign, Urbana, IL 61801

Communicated by Hans Frauenfelder, May 31, 1990 (received for review March 2, 1990)

ABSTRACT **Orientation-selective cells in the striate cortex of higher animals are organized as a hierarchical topographic map of two stimulus features: (*i*) position in visual space and (*ii*) orientation. We show that the observed structure of the topographic map can arise from a principle of continuous mapping. For the realization of this principle we use a mathematical model that can be interpreted as an adaptive process changing a set of synaptic weights, or synaptic connection strengths, between two layers of cells. The patterns of orientation preference and selectivity generated by the model are similar to the patterns seen in the visual cortex of macaque monkey and cat and correspond to a neural projection that maps a more than two-dimensional feature space onto a two-dimensional cortical surface under the constraint that shape and position of the receptive fields of the neurons vary smoothly over the cortical surface.**

The striate cortex of higher animals contains a topographic representation of visual space in which are embedded neighborhood-preserving maps of several variables describing features such as position in visual space, line orientation, movement direction, and ocularity (1–4). The representation of the multidimensional feature space onto the two-dimensional cortical sheet is achieved in a hierarchical fashion (5). The topographic projection of the retina establishes a primary order, and for each small region of the visual field there are patches or stripes of cells with similar feature preference (1, 3). In this report we restrict our discussion to the features position and orientation.

Distribution of orientation-selective cells within the visual cortex of the cat, macaque monkey, and the tree shrew has been characterized by several groups (1, 3, 5–7). Various mechanisms have been proposed to explain the input-driven formation and the plasticity of cortical maps, and simulations have shown that some of these mechanisms can account for various aspects of the observed organization (8–14). In this article we will investigate a neural-network model for the formation of orientation columns. The model is based on the self-organizing feature map algorithm (15, 16), which incorporates several of the proposed mechanisms (9). This investigation differs from the previously published work in several respects—the most important being that (*i*) the combined formation of the retinotopic projection and the orientation column system are studied, (*ii*) the interactions between both maps are presented, and (*iii*) the variations in the spatial distribution of feature-selective cells found in different species are considered.

The Principle of Continuous Mapping. Are there general principles that can explain the formation of these highly ordered patterns? To investigate this question we consider a mathematical model (based on refs. 15 and 16) that determines cell properties along the cortical surface such that the "optimal stimuli" of the cell vary as smoothly as possible. This constraint can be interpreted as establishing a mapping from a higher-dimensional feature space characterizing the stimuli onto the two-dimensional cortical surface such that continuity is maximized. The required computations can be interpreted as adaptive processes. Such processes have been suggested to play an important role in the formation of retinotopic projections (9, 11) and orientation columns (10). Note that in the model under consideration the principle of continuous mapping does not operate on the level of stimulus features (see, for example, ref. 9) but on the level of optimal stimuli, leading to the phenomenon of "fractures" in the feature map, to be discussed below.

The Mathematical Model. The model is related to earlier modeling efforts focusing on a single stimulus feature (10, 11) and is based on the algorithm of self-organizing feature maps (9, 15, 16). The model involves a two-layer system of cells (Fig. 1). The first layer serves as an input layer in which the randomly distributed cells encode some spatial intensity pattern (referred to as stimulus). The activity of a cell i in this layer is denoted by r_i, $i = 1, 2, \ldots, N_I$. The stimulus features, describing this activity pattern, are mapped onto the cells of the (two-dimensional) second layer. For each such second-layer cell, the model defines a best-stimulus-position s_{kl} and a best-stimulus-orientation β_{kl} (where k, l represent the two coordinates locating the cell in the network layer), by using a set of weights w_{kli}, $i = 1, 2 \ldots N_o$ associated with the cell (k, l). The definition of s_{kl} and β_{kl} is motivated by an interpretation of these weights as synaptic connection strengths from input cells i to cells (k, l) in the network layer. According to such interpretation the weighted sum $o_{kl} = \Sigma_i w_{kli} r_i$ measures the external input converging at a cell (k, l). For all activity patterns with fixed sum $\Sigma_i r_i^2$, the pattern $r_i = C \cdot w_{kli}$, where C = an arbitrary constant, leads to a maximum of o_{kl} and will be referred to as the optimal stimulus of that cell (k, l). The centroid of the optimal stimulus and the orientation of the major principal axis of the optimal stimulus define s_{kl} and the angle β_{kl}, respectively.

The postulated principle of continuous mapping is not enforced for the values s_{kl} and β_{kl} directly but is enforced for the weights w_{kli}. A set of stimulus patterns drives the model to adapt its weight values such that (*a*) the variation of w_{kli} as a function of cell position (k, l) in the second layer is as continuous as possible (index i is considered as fixed); (*b*) the range of the resulting best positions and best orientations for all cells matches the range over which these feature combinations vary in the set of stimulus patterns. These are two complementary requirements: condition *a* favors uniformity, whereas condition *b* demands diversity for the weight values w_{kli} (9).

To satisfy conditions *a* and *b* we use a variant of the self-organizing feature map algorithm (9, 15, 16). According to this algorithm, a set of initial weight values is iteratively refined by a sequence of discrete steps. For each step, a stimulus pattern is selected at random from a given ensemble, and the weights are updated according to the equation:

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

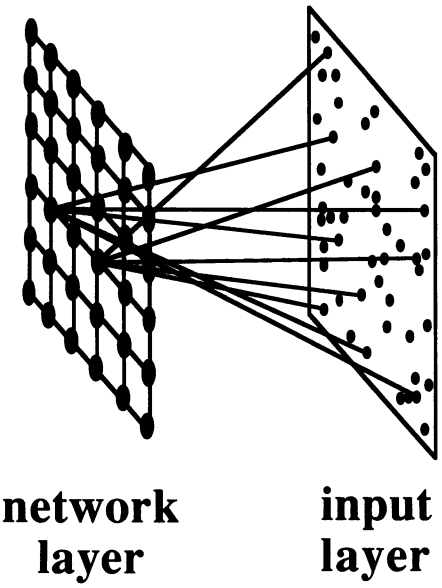


FIG. 1. Two-layered model network. The activity pattern from cells in the input layer gives rise to a map of the stimulus features in the network layer as is described in text.

$$w_{kli}(t+1) = C_{kl}(t)(w_{kli}(t) + \varepsilon(t)h_{kl;rs}r_i). \quad [1]$$

Prefactors $C_{kl}(t)$ and $\varepsilon(t)$ serve to keep the sum $\Sigma_i w_{kli}^2$ constant and to scale the size of the weight change, respectively. Indices (r, s) denote the cell where input o_{rs} is a global maximum. The function $h_{rs;kl}$ has a maximum at $(r, s) = (k, l)$ and decays rapidly to zero for distances above σ, where σ denotes the length scale below which one assumes that intracortical mechanisms keep the response properties of the cells correlated (9, 10). The continuity requirement condition *a* is realized by the accumulation of local smoothing operations on the spatial variation of the weights w_{kli} (second term in Eq. 1). Because each smoothing operation is always limited to an area of approximate radius σ, the competition between patterns of change weights in their favor (affected by the normalization prefactor C_{kl}) leads to the diversity required by condition *b*.

This realization of the principle of continuous mapping allows one to interpret the iterative computation procedure, Eq. 1, in terms of a Hebbian rule: r_i is the "presynaptic activity" delivered from input cell i, whereas intracortical mechanisms would be required to adjust the postsynaptic activity of cell (k, l) to a value proportional to $h_{kl;rs}$ (9).

Stimulus Patterns and Model Parameters. A pattern of elliptic shape is defined by three independent random numbers, each chosen from a uniform distribution: A pair of coordinates $\mathbf{x}^{\text{stim}} = (x_1^{\text{stim}}, x_2^{\text{stim}})^{\text{T}}$ in the unit square (stimulus center) and an angle $\alpha \in [0, 180°]$ (stimulus orientation). Cell activities r_i were then defined by

$$r_i = \exp\left[-\frac{1}{\sigma_1^2}((x_{i1} - x_1^{\text{stim}})\cos\alpha - (x_{i2} - x_2^{\text{stim}})\sin\alpha)^2 - \frac{1}{\sigma_2^2}((x_{i1} - x_1^{\text{stim}})\sin\alpha + (x_{i2} - x_2^{\text{stim}})\cos\alpha)^2 \right], \quad [2]$$

where $\mathbf{x}_i = (x_{i1}, x_{i2})^{\text{T}}$ denote the position of cell i 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.

The parameter $\varepsilon(t)$ decreases linearly from an initial value ε_1 to a final value ε_F. Function $h_{rs;kl}$ is assumed as a Gaussian

$$h_{rs;kl}(t) = \exp\left[-\frac{(r-k)^2}{\sigma_{h1}^2(t)} - \frac{(s-l)^2}{\sigma_{h2}^2(t)} \right]. \quad [3]$$

Parameters σ_{h1} and σ_{h2} can be used to specify an anisotropy of the intracortical connectivity. In the present simulation σ_{h1} and σ_{h2} were changed from larger initial values $\sigma_{h1,I}$ and $\sigma_{h2,I}$ to smaller, final values $\sigma_{h1,F}$ and $\sigma_{h2,F}$ (15, 16).

RESULTS AND DISCUSSION

Formation of Orientation Columns. Fig. 2*A* shows the final spatial distribution of orientation preference (color) and selectivity (saturation) for cells initially unspecific both to stimulus orientation and position. The presence of only very small "dark" areas indicates that almost all cells have become orientation specific. These cells form domains of continuously changing orientation, in which iso-orientation regions are organized as parallel slabs. The slabs start and end at foci 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. Fig. 2*B* displays the distribution of cells tuned to a small interval of orientation values and closely resembles the 2-deoxyglucose patterns found after exposing the animal to a global visual pattern of parallel stripes. The simulation results closely resemble the maps found experimentally in the macaque monkey (e.g., refs. 1 and 4). Fig. 3 simulates an electrode penetration tangentially to the cell layer and shows smooth change of orientation preference, smooth change with reversals of orientation shifts, and interruptions by sudden breaks, all very similar to variations found in experiments (5).

To differentiate the regions of smooth and rapid orientation shift more clearly, a gradient filter was applied to the orientation values (Fig. 2*C*). Foci with rapid orientation changes (white "spots") are often connected by bands of rapid orientation shift, where the orientation gradient can become very large ("fractures"). Areas containing unspecific cells coincide with the regions of rapid orientation change. This can be understood if one observes that for almost-circular receptive fields small changes in the field shape are sufficient to greatly change the direction of its major principal axis. Because even for such receptive fields, intracortical mechanisms are likely to greatly enhance the otherwise low orientation specificity of the associated cells (17), fractures between orientation domains may result from much smoother variations in the shape of the receptive fields.

The optimal stimuli of orientation-selective cells in the model generally reflect the shape of the presented stimuli. Cells located within the foci, however, have circular-shaped best stimuli, although circular stimuli have not been presented to the network. Their presence ensures a smooth progression of the shape of the best stimuli also between cells of orthogonal-orientation preference and, therefore, contributes essentially to the realization of the principle of continuous mapping.

The overall preservation of the lattice topology and the absence of any larger discontinuities in the mapping of the stimulus dimensions position (Fig. 2*D*) demonstrates that best position plays the role of the primary-stimulus variable and varies in a topographic fashion across the cell layer. On length scales below the diameter of a hypercolumn, however, numerous local distortions are visible. The regions of rapid position change (large meshes) correspond to regions of smooth shift in receptive field shape and vice versa.

Analytical calculations (18) as well as earlier simulations (9) for a low-dimensional analogue of the presented model show that such periodic distortions appear when the variance of the secondary-stimulus variable reaches a level comparable to the change of the primary-stimulus variable associated

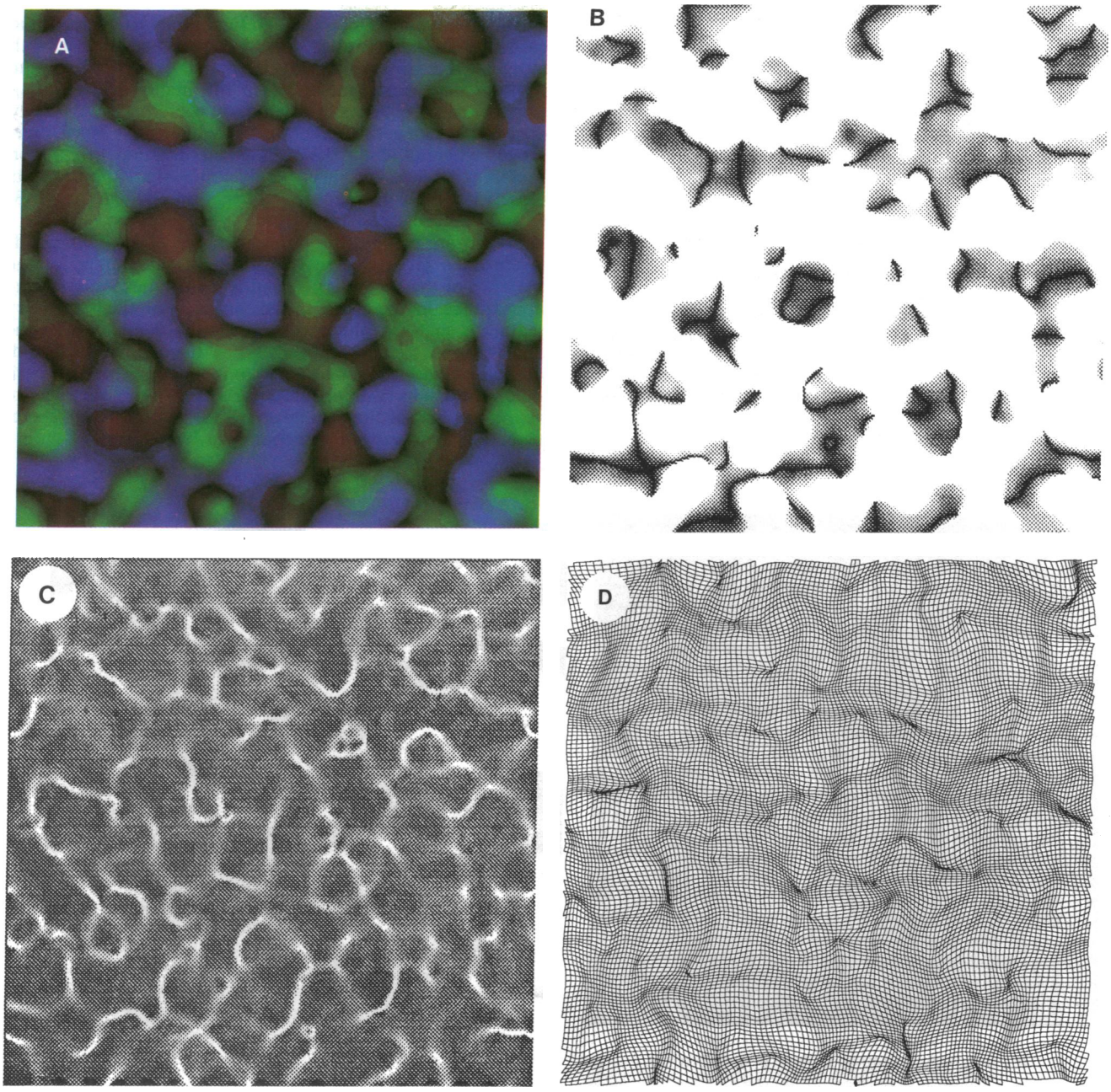


FIG. 2. Spatial organization of the final feature map after 30,000 adaptive steps using an isotropic neighborhood function ($\sigma_{h1} = \sigma_{h2}$). The network layer contains 65,536 cells (arranged on a 256 × 256 square lattice with periodic boundary conditions), and the input layer contains 900 randomly distributed cells. The initial connection strengths were chosen randomly from the interval (0, 1) and normalized to unity. Parameters of the stimulation were as follows: $\sigma_h(0) = 240$; $\sigma_h(15{,}000) = 60$; $\sigma_h(30{,}000) = 2$; $\varepsilon_I = 0.09$; $\varepsilon_F = 0.02$; $\sigma_1 = 0.23$; and $\sigma_2 = 0.09$. (*A*) Spatial pattern of orientation preference and selectivity. 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° with the vertical). The specificity of each cell was measured by the ratio of the variances of the optimal stimulus along its major and minor principal axis. Its value is indicated by brightness (dark, unspecific; bright, specific). (*B*) Spatial distribution of cells in the network layer with deviations of orientation preference near a fixed direction (60°). Deviations from 60° are indicated by gray values (black, 0°; white, ±30° or more). (*C*) Gradient of the orientation values shown in *A*. Bright and dark regions indicate areas of rapid and smooth change in orientation preference, respectively. (*D*) Best-position map. Every second lattice position (k, l) of the network layer is projected to its best position $\mathbf{s}_{kl}$ in the input layer. Points belonging to cells that are nearest neighbors in the network layer are connected by links.

with progressing by a certain distance d^* in the map.† The distance d^* is essentially the range σ of the function $h_{rs;kl}$ (18) and corresponds to the distance over which response properties of cells are kept correlated. Therefore, for fixed resolution of the primary-stimulus variable, the threshold in the variance of the secondary-stimulus variable above which it becomes represented in the map varies linearly with the length scale σ. The resulting periodic distortions in the position map should be experimentally detectable in a high-resolution-mapping study of receptive-field centroids.

Anisotropic Neighborhood Function. Fig. 4 shows the spatial organization of a typical feature map obtained using an

†Usually, there is an ambiguity in comparing magnitudes of change of different stimulus features. Here this ambiguity is absent because from the mathematical analysis it follows that the variances of the two features must be compared in the natural metric of the vector space spanned by the activity patterns r_i driving the Eq. 1 process.

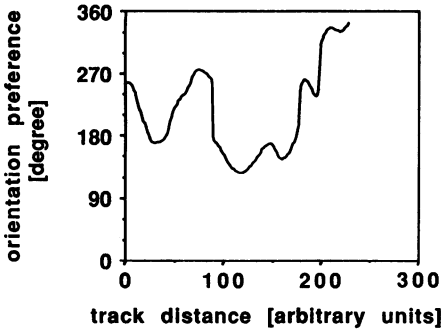
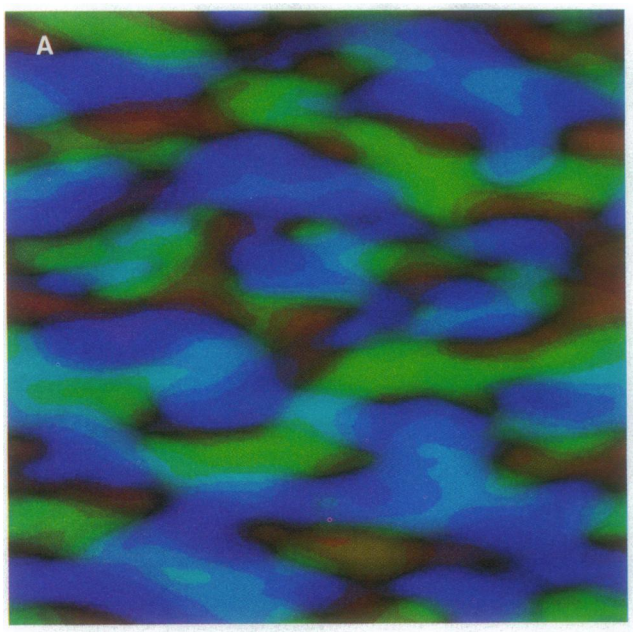
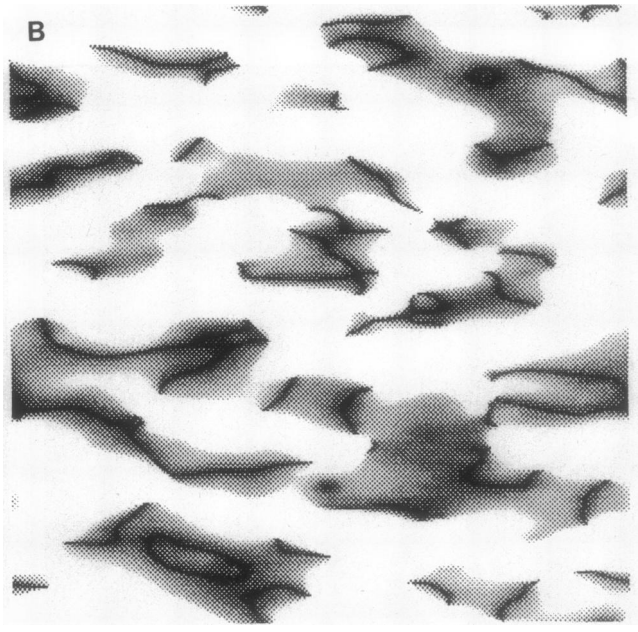


FIG. 3. Typical graph of orientation preference vs. distance along a straight line (electrode track) within the map shown in Fig. 2*A*. The electrode penetration runs tangentially to the cell layer. At left is shown a smooth change of orientation preference. Farther center appear fractures and regions with smooth reversals of orientation shifts.

anisotropic function $h_{rs;kl}$ (Eq. 1). In contrast to the isotropic case (Fig. 2), a long-range order has been established, with most of the iso-orientation slabs and distortions of the position map running parallel to the major axis of $h_{rs;kl}$. Consequently, changes of orientation preference are mainly observable along the direction for which $h_{rs;kl}$ is of short range, whereas along the orthogonal-direction orientation preference tends to vary slowly. The simulation results closely resemble the maps found experimentally in the cat (3, 6) and the tree shrew (7).

For an anisotropy, the different range of $h_{rs;kl}$ along different axes gives rise to two different thresholds for the variance of the secondary-stimulus feature necessary for its appearance in the map. If the variance of the secondary feature is such that the lower threshold, associated with the short-ranged axis, is exceeded but the higher threshold, associated with the long-ranged axis, is not exceeded, periodic distortions will arise; these distortions will lead to representation of the secondary feature in the map. Therefore, the short-ranged direction will be associated with a





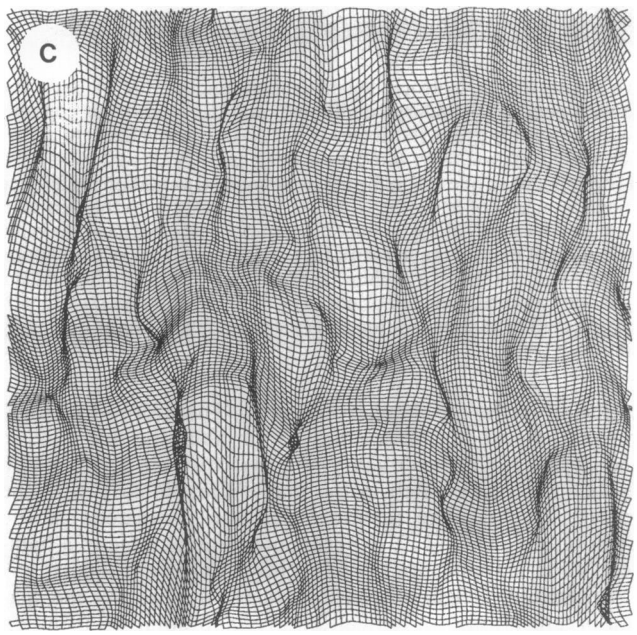


FIG. 4. Final feature map after 30,000 adaptive steps, using an anisotropic neighborhood function ($\sigma_{h1}(t)/\sigma_{h2}(t) = 2$; major axis is horizontal). Parameters of the simulation were $\sigma_{h1}(0) = 300$; $\sigma_{h1}(15{,}000) = 80$; $\sigma_{h1}(30{,}000) = 3$; $\varepsilon_I = 0.09$; $\varepsilon_F = 0.02$; $\sigma_1 = 0.23$; and $\sigma_2 = 0.09$. (*A*) Spatial pattern of orientation preference and selectivity. (*B*) Spatial distribution of cells with orientation preference near the vertical (90°). (*C*) Projection of the network onto the input layer. Visualization of all data is as described for Fig. 2.

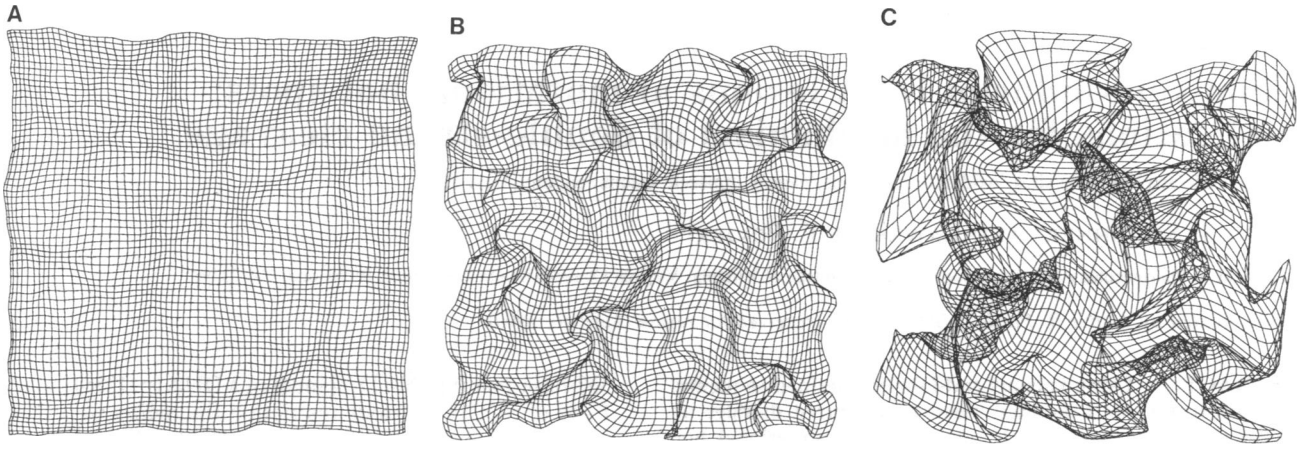


FIG. 5. Final position map, showing gradual disappearance of feature hierarchy for stimuli with eccentricity $\sigma_1/\sigma_2 = 1.7$ (*A*); 2.8 (*B*); and 5 (*C*). The initial connectivity was retinotopic. Parameters of the simulation were as follows: 20,000 iterations, 16,384 units (128 × 128 lattice with free-boundary conditions), $\sigma_{h1}(0) = 40$; $\sigma_{h1}(10{,}000) = 20$; $\sigma_{h1}(20{,}000) = 2$; $\varepsilon_1 = 0.07$; $\varepsilon_F = 0.02$; and $\sigma_1 = 0.16$, 0.23, and 0.35, respectively. The position of every second cell is shown.

variation of the secondary feature, which explains the alignment of the iso-orientation slabs with the major axis of $h_{rs;kl}$.

Origin of Feature Hierarchy. In determining the map location of any particular feature combination the primary factor is the feature position, and the secondary factor is the feature orientation. Because the model does not in any obvious way single out either position or orientation, the question about the origin of the observed feature hierarchy arises.

A two-dimensional map admits two primary features. It has been suggested (9) that these features can be characterized by independence and maximal variance. Further independent features begin to appear as secondary features when their variance exceeds a threshold (18). If, however, their variance becomes comparable with that of the former primary features, the hierarchy breaks down and can even be reversed when the variance of the former secondary features dominates.

For the ellipsoidal stimuli Eq. 2, the variance of the feature position is independent of σ_1 and σ_2 as long as both are well below the size of the input layer. The variance of the orientation feature is zero for circular stimuli, increases with eccentricity, and approaches the variance of the position feature in the limit of infinite eccentricity. This explains why orientation cannot become a primary feature instead of one of the two position coordinates. However, for sufficiently elongated input patterns the variance of orientation, though still smaller, can become comparable to the variance of position, and the feature hierarchy disappears (Fig. 5 *A–C*). The increased distortions and backtrackings that emerge in best position show how the primary role of position gradually gives way to a more equal representation of position and orientation as the eccentricity of the patterns enlarges.

This relationship leads to a conclusion for the case that the formation of the visual map employs externally or internally (as proposed in ref. 14) generated activity patterns driving a process of a type described by Eq. 1. The eccentricity of those 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.

Conclusion. The close resemblance found for the maps obtained from the process described by Eq. 1 and those formed in the visual cortex of various animals suggests that many aspects of cortical feature maps are a consequence of a principle of continuous mapping. The particular realization by Eq. 1 of this principle suggests that an adaptive process, either pre- or postnatally driven, may help to shape the structure of cortical maps.

Note. After completion of this work, an independent study of a theoretical framework for cortical feature maps was published (19). Although both approaches share important aspects, significant differences are also present: (*i*) The present work avoids the use of (low-dimensional) preprocessed feature vectors. Instead, the map-formation process is based on the original, high-dimensional stimulus patterns solely. (*ii*) Differences between the orientation map of the monkey and of the cat are discussed and reproduced by the present model.

This research has been supported by the University of Illinois at Urbana–Champaign. Computer time on the Connection Machine CM-2 has been made available by the National Center for Supercomputer Applications. The authors thank R. Kufrin and G. Quinn for their help and support in all technical matters concerning the use of the Connection Machine system and the Boehringer Ingelheim Fonds for support of this project by a scholarship to K.O.

1. Blasdel, G. G. & Salama, G. (1986) *Nature (London)* **321,** 579–585.
2. van Essen, D. (1985) in *Cerebral Cortex*, eds. Peters, A. & Jones, E. G. (Plenum, New York), Vol. 3, pp. 259–329.
3. Löwel, S., Freeman, B. & Singer, W. (1987) *J. Comp. Neurol.* **255,** 401–415.
4. Hubel, D., Wiesel, D. N. & Stryker, P. N. (1978) *J. Comp. Neurol.* **177,** 361–380.
5. Hubel, D. H. & Wiesel, T. N. (1974) *J. Comp. Neurol.* **158,** 267–294.
6. Swindale, N. V., Matsubara, J. A. & Cynader, M. S. (1987) *J. Neurosci.* **7,** 1414–1427.
7. Skeen, L. C., Humphrey, A. L., Norton, T. T. & Hall, W. C. (1978) *Brain Res.* **142,** 538–545.
8. Takeuchi, A. & Amari, S. (1979) *Biol. Cybern.* **35,** 63–72.
9. Kohonen, T. (1983) *Self-Organization and Associative Memory* (Springer, New York).
10. von der Malsburg, C. (1973) *Kybernetik* **14,** 85–100.
11. Willshaw, D. J. & von der Malsburg, C. (1976) *Proc. R. Soc. London Ser. B.* **194,** 431–445.
12. Braitenberg, V. & Braitenberg, C. (1979) *Biol. Cybern.* **33,** 179–186.
13. Swindale, N. V. (1982) *Proc. R. Soc. London. Ser. B* **215,** 211–230.
14. Linsker, R. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 8779–8783.
15. Kohonen, T. (1982) *Biol. Cybern.* **43,** 59–69.
16. Kohonen, T. (1982) *Biol. Cybern.* **44,** 135–140.
17. Silito, A. M. (1984) in *Cerebral Cortex*, eds. Peters, A. & Jones, E. G. (Plenum, New York), Vol. 2, pp. 91–117.
18. Ritter, H. & Schulten, K. (1989) *Biol. Cybern.* **60,** 59–71.
19. Durbin, R. & Mitchison, M. (1990) *Nature (London)* **343,** 644–647.