ELSEVIER

Neurocomputing 32–33 (2000) 83–90

NEUROCOMPUTING

www.elsevier.com/locate/neucom

Dynamics of cortical map development in the elastic net model☆

Andrei Cimponeriu[a,b], Geoffrey J. Goodhill[c,*]

[a]*Mount Sinai School of Medicine, Department of Ophthalmology, 1 Gustave, L. Levy Place, Box 1183, New York, NY 10029-6500, USA*

[b]*Department of Electronics and Telecommunications, Polytechnic University of Timisoara, Blv. V.Parvan 2, 1900 Timisoara, Romania*

[c]*Georgetown Institute for Cognitive and Computational Sciences, Georgetown University Medical Center, 3970 Reservoir Road NW, Washington DC 20007, USA*

Accepted 13 January 2000

Abstract

We examine the dynamics of development of cortical maps of spatial position, ocular dominance and orientation preference in the elastic net model. Analysis and simulation results are presented for the order in which each of these maps first emerges as a function of the parameters of the feature space.

Keywords: Visual cortex; Ocular dominance columns; Orientation columns; Feature space

1. Introduction

Feature-space models of cortical map development based on the elastic net and Kohonen algorithms [10,5,16] currently give the closest match to experimental data regarding the spatial structure of orientation (OR) and ocular dominance (OD) maps [8,21]. Hoffsümmer et al. [12,13] have investigated the dynamics of the joint development of these maps in a continuous version of the elastic net model. Biologically, it is possible that OR and OD may develop in a different order in different species, and Hoffsümmer et al. found that the final map structure in the model depends on this order of development. Here we analyze how the ordering of map development is

☆This work was supported by NSF grant IBN-9808364.

* Corresponding author.

E-mail address: geoff@giccs.georgetown.edu (G.J. Goodhill).

0925-2312/00/$ - see front matter
PII: S0925-2312(00)00147-8

controlled by the parameters of the feature space in a discrete version of the elastic net by extending Durbin et al. [6] analysis of the behavior of the net for the traveling salesman problem to its behavior in high-dimensional feature spaces. The resulting predictions compare favorably with simulation results.

2. The elastic net model of cortical mapping

The elastic net is a biologically motivated method for solving combinatorial optimization problems such as the traveling salesman problem (TSP) [7]. It has been reviewed many times (e.g. [8,21]) and therefore it will only be described briefly here. It consists of a set of $\boldsymbol{y}_j$, $j = 1, \ldots, N$ vectors, modeling the receptive fields of cells in the primary visual cortex (V1). The net learns a set of $\boldsymbol{x}_i$, $i = 1, \ldots, M$ prototypes, which represent the visual input. Both the prototypes and the cortical cell vectors consist of several features, specifically the x and y retinal position, the ocularity, and the OR preferred angle and "strength" as introduced by Swindale [20]. Learning is performed by minimizing an energy function E given by

$$E = -\alpha k \sum_i \log \sum_j \Phi(\boldsymbol{x}_i, \boldsymbol{y}_j, k) + \frac{\beta}{2} \sum_{j' \in Nb(j)} \|\boldsymbol{y}_{j'} - \boldsymbol{y}_j\|^2, \tag{1}$$

where α and β are constants, $Nb(j)$ represents $\boldsymbol{y}_j$'s neighbors, and Φ is given by

$$\Phi(\boldsymbol{x}_i, \boldsymbol{y}_j, k) = \exp\left(\frac{-\|\boldsymbol{x}_i - \boldsymbol{y}_j\|^2}{2k^2}\right). \tag{2}$$

The parameter k gives the extent of the neighborhood in the input space over which each $\boldsymbol{y}_j$ is excited, i.e. the extent of its receptive field. In the "annealed" version of the elastic net model, k is gradually reduced; since it acts as a scale for the difference between $\boldsymbol{x}_i$ and $\boldsymbol{y}_j$'s features, this can be seen by increasing the power of a magnifying glass through which the net's vectors can be distinguished from the prototypes; the smaller k, the greater the magnification, and the smaller the distance between them, minimized by the relaxation of E. To solve the TSP problem, $M \leqslant N$; in this case, when k becomes very small, each $\boldsymbol{y}_j$ is matched to an $\boldsymbol{x}_i$ [6,19]. Meanwhile, the "length" of the cortical map, given by the second term in the energy function, is kept small, requiring that neighboring cortical cells learn close features. The elastic net is closely related to other models based on local learning rules [18,24,4,25], and implements a form of deterministic annealing [17].

Two possible ways in which the prototypes can be distributed is randomly, sampling from a continous distribution of feature values (e.g. [13,23]), or regularly, where feature values are evenly spaced (e.g. [10,11]). Since the interpretation of the input space is somewhat abstract, it is hard to give biological justifications for selecting one type of distribution over the other. Here we use a regular distribution; this allows some simple expressions to be derived for how the order of development depends in the feature space parameters.

3. Analysis

Durbin et al. [6] analyzed the dynamics of the elastic net as applied to the TSP. They showed that for a sufficiently large value of k the energy function has one minimum, with all points in the elastic net at the origin. As k is reduced the energy function bifurcates. These bifurcation points occur when the Hessian matrix of the energy function becomes singular, i.e. has a zero eigenvalue (for large k all eigenvalues are positive). At the first bifurcation, the net starts to expand along the direction of the principal eigenvector of the covariance matrix of the set of prototypes; a first map has emerged. As k is reduced further the energy function bifurcates again and a second map emerges, and so on. This type of behavior has also been studied for more general deterministic annealing methods (reviewed in [17]).

The full Hessian is complicated; however, Durbin et al. showed that when the net lies entirely at the origin, $\{\boldsymbol{y}_j = 0\}_{j=1\ldots N}$, the Hessian has a much simpler form and the prototypes distribution enters only via its covariance matrix. Let $\lambda^{\mathrm{E}}_{\min}$ be the smallest eigenvalue of the Hessian and $\lambda^{\mathrm{P}}_{\max}$ the largest eigenvalue of the city/prototypes covariance matrix. Then:

$$\lambda^{\mathrm{E}}_{\min} = \frac{\alpha N}{kM} - \frac{\alpha \lambda^{\mathrm{P}}_{\max}}{k^3}\frac{N}{M} + 8\beta \sin^2\left(\frac{\pi}{M}\right). \tag{3}$$

(Note that there are some typos in Durbin et al. [6] which are corrected here.) Thus the value of k for which $\lambda^{\mathrm{E}}_{\min} = 0$ gives the point at which the net breaks out from the origin. For k small this condition becomes simply $k = \sqrt{\lambda^{\mathrm{P}}_{\max}}$ (this relation can also be derived from a more general consideration of deterministic annealing [17]).

3.1. The first emerging map

For our analysis, the prototypes are placed on a five-dimensional regular grid. The individual components of the grid are the x and y positions, evenly-spaced with a step d to a total of unit length (i.e. $n = 1 + 1/d$ positions in each direction), $\mathrm{OD} \in \{-l, +l\}$, and $r\sin(\theta)$, $r\cos(\theta)$, where θ takes m evenly-spaced values on $[0, 2\pi)$, and r is the "strength" of OR tuning (since orientation preference is π periodic, we use the standard device of doubling this angle to give an appropriate 2π periodicity). All these components are mutually orthogonal. This regular distribution of the prototypes allows the covariance matrix and its eigenvalues to be calculated analytically in terms of d, l and r, and thus analytical values to be derived for the critical k value at which the first map starts to expand. Because of the orthogonality of the components the covariance matrix is diagonal. It can be straightforwardly obtained as

$$C = \begin{pmatrix} \frac{n^2 d_x^2}{12} & 0 & 0 & 0 & 0 \\ 0 & \frac{n^2 d_y^2}{12} & 0 & 0 & 0 \\ 0 & 0 & \frac{r^2}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{r^2}{2} & 0 \\ 0 & 0 & 0 & 0 & l^2 \end{pmatrix}, \tag{4}$$

where the first two rows/columns give the variance in the x, y directions, the next two the variance along the polar coordinates of the (r, θ) plane, and the last the variance along the OD direction. For generality we have included the possibility that the spacing of feature points along the x and y component may be different (cf. [9]). The critical values k_{xy}, k_{or} and k_{od} at which initial movement from the origin occurs along each of these dimensions are thus

$$k_{xy} = \frac{nd}{\sqrt{12}}, \qquad k_{\mathrm{or}} = \frac{r}{\sqrt{2}}, \qquad k_{\mathrm{od}} = l. \tag{5}$$

Note the following:

- If the spatial dimensions are of unit length then $nd \approx 1$ and k_{xy} becomes $\sqrt{1/12} = 0.29$. Our previous applications of the elastic net to the formation of ocular dominance columns have usually started at $k_{\mathrm{init}} = 0.2$ with $l < k_{\mathrm{init}}$ [10,11,9]. Thus, as expected, in those cases we saw an immediate expansion along the spatial dimensions without expansion along the ocular dominance dimension.
- These values allow conditions to be derived on d, r and l so as to ensure that one particular feature map develops first. Biologically it is clear that the topographic map develops first, thus the biologically appropriate parameter regime (in the limit of small d) is $r < 0.41$ and $l < 0.29$.

3.2. Subsequent maps

Once it is true that $\{\boldsymbol{y}_j \neq 0\}_{j=1,\ldots,N}$ the subsequent null eigenvalues of the Hessian are no longer easy to compute analytically. However, the bifurcation points can be studied by simulation as described below. Since the Hessian depends on the $\boldsymbol{y}_j$ values, its eigenvalues would have to be calculated during the simulation.

4. Simulations

We used the following simulation parameters: $\alpha = 0.2$, $\beta = 2.0$, k was started at an initial value of 0.5 and was subsequently reduced by multiplying it by 0.99 after each iteration, $d = 0.05$, $m = 6$, $r \in [0.08, 0.28]$, $l \in [0.08, 0.2]$. The cortical sheet was a square array of size 72×72 cells, so that there were roughly the same number of cortical cells (5, 184) as prototype vectors (5, 292). The neighborhood of a cortical cell was taken to be its four nearest-neighbors in the cortical array. The net was started with random OR and OD components and a crude retinotopy.

As predicted analytically, the net initially collapsed to a point at the center of the feature space. We examined the development of the subsequent maps by observing each feature separately. As k decreased to its first critical value the retinotopic map started to emerge at $k = 0.29$, as predicted by Eq. (5). Fig. 1 shows a plot of the "length" term of the energy function within the biological parameter regime. Initially the length becomes very small as the net collapses to a point. It then expands again as the retinotopic map forms; the bottom of the well corresponds to k_{xy}. Expansion then

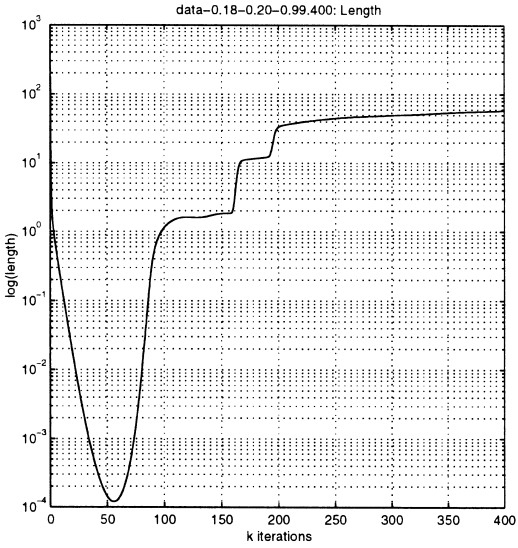

Fig. 1. The evolution of the "length" term of the energy function in a typical simulation, ($l = 0.18$, $r = 0.20$). The retinotopic map emerges first, followed by the OD and OR maps, respectively.

slows down until a sudden jump when the second map forms at the next critical value (in this case OD). A final jump occurs when the next map forms at the next critical value (in this case OR). Beyond this point k is very small and the maps remain quite stable.

We also ran some simulations in the non-biological regime where the OR or OD map form before the retinotopic map. As predicted by our calculations, whereas if $r < 0.41$ and $l < 0.29$ the retinotopic map emerged first, if $r > 0.41$ and $l < r/\sqrt{2}$ the OR map emerged first (i.e. each cortical receptive field had a different orientation but the same retinotopic location), and if $l > 0.29$ and $r < l\sqrt{2}$ the OD map emerged first.

We investigated by simulation the dependence of the critical values of k on the parameters of the feature space. Fig. 2 shows the dependence k_{or} and k_{od} on l and on r after the retinotopic map has formed. Since on the ordinates k's are represented as increasing, while in simulations k decreases, these plots should be read from up to down: for example on the left figure, for a given value of l, e.g. 0.10, the first map to emerge is the retinotopic one, at $k_{xy} = 0.29$, the second is the OR map, at $k_{\mathrm{or}} = 0.075$, and the last is the OD map, at $k_{\mathrm{od}} = 0.051$.

These simulation values for k_{or} and k_{od} do not match the numerical values predicted by Eq. (5), since the predictions are valid only if OR and OD, respectively, are the first map to emerge. However, surprisingly the analysis still correctly predicts several more general aspects of these graphs. Firstly, in both analysis and simulations k_{or} depends only on r and k_{od} depends only on l. Secondly, in both analysis and simulations this is a linear dependence, though with a change in the slope in the simulations at (OR) or near (OD) the cross-over point. Thirdly, in the simulations the slope of the k_{od} versus

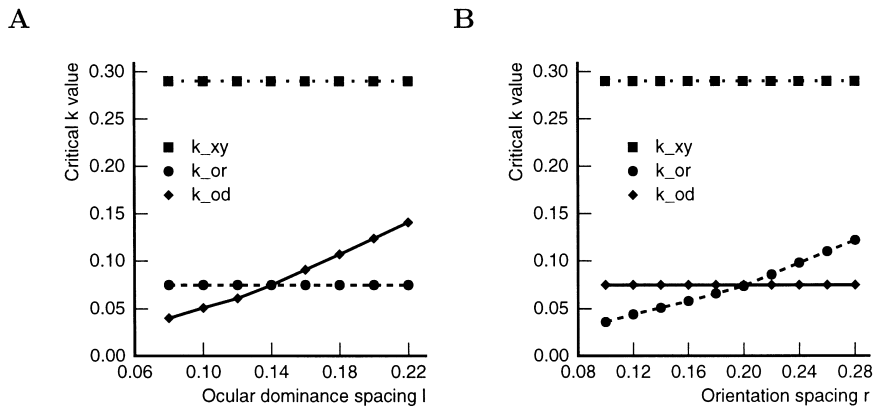

Fig. 2. The dependence of the critical values on feature space parameters: (A) varying OD parameter l with $r = 0.20$; (B) varying OR parameter r with $l = 0.14$.

l line is ≈ 0.5 before cross-over and ≈ 0.8 after cross-over, while the slope of the k_{or} versus r line is ≈ 0.35 before cross-over and ≈ 0.6 after cross-over. Thus, the ratio of the k_{od} and k_{or} slopes both before and after cross-over is $\approx \sqrt{2}$, as in Eq. (5).

5. Discussion

A generalization of the analysis of Durbin et al. [6] makes good predictions for the critical values of the annealing parameter k at which the first map emerges. In addition, there is an intriguing similarity between the functional form of these predictions for k_{or} and k_{od} and the simulation results for the development of these subsequent maps. This suggests that theoretical analysis of the critical points after a retinotopic map has formed may be tractable, even though in principle it appears much more difficult than when all points in the net reside at a point.

The elastic net model of cortical map development can also be run in a non-annealed version, where a constant value of $k < k_{xy}, k_{\mathrm{or}}, k_{\mathrm{od}}$ is used throughout the simulation. In this case, all maps develop simultaneously and the issue of the staging of map development does not arise. Wolf and Geisel [23] have observed that in this region there is the phenomenon of pinwheel annihilation (the maps never stabilize). This does not occur in the annealed version since the maps become "frozen" shortly after their initial formation, and hence we did not observe pinwheel annihilation in the simulations reported here (data not shown). Optical imaging data from ferrets suggests that the emerging maps remain highly stablen [2].

Experimental data has not yet established the precise ordering of OR and OD map development. In monkeys the OD map emerges before birth [14] and the OR map is present very soon after birth [22,1], though it may emerge earlier. In cat, although it

was originally thought that the OD map emerges starting at about 3 weeks after birth [15], more recent optical imaging data suggests that both OR and OD maps are present as early as two weeks after birth [3]. In this paper we have expanded on the work of Hofsümmer et al. [12,13] and analyzed the parameters controlling the order of map development in a slightly different version of the elastic net model; we must now await further experimental data to establish the link between these parameters and the development of real maps in different species.

References

[1] G.G. Blasdel, K. Obermayer, L. Kiorpes, Organization of ocular dominance and orientation columns in the striate cortex of neonatal macaque monkeys, Visual Neurosci. 12 (1995) 589–603.
[2] B. Chapman, M.P. Stryker, T. Bonhoeffer, Development of orientation preference maps in ferret primary visual cortex, J. Neurosci. 16 (1996) 6443–6453.
[3] M.C. Crair, D.C. Gillespie, M.P. Stryker, The role of visual experience in the development of columns in cat visual cortex, Science 279 (1998) 566–570.
[4] P.S. Dayan, Arbitrary elastic topologies and ocular dominance, Neural Comput. 5 (1993) 392–401.
[5] R. Durbin, G. Mitchison, A dimension reduction framework for understanding cortical maps, Nature 343 (1990) 644–647.
[6] R. Durbin, R. Szeliski, A. Yuille, An analysis of the elastic net approach to the traveling salesman problem, Neural Comput. 1 (1989) 348–358.
[7] R. Durbin, D.J. Willshaw, An analogue approach to the travelling salesman problem using an elastic net method, Nature 326 (1987) 689–691.
[8] E. Erwin, K. Obermayer, K. Schulten, Models of orientation and ocular dominance columns in the visual cortex: a critical comparison, Neural Comput. 7 (1995) 425–468.
[9] G.J. Goodhill, K.R. Bates, P.R. Montague, Influences on the global structure of cortical maps, Proc. Roy. Soc. Ser. B 264 (1997) 649–655.
[10] G.J. Goodhill, D.J. Willshaw, Application of the elastic net algorithm to the formation of ocular dominance stripes, Network 1 (1990) 41–59.
[11] G.J. Goodhill, D.J. Willshaw, Elastic net model of ocular dominance: Overall stripe pattern and monocular deprivation, Neural Comput. 6 (1994) 615–621.
[12] F. Hoffsümmer, F. Wolf, T. Geisel, S. Löwel, K. Schmidt, Sequential bifurcation of orientation — and ocular dominance maps, in: ICANN95: Proceedings of the International Conference on Artificial Neural Networks, Vol. 1, Paris, 1995. p. 535
[13] F. Hoffsümmer, F. Wolf, T. Geisel, S. Löwel, K. Schmidt, Sequential bifurcation and dynamic rearrangement of columnar patterns during cortical development, in: Jim Bower (Ed.), Computation and Neural Systems, 1996..
[14] J.C. Horton, D.R. Hocking, An adult-like pattern of ocular dominance columns in striate cortex of newborn monkeys prior to visual experience, J. Neurosci. 16 (1996) 1791–1807.
[15] S. LeVay, M.P. Stryker, C.J. Shatz, Ocular dominance columns and their development in layer IV of the cat's visual cortex: a quantitative study, J. Comp. Neurol. 179 (1978) 223–244.
[16] K. Obermayer, G.G. Blasdel, K. Schulten, Statistical-mechanical analysis of self-organization and pattern formation during the development of visual maps, Phys. Rev. A 45 (1992) 7568–7589.
[17] K. Rose, Deterministic annealing for clustering, compression, classification, regression and related optimization problems, Proc. IEEE 86 (1998) 2210–2239.
[18] P.D. Simic, Statistical mechanics as the underlying theory of elastic and neural optimisations, Network 1 (1990) 89–103.
[19] M.W. Simmen, Parameter sensitivity of the elastic net approach to the travelling salesman problem, Neural Comput. 3 (1991) 363–374.
[20] N.V. Swindale, A model for the formation of orientation columns, Proc. Roy. Soc. London 215 (1982) 211–230.

[21] N.V. Swindale, The development of topography in the visual cortex: a review of models, Network 7 (1996) 161–247.
[22] T.N. Wiesel, D.H. Hubel, Ordered arrangement of orientation columns in monkeys lacking visual experience, J. Comp. Neurol. 158 (1974) 307–318.
[23] F. Wolf, T. Geisel, Spontaneous pinwheel annihilation during visual development, Nature 395 (1998) 73–78.
[24] A.L. Yuille, Generalized deformable models, statistical physics, and matching problems, Neural Comput. 2 (1990) 1–24.
[25] A.L. Yuille, J.A. Kolodny, C.W. Lee, Dimension reduction, generalized deformable models and the development of ocularity and orientation, Neural Networks 9 (1996) 309–319.