# Neural pattern formation via a competitive Hebbian mechanism

K. Obermayer [a,*], T. Sejnowski [a], G.G. Blasdel [b]

[a] *Howard Hughes Medical Institute and Salk Institute, La Jolla, CA 92037, USA*
[b] *Department of Neurobiology, Harvard Medical School, Boston, MA 02115, USA*

## Abstract

In this contribution we investigate a simple pattern formation process [9,10] based on Hebbian learning and competitive interactions within cortex. This process generates spatial representations of afferent (sensory) information which strongly resemble patterns of response properties of neurons commonly called brain maps. For one of the most thoroughly studied phenomena in cortical development, the formation of topographic maps, orientation and ocular dominance columns in macaque striate cortex, the process, for example, generates the observed patterns of receptive field properties including the recently described correlations between orientation preference and ocular dominance. Competitive Hebbian learning has not only proven to be a useful concept in the understanding of development and plasticity in several brain areas, but the underlying principles have have been successfully applied to problems in machine learning [22]. The model's universality, simplicity, predictive power, and usefulness warrants a closer investigation.

*Key words:* Self-organization; Neural development; Neural network; Computer simulation; Cortical map; Orientation column; Ocular dominance column; Macaque

## 1. Introduction

Pattern formation by self-organization is a common phenomenon in brain development. The enormous number of neurons and their connections makes it impossible for organisms to completely specify neural connectivity patterns within the genome. Instead organisms seem to code for processes which interact with themselves and with the environment thereby generating the patterns which are observed in the brain. These processes are supposedly much simpler to describe than the actual patterns, and it is our hope that there are only few and that they can be cast into a small set of simple rules.

One prominent pattern formation process is the ability of the brain to form spatial representations of sensory information. These patterns are commonly called brain maps and are characterized by the facts (i) that neurons within them respond strongly to certain features in the afferent signals, and (ii) that neurons with different response properties are arranged across cortex in a system-

atic fashion. Well known examples are the multiple topographic representations of the body surface in the somatosensory areas [8], the various types of acoustic maps in the auditory areas [20], and the retinotopic maps [26], as well as orientation and ocular dominance columns [3] in the primary visual areas of several species. Modelling studies have shown that the formation and the plasticity properties of brain maps can be described by simple dynamical systems (see e.g. [14,18,25]). It thus seems that the phenomenon of brain maps and their formation provide an example for our hypothesis: that a small set of 'primitives' of information processing and development exists which could account for the variety of phenomena encountered in brain development and function.

In the following we will investigate a promising candidate for such a developmental primitive: a simple pattern formation process which was originally invented by T. Kohonen [9,10]. This process describes certain activity dependent developmental processes in the neocortex and has proven to be particularly successful in modelling the distribution of neurons with various response properties in so called cortical maps. In section 2 we will introduce the process and describe a few of its properties. In section 3 we will describe a particularly well studied application: the

* Corresponding author. Present address: Technische Fakultät, Universität Bielefeld, Universitätsstrasse 25, 33615 Bielefeld, Germany. Fax: (49) (521) 106-6011. e-mail: oby@techfak.uni-bielefeld.de

formation of orientation and ocular dominance columns in monkey striate cortex. Section 4 will summarize our findings.

## 2. A simple self-organizing process

Let us consider an idealized neural structure whose properties are affected by incoming sensory signals (Fig. 1). The structure is two-dimensional in order to account for the columnar organization of neural response properties found in several cortical areas. We assume that the neural structure responds to afferent activity patterns, for example to sensory signals, and we describe its response properties at each location by a set of parameters, which we will call 'receptive field properties' from now on. These parameters change their values as a result of sensory stimulation: afferent signals generate an activation pattern across the neural structure and both, the afferent signals and the activation pattern, drive the parameters slowly towards their stationary values. The final, stable pattern of receptive field properties in our model is then identified with the mature brain map, while the path leading to this pattern is identified with the process of cortical map development.

Next we have to make a few assumptions about the nature of this process. These are:
1. Locality: the change in receptive field properties depends only on the local sensory input and the local activiation at a given time. Hebb's rule [7], which we will use below, is local.
2. Global competition: the lateral interactions between distant regions in the neural structure are such that neurons in these regions effectively inhibit each other. These type of interactions enhance the activation of already highly active areas at the expense of areas which are activated less.
3. Local cooperation: the lateral interaction between closeby regions in the neural structure is such that neu-
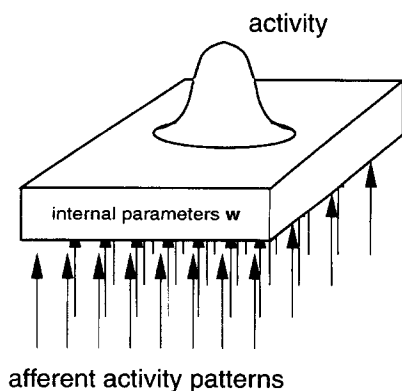


Fig. 1. Idealized neural structure.

rons effectively excite each other. These types of interactions lead to the spread of activity laterally.

Formally [16], we describe receptive field properties at each location $\vec{r}$ in the neural structure by a feature vector $\vec{w}_{\vec{r}}$, whose components denote receptive field properties. In our example from macaque visual cortex, which is shown in section 3, the components denote the position of the receptive field in retinotopic space, orientation preference, orientation tuning strength and ocular dominance. Afferent signals are described by a pattern vector $\vec{v}$ of similar dimension which is drawn at random according to a given probability distribution $P(\vec{v})$. Pattern formation is an iterative process and receptive field properties are changed at every step according to:

$$\frac{d\vec{w}_{\vec{r}}}{dt} = \int \vec{v} h(\vec{r}, \vec{s}(\vec{v})) P(\vec{v})d\vec{v} - \vec{w}_{\vec{r}} \int h(\vec{r}, \vec{s}(\vec{v})) P(\vec{v})d\vec{v},$$

(1)

where

$$\vec{s}(\vec{v}) = \min_{\vec{r}} |\vec{v} - \vec{w}_{\vec{r}}|.$$

(2)

$\vec{s}(\vec{v})$ is the location in the map where the neurons respond best to the stimulus $\vec{v}$. The function $h(\vec{r},\vec{s}(\vec{v}))$ describes the 'activity' at location $\vec{r}$ for the presentation of pattern $\vec{v}$. It will be greatest (and the learning fastest) for locations near the best response in the map. Rule 2 is called 'winner take all'-rule and formally implements global competition, while the finite range of the function $h(\vec{r},\vec{s})$ ensures local cooperation. Locality, finally, is implemented by rule 1, which is a Hebbian learning rule with a decay term. $\vec{v}$ is the presynaptic term, $h(\vec{r},\vec{s}(\vec{v}))$ is the postsynaptic term, and the decay term limits the learning when the network reaches a mature state. If $h(\vec{r},\vec{s})$ is a Gaussian function of the argument $\vec{r} - \vec{s}$, the process (1)–(2) is amenable to mathematical analyses [4,5,16,22,24].

Note, that we implicitly assume that suitable lateral interactions are present before the cortical map is mature. We have not made any statement about the nature of those interactions, however: activity could be electrical as well as chemical, and interactions between cells could arise due to synaptic connections between neurons, due to electrical coupling, or due to diffusive substances. It is important, however, that the coupling between neurons is nonlinear and gives rise to global competition and local cooperation.

Given these assumptions and an appropriate implementation, receptive field properties and their (stationary) distribution across the neural structure are determined by the properties of the ensemble of afferent patterns. When these afferent patterns are generated by localized stimuli on a sensory surface, for example, the process generates spatial receptive fields and a topographic map of this surface [17]; when these patterns are generated by a moving loud-

speaker in front of a stereo microphone, the same process generates receptive fields tuned to intensity differences between both microphones and a topographic map of space

based on these acoustic signals [22], and when the afferent patterns encode suitably chosen text fragments, the process generates semantic maps [21], in which 'receptive
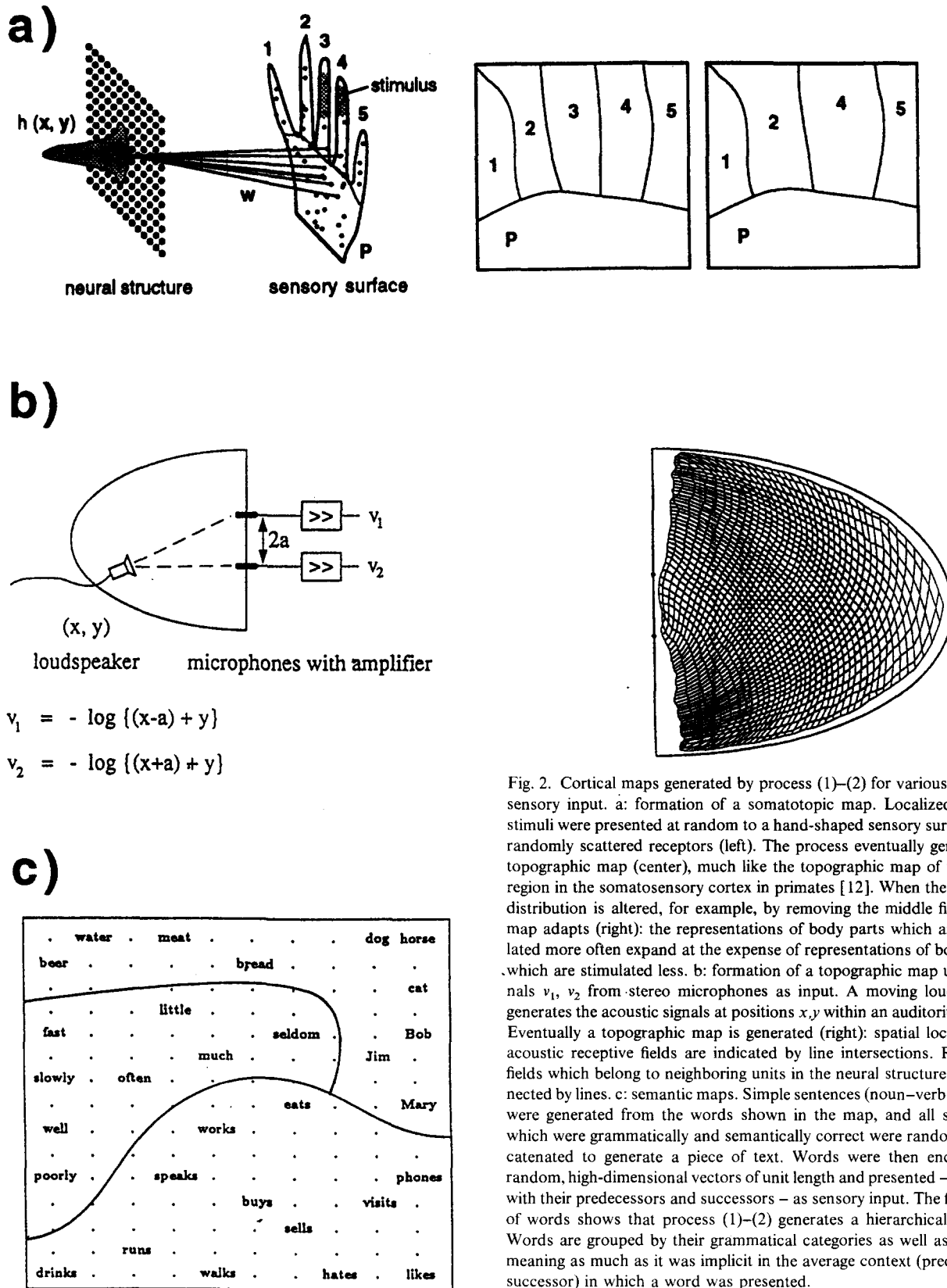


a)

h (x, y)

neural structure          sensory surface

b)

(x, y)

loudspeaker          microphones with amplifier

$v_1 = - \log \{(x-a) + y\}$

$v_2 = - \log \{(x+a) + y\}$

c)



Fig. 2. Cortical maps generated by process (1)–(2) for various kinds of sensory input. a: formation of a somatotopic map. Localized 'tactile' stimuli were presented at random to a hand-shaped sensory surface with randomly scattered receptors (left). The process eventually generates a topographic map (center), much like the topographic map of the hand region in the somatosensory cortex in primates [12]. When the stimulus distribution is altered, for example, by removing the middle finger, the map adapts (right): the representations of body parts which are stimulated more often expand at the expense of representations of body parts which are stimulated less. b: formation of a topographic map using signals $v_1$, $v_2$ from stereo microphones as input. A moving loudspeaker generates the acoustic signals at positions x,y within an auditorium (left). Eventually a topographic map is generated (right): spatial locations of acoustic receptive fields are indicated by line intersections. Receptive fields which belong to neighboring units in the neural structure are connected by lines. c: semantic maps. Simple sentences (noun–verb–adverb) were generated from the words shown in the map, and all sentences which were grammatically and semantically correct were randomly concatenated to generate a piece of text. Words were then encoded as random, high-dimensional vectors of unit length and presented – together with their predecessors and successors – as sensory input. The final map of words shows that process (1)–(2) generates a hierarchical pattern: Words are grouped by their grammatical categories as well as by their meaning as much as it was implicit in the average context (predecessor, successor) in which a word was presented.
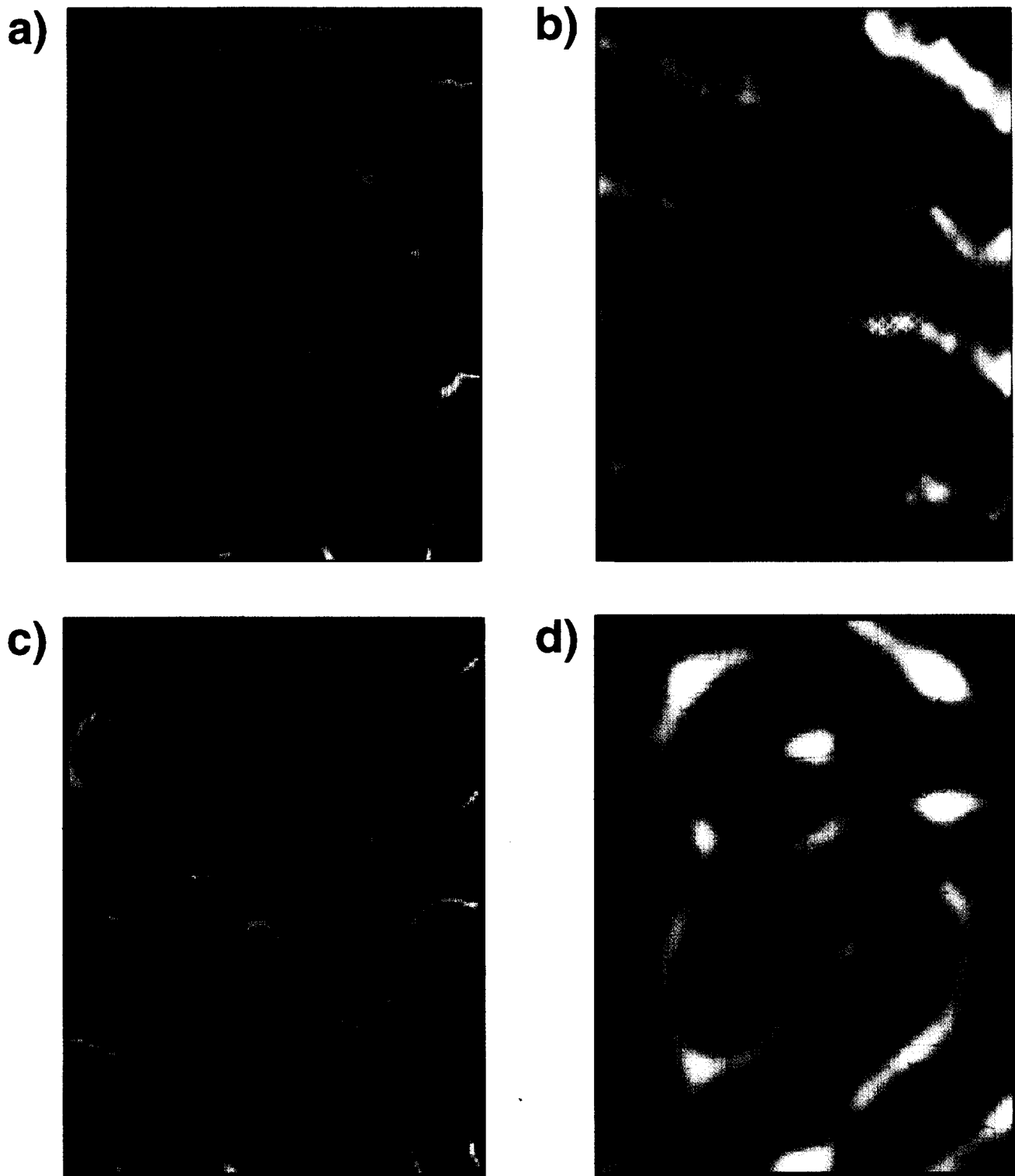
Fig. 3. Orientation and ocular dominance maps found in the upper layers of macaque striate cortex in comparison with the corresponding patterns generated by process (1)–(2) for $d_{OR}$ and $d_{OD}$ above threshold. a: spatial pattern of orientation preference. The data were recorded optically [3] from area 17 of an adult macaque and show a region 3.1 mm × 4.3 mm in size, located close to the border to area 18 and close to midline. Hue values code for orientation preferences, where the interval of 180° for orientation preferences is represented by the colors of the color circle. Complementary colors denote orthogonal orientation preferences. Arrows indicate singularities (2), linear zones (1), saddle points (3) and fractures (4). b: spatial pattern of ocular dominance. The pattern was recorded optically from the same cortical region as shown in (a). Bright and dark areas denote preferences for the ipsi- and the contralateral eye, respectively. c: orientation preference map generated by process (1)–(2) for the parameters given in the text. Arrows indicate singularities (1), linear zones (2), saddle points (4) and fractures (3). d: ocular dominance map generated by process (1)–(2) for the parameters given in the text. c and d show an enlarged section of the complete pattern, 90 × 120 elements in size.

fields' for words develop, which are then spatially grouped according to grammatical and semantical categories. Fig. 2 shows a few results from numerical simulations: one and the same simple self-organizing process is able to generate a large variety of patterns depending on the afferent signals it acts on.

## 3. Comparison with experiments: formation of orientation and ocular dominance columns

In this section we will study the case of five properties characterizing the response properties of neurons in primary visual cortex: position of a receptive field in retinotopic space, orientation preference and and orientation tuning strength and ocular dominance.

Receptive field properties at each cortical location $\vec{r}$ are described by a five-dimensional vector $\vec{w}_{\vec{r}}$, whose components denote the position $(x_{\vec{r}}, y_{\vec{r}})$ of the receptive field in retinotopic space, orientation preference $\Theta_{\vec{r}}$ and tuning strength $q_{\vec{r}}$ in their cartesian coordinates $(q_{\vec{r}} \cos 2 \Theta_{\vec{r}}, q_{\vec{r}} \sin 2 \Theta_{\vec{r}})$ [25], and ocular dominance $z_{\vec{r}}$. Afferent signals are described by a similar five-dimensional vector $\vec{v}$ cho-

sen at random and with equal probability from a manifold given by: $x,y \in [0, 512]$, $q \leq 20$, $z \in [-15, 15]$. Numerical simulations of process (1)–(2) were carried out on a CM5 parallel computer using a neural structure of $512 \times 512$ units and periodic boundary conditions. Eqn. 1 was solved using stochastic approximation, i.e., the change of the feature vectors was calculated using:

$$\vec{w}_{\vec{r}}(t + 1) = \vec{w}_{\vec{r}}(t) + \varepsilon h(\vec{r}, \vec{s}(\vec{v}))(\vec{v}(t + 1) - \vec{w}(t)), \quad \varepsilon < 1. \tag{3}$$

Parameters were: $\varepsilon = 0.05$, $9 \cdot 10^7$ iterations, retinotopic initial conditions. $h(\vec{r},\vec{s})$ was a Gaussian function with half-width $\sigma_h = 5$.

Fig. 3 shows a comparison between an orientation and ocular dominance column systems generated by process (1)–(2) and experimental data recorded optically from the superficial layers of macaque striate cortex. From visual comparison in becomes apparent that at least all of the prominent elements of organization are present [1,2,15]. The orientation maps exhibit singularities and fractures, as well as linear zones and saddle points.

Singularities are pointlike discontinuities in the map around which orientation preference change by $\pm 180°$ around a closed path. They are presumably part of the
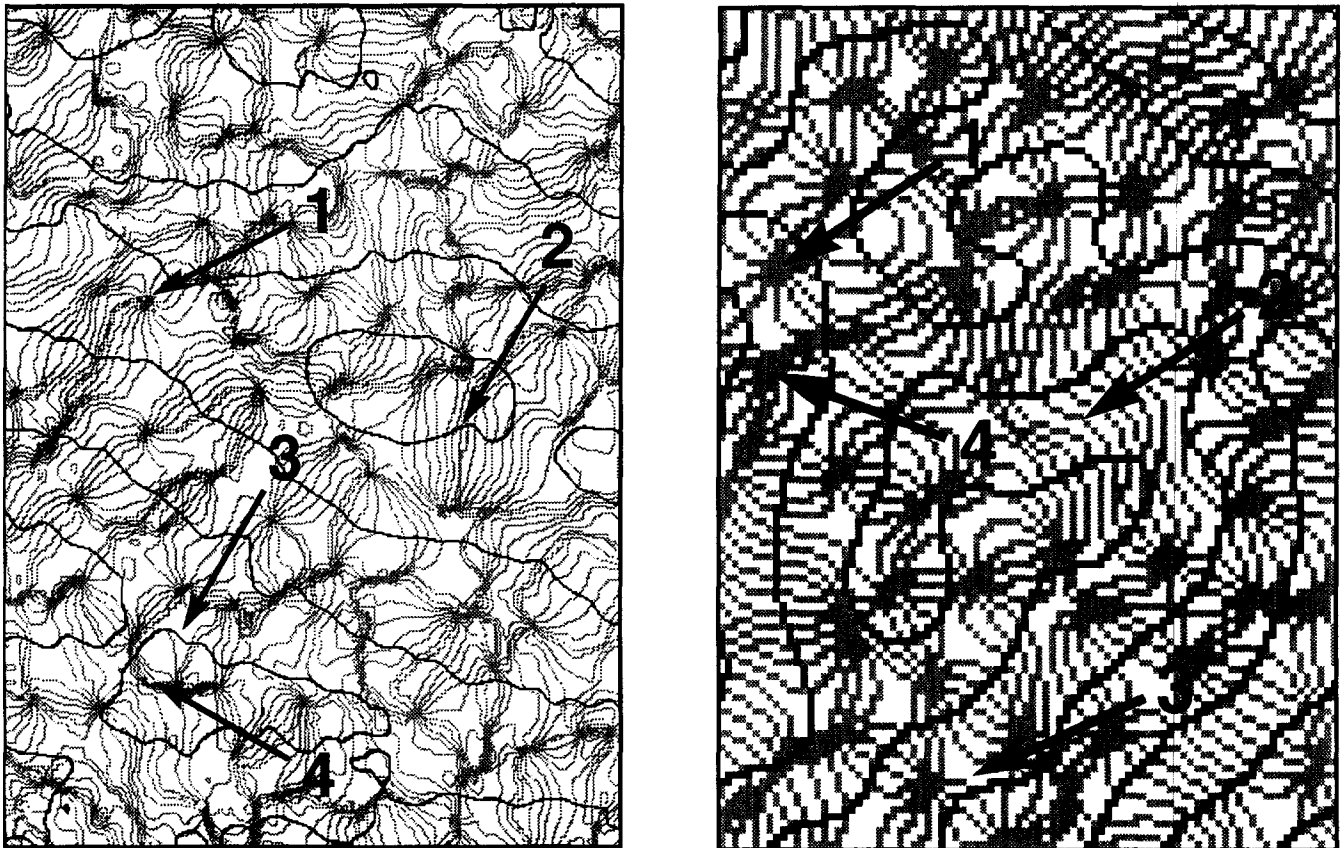


Fig. 4. Contour plots for the orientation maps shown in Fig. 4 in overlay with the borders of the ocular dominance bands. Iso-orientation lines are shown in gray. The borders of the ocular dominance bands are marked black. a: orientation and ocular dominance in macaque striate cortex. Iso-orientation contours indicate intervals of 11.25°. b: orientation and ocular dominance generated by process (1)–(2). Iso-orientation contours indicate intervals of 18°. Arrows indicate singularities (1), linear zones (2), saddle points (3) and fractures (4).

macaque's texture and color processing system. Linear zones are areas up to 1 mm in size, within which parallel iso-orientation bands exist. They are presumably part of a contour detection system. Saddle points are patches of similar orientation preference where orientation preferences change clockwise along one axis for movements in either direction from the middle and counterclockwise along the orthogonal axis, while fractures are line-like features across which orientation preferences change rapidly. The functional role of the latter features is not known.

The ocular dominance maps show the typical pattern of locally parallel bands which sometimes branch and terminate. A more careful analysis reveals that the patterns are characterized by global disorder as well as by a typical spatial frequency and that, consequently, power spectra of the orientation map show a pronounced peak [15,16]. Interestingly, there is no dependence of orientation preference on the spatial layout across cortex, an artefact which has plaqued models of development [11,13,27] using linear interaction functions [6].

Fig. 4 shows a contour plot of the pattern of orientation preference in overlay with the borders of ocular dominance bands. The predicted patterns are characterized by correlations between the orientation preference and ocular dominance maps which are similar to the correlations observed in the experimental data: singularities have a tendency to align with the centers of ocular dominance bands and in the linear zones, where parallel iso-orientation bands exist, these bands intersect ocular dominance bands at approximately right angles [15,16].

Let us study the stationary patterns as a function of the ensemble of afferent patterns more formally. The ensemble of afferent patterns—each of which is characterized by five parameters—forms a cloud of points in a five-dimensional space. Let $m_j$ denote its second order moment along axis $j$. For reason of simplicity let us also assume that the principal axes of this cloud are colinear with the coordinate axes, and that $m_1 = m_2$ and $m_3 = m_4$. The ratios $d_{OR} = m_3/m_1$ and $d_{OD} = m_5/m_1$ are then order parameters of the process [16] (Fig. 5): If both of them are smaller than a certain threshold the process generates a topographic map but no orientation and ocular dominance columns. If $d_{OR}$ $(d_{OD})$, however, is above threshold orientation columns (ocular dominance columns) form and the topographic map is distorted locally.

## 4. Final comments

We have investigated a simple self-organizing process which is based on a few general assumptions, among them local learning rules, and effective interactions between neurons which are based on global competition and local
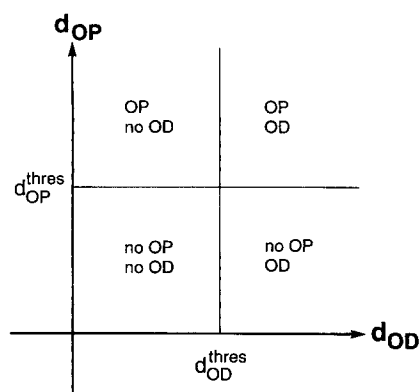


Fig. 5. The four parameter regimes of the process (1)–(2). $d^{OP}$ and $d^{OD}$ both have to be above threshold for stable orientation (OP) and ocular dominance (OD) columns to form.

cooperation. This process acts on an ensemble of afferent patterns and produces a spatial map of neural response properties based on the nature of these signals. For example, the map of features in the macaque's area V2 will be organized quite differently from that in V1 because there are different afferent patterns which drive map formation. Process (1)–(2) can also account–at least qualitatively–for some phenomena in cortical plasticity as have been investigated in the deprivation experiment paradigm [17,19,23]. The model's universality, simplicity and predictive power provide evidence for the hypothesis that there exist only a few simple mechanisms which account for a variety of phenomena encountered in brain development and function.

## Acknowledgements

## References

[1] Blasdel, G.G., Differential imaging of ocular dominance and orientation selectivity in monkey striate cortex, J. Neurosci., 12 (1992) 3117–3138.
[2] Blasdel, G.G., Orientation selectivity, preference, and continuity in monkey striate cortex, J. Neurosci., 12 (1992) 3139–3161.
[3] Blasdel, G.G. and Salama, G., Voltage sensitive dyes reveal a modular organization in monkey striate cortex, Nature, 321 (1986) 579–585.
[4] Erwin, E., Obermayer, K. and Schulten, K., Self-organizing maps:

ordering, convergence properties and energy functions, *Biol. Cybern.*, 67 (1992) 47–55.

[5] Erwin, E., Obermayer, K. and Schulten, K., Self-organizing maps: stationary states, metastability and convergence rate, *Biol. Cybern.*, 67 (1992) 35–45.

[6] Erwin, E., Obermayer, K. and Schulten, K., Models of orientation and ocular dominance columns in the visual cortex: a critical comparison, in press.

[7] Hebb, D., *The Organisation of Behaviour*, Wiley, New York, 1949.

[8] Kaas, J.H., Nelson, R.J., Sur M., Lin, C.S. and Merzenich, M.M., Multiple representations of the body within the primary somatosensory cortex of primates, *Science*, 204 (1979) 521–523.

[9] Kohonen, T., Analysis of a simple self-organizing process, *Biol. Cybern.*, 44 (1982) 135–140.

[10] Kohonen, T., Self-organized formation of topologically correct feature maps, *Biol. Cybern.*, 43 (1982) 59–69.

[11] Linsker, R., From basic network principles to neural architecture: emergence of orientation columns, *Proc. Natl. Acad. Sci. USA*, 83 (1986) 8779–8783.

[12] Merzenich, M.M., Nelson, R.J., Kaas, J.H., Stryker, M.P., Jenkins, W.M., Zook, J.M., Cynader, M.S. and Schoppmann, A., Variability in hand surface representations in areas 3b and 1 in adult and squirrel monkeys, *J. Comp. Neurol.*, 258 (1987) 281–296.

[13] Miller, K.D., A model for the development of simple cell receptive fields and the ordered arrangements of orientation columns through activity-dependent competition between on- and off-center inputs, *J. Neurosci.*, 14 (1994) 409–441.

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

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

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

[17] Obermayer, K., Ritter, H. and Schulten, K., Large-scale simulations of self-organizing neural networks on parallel computers: application to biological modelling, *Par. Comp.*, 14 (1990) 381–404.

[18] 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 (1990) 8345–8349.

[19] Obermayer, K., Ritter, H. and Schulten, K., A model for the development of the spatial structure of retinotopic maps and orientation columns, *IEICE T Fune A*, 75 (1992) 537–545.

[20] Olsen, J.F., Knudsen, E.I. and Esterly, S.D., Neural maps of interaural time and intensity differences in the optic tectum of the barn owl, *J. Neurosci.*, 9 (1989) 2591–2605.

[21] Ritter, H. and Kohonen, T., Self-organizing semantic maps, *Biol. Cybern.*, 61 (1989) 241–254.

[22] Ritter, H., Martinetz, T. and Schulten. K., *Neural Computation and Self-Organizing Maps: An Introduction*, Addison–Wesley, New York, 1992.

[23] Ritter, H., and Schulten, K., On the stationary state of Kohonen's self-organizing mapping, *Biol. Cybern.*, 54 (1986) 99–106.

[24] Ritter, H., and Schulten, K., Convergence properties of Kohonen's topology conserving maps: fluctuations, stability, and dimension selection, *Biol. Cybern.*, 60 (1988) 59–71.

[25] Swindale, N.V., A model for the formation of orientation columns, *Proc. R. Soc. Lond. B*, 215 (1982) 211–230.

[26] Tusa, R.J., Palmer, L.A. and Rosenquist, A.C., The retinotopic organization of area 17 (striate cortex) in the cat. *J. Comp. Neurol.*, 193 (1978) 213–236.

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