Stripe-rearing changes multiple aspects of the structure of primary visual cortex

Nicholas J. Hughes [a], Jonathan J. Hunt [a], Shaun L. Cloherty [b,c], Michael R. Ibbotson [b,c], Frank Sengpiel [d], Geoffrey J. Goodhill [a,e,*]

[a] Queensland Brain Institute, The University of Queensland, St Lucia, Queensland 4072, Australia
[b] National Vision Research Institute, Australian College of Optometry, Carlton, Victoria 3053, Australia
[c] Department of Optometry and Vision Science, University of Melbourne, Parkville, Victoria 3010, Australia
[d] School of Biosciences, Cardiff University, Cardiff, CF10 3AX Wales, UK
[e] School of Mathematics and Physics, The University of Queensland, St Lucia, Queensland 4072, Australia

ARTICLE INFO

ABSTRACT

An important example of brain plasticity is the change in the structure of the orientation map in mammalian primary visual cortex in response to a visual environment consisting of stripes of one orientation. In principle there are many different ways in which the structure of a normal map could change to accommodate increased preference for one orientation. However, until now these changes have been characterised only by the relative sizes of the areas of primary visual cortex representing different orientations. Here we extend to the stripe-reared case a recently proposed Bayesian method for reconstructing orientation maps from intrinsic signal optical imaging data. We first formulated a suitable prior for the stripe-reared case, and developed an efficient method for maximising the marginal likelihood of the model in order to determine the optimal parameters. We then applied this to a set of orientation maps from normal and stripe-reared cats. This analysis revealed that several parameters of overall map structure, specifically the difference between wavelength, scaling and mean of the two vector components of maps, changed in response to stripe-rearing, which together give a more nuanced assessment of the effect of rearing condition on map structure than previous measures. Overall this work expands our understanding of the effects of the environment on brain structure.

© 2014 Elsevier Inc. All rights reserved.

Introduction

A critical question for understanding brain plasticity is to understand how the sensory environment influences brain structure. An important and common model system in this regard is the effect of visual activity during early life on the spatial arrangement of topographic maps in the primary visual cortex of mammals such as cats, ferrets and monkeys (Espinosa and Stryker, 2012). In particular, the spatial arrangement of preference for orientation (Blasdel and Salama, 1986), direction (Weliky et al., 1996), spatial frequency (Hübener et al., 1997; Issa et al., 2000) and ocular dominance (Anderson et al., 1988; Bonhoeffer et al., 1995), as well as the relationships between these maps (Shmuel and Grinvald, 1996; Hübener et al., 1997), has been widely studied.

A particularly striking example of the effect of the environment on these maps is stripe-rearing, in which animals are exposed only to edges of a particular orientation during the critical period (Blakemore and Cooper, 1970; Hirsch and Spinelli, 1970). This leads to an increase in the proportion of primary visual cortical neurons preferring the reared orientation, at the expense of other orientations (Sengpiel et al., 1999). However, so far this proportion is the only quantitative measure that has been used to characterise the difference in structure between normal and stripe-reared orientation maps. This measure is not robust to changes in the subjective spatial filtering applied to maps, which are traditionally obtained by vector averaging noisy data from single-condition optical imaging experiments. Furthermore this measure does not provide a method for determining precisely how the structure of maps changes to accommodate the larger preference for a single orientation which is induced by stripe-rearing.

Recently a novel method was introduced which improves on vector averaging for determining orientation map structure, using a Bayesian approach based on Gaussian processes (Macke et al., 2011). This method uses prior knowledge about map structure in a principled way rather than via subjectively-chosen smoothing parameters, and provides quantitative error estimates for the resulting map. However, this method contained assumptions which are violated by maps from abnormally-reared animals.

We therefore first generalised this Gaussian process method to a much broader class of visual map data, including abnormal rearing

* Corresponding author at: Queensland Brain Institute, The University of Queensland, St Lucia, Queensland 4072, Australia.
E-mail address: g.goodhill@uq.edu.au (G.J. Goodhill).

conditions. We then used it to interrogate map structure in data from the stripe-reared cats of Sengpiel et al. (1999). Besides allowing for more accurate estimates of the maps than in the original study, the method revealed that three parameters of map structure change with stripe-rearing, rather than just the single parameter of the proportion of neurons preferring the reared orientation. These three parameters are the difference between the two vector components in terms of their wavelengths, scalings and means. Together these three parameters provide a much richer description of both natural variability between maps from different individuals, and how maps change in response to altered environments.

Materials and methods

Orientation preference maps

We represent an orientation preference map as a 2D complex field $\mathbf{m}(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^2$ (mathematically equivalent to a 2D vector field), where the preferred orientation and the strength of that preference at each point are given by half the argument and the magnitude, respectively, of the complex field at that point. The argument is halved as orientation is periodic in π rather than 2π radians.

The classical method of estimating orientation preference maps from imaging data is known as vector averaging,

$$\mathbf{m}(\mathbf{x}) = \frac{1}{N} \sum_{j=1}^{N} \left(r_j(\mathbf{x}) \exp\left(2\theta_j i\right) \right), \tag{1}$$

where $r_j(\mathbf{x})$ is the cortical response at location $\mathbf{x} \in \{\mathbf{x}_1, ..., \mathbf{x}_n\} \subset \mathbb{R}^2$ during the jth trial out of N, where a stimulus of orientation θ_j was presented. This map is simply the vectorially-averaged raw response to the stimuli. To account for noise in the result, vector averaged maps are usually then high- and low-pass Gaussian spatially filtered.

Gaussian process regression map estimation

Macke et al. (2011) introduced a novel method for estimating orientation preference maps from optical imaging data by applying Gaussian process (GP) regression (Rasmussen and Williams, 2006) to the problem of estimating a 2D vector field from noisy data. The important findings of the paper are the choice of an appropriate prior, a method for estimating the parameters of the prior, an appropriate noise model and a method for fitting it to the data, and the use of approximation techniques for dealing with large datasets. Here we define the model equivalently but in our own notation for the convenience of the extensions we make below.

Hereafter let $\mathbf{r}_j = (r_j(\mathbf{x}_1), ..., r_j(\mathbf{x}_n))^\top$ refer to the vector of responses at each of the n observed points in the cortex (i.e., the experimental image reshaped as a vector). Let $\mathbf{m}(\mathbf{x})$ be the true orientation preference map underlying the data, and let $\mathbf{m}$ refer to the flattened and concatenated vector of the true map's components, that is,

$$\mathbf{m} = (\mathrm{real}(\mathbf{m}(\mathbf{x}_1)), ..., \mathrm{real}(\mathbf{m}(\mathbf{x}_n)), \mathrm{imag}(\mathbf{m}(\mathbf{x}_1)), ..., \mathrm{imag}(\mathbf{m}(\mathbf{x}_n)))^\top.$$

Then we assume that the data can be written as $\mathbf{r}_j = V_j^\top \mathbf{m} + \epsilon_j$, where $\epsilon_j \sim \mathcal{N}(0, \Sigma_\epsilon)$ and

$$V_j = \left(\cos 2\theta_j, \sin 2\theta_j\right)^\top \otimes \mathbb{I}_n,$$

where $\mathbb{I}_n$ is an $n \times n$ identity matrix and $\otimes$ is the Kronecker product. That is, we observe the component of the underlying map in the θ_j direction corrupted by some correlated Gaussian noise.

A prior defining a distribution of maps is now defined as a Gaussian process, defined by an identically zero mean function and a covariance function K which specifies the covariance between points in the map both within and across the two map vector components. Macke et al. (2011) used a difference of Gaussians covariance function,

$$K(\mathbf{x}, \mathbf{x}', k, k') = \delta_{k,k'} \sum_{a,b=1}^{2} \frac{\alpha_a \alpha_b}{2\pi(\sigma_a^2 + \sigma_b^2)} \exp\left(-\frac{|\mathbf{x}-\mathbf{x}'|^2}{2(\sigma_a^2 + \sigma_b^2)}\right),$$

where $\delta_{i,j}$ is the Kronecker delta, $\mathbf{x}$ and $\mathbf{x}'$ are the two positions in the map, and k and k' are either 1 or 2, the real and imaginary components of the map respectively. The assumption that the components are uncorrelated gives $K(\mathbf{x}, \mathbf{x}', 1, 2) \equiv K(\mathbf{x}, \mathbf{x}', 2, 1) \equiv 0$. Macke et al. (2011) set $\alpha_1 = -\alpha_2$ so that K has zero mean, and $\sigma_2 = 2\sigma_1$ to reduce the number of parameters. These parameters can be interpreted as scaling parameters α_is and filter widths σ_is, which correspond to the wavelength of the map. Macke et al. (2011) showed that this function fit the empirical autocovariance functions of their data well. We found the same result with our data, and additionally found that letting all four parameters (i.e., both α_i and both σ_i) vary independently did not substantially improve the fit to the empirical functions (data not shown). We therefore chose to use the same covariance function as Macke et al. (2011) but with all the multiplying constants (i.e., both α_i, the sum of squared σ_i and 2π) collected into a single scaling parameter coefficient, which we call α, and set $\sigma := \sigma_2 = 2\sigma_1$ for notational simplicity:

$$K(\mathbf{x}, \mathbf{x}', k, k') = \delta(k, k')\alpha^2 \left[\exp\left(\frac{-|\mathbf{x}-\mathbf{x}'|^2}{4\sigma^2}\right) + \frac{1}{4}\exp\left(\frac{-|\mathbf{x}-\mathbf{x}'|^2}{16\sigma^2}\right) - \frac{4}{5}\exp\left(\frac{-|\mathbf{x}-\mathbf{x}'|^2}{10\sigma^2}\right) \right]. \tag{2}$$

To perform Gaussian process regression, the prior distribution defined at the observed points $X = \{\mathbf{x}_1, ..., \mathbf{x}_n\}$ is conditioned on the data to give a posterior distribution, the mean of which is the Bayes optimal estimate of the underling map $\mathbf{m}$. The prior on the set of observed points is $p(\mathbf{m}) = \mathcal{N}(\mathbf{m}; 0, \Sigma_{\mathrm{prior}})$, where the covariance matrix is

$$\sum\nolimits_{\mathrm{prior}} = \begin{pmatrix} K(X, X, 1, 1) & K(X, X, 1, 2) \\ K(X, X, 2, 1) & K(X, X, 2, 2) \end{pmatrix}.$$

This matrix needs to be calculated, stored and inverted, which is a computationally demanding task. To make this tractable, Macke et al. (2011) used a low-rank approximation to the prior, generated by an incomplete Cholesky decomposition. This results in a prior of the form $\sum_{\mathrm{prior}} = D + GG^\top$, where D is $2n \times 2n$ and diagonal, and G is $2n \times q$, with $q \ll 2n$. The eigenspectrum of the chosen covariance function goes to zero sufficiently quickly to make the incomplete Cholesky decomposition a good approximation. To estimate the parameters α and σ of the prior covariance function (Eq. 2), Macke et al. (2011) fit the covariance function to the average of the radial component of the autocorrelation function of the two components of the filtered, vector averaged map.

As per Bayes' rule, the posterior distribution is proportional to the product of the prior and the likelihood,

$$p(\mathbf{m}|\mathbf{r}_1, ..., \mathbf{r}_N) \propto p(\mathbf{m})p(\mathbf{r}_1, ..., \mathbf{r}_N|\mathbf{m}).$$