A single functional model accounts for the distinct properties of suppression in cortical area V1

M.W. Spratling *

King's College London, Department of Informatics and Division of Engineering, London, UK

ARTICLE INFO

ABSTRACT

Cross-orientation suppression and surround suppression have been extensively studied in primary visual cortex (V1). These two forms of suppression have some distinct properties which has led to the suggestion that they are generated by different underlying mechanisms. Furthermore, it has been suggested that mechanisms other than intracortical inhibition may be central to both forms of suppression. A simple computational model (PC/BC), in which intracortical inhibition is fundamental, is shown to simulate the distinct properties of cross-orientation and surround suppression. The same model has previously been shown to account for a large range of V1 response properties including orientation-tuning, spatial and temporal frequency tuning, facilitation and inhibition by flankers and textured surrounds as well as a range of other experimental results on cross-orientation suppression and surround suppression. The current results thus provide additional support for the PC/BC model of V1 and for the proposal that the diverse range of response properties observed in V1 neurons have a single computational explanation. Furthermore, these results demonstrate that current neurophysiological evidence is insufficient to discount intracortical inhibition as a central mechanism underlying both forms of suppression.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Inhibitory mechanisms intrinsic to primary visual cortex (V1) were originally believed to be responsible for both cross-orientation suppression (Bonds, 1989; DeAngelis, Robson, Ohzawa, & Freeman, 1992; Morrone, Burr, & Maffei, 1982) and surround suppression (DeAngelis, Freeman, & Ohzawa, 1994; Fitzpatrick, 2000; Knierim & van Essen, 1992). Subsequent work has demonstrated that these two forms of suppression exhibit a number of distinct properties, and hence, may result from different underlying mechanisms. A range of alternative mechanisms have been proposed for each type of suppression. Specifically, it has been suggested that cross-orientation suppression might arise from attenuation of the feedforward input due to depression of the thalmacortical synapses (Carandini, Heeger, & Senn, 2002; Freeman, Durand, Kiper, & Carandini, 2002) or a reduction in feedforward drive to cortical cells caused by contrast saturation in lateral geniculate nucleus (LGN) cells (Li, Thompson, Duong, Peterson, & Freeman, 2006; Priebe & Ferster, 2006). It has also been suggested that surround suppression might be mediated by inhibitory mechanisms intrinsic to V1 but driven by feedback from extrastriate cortex (Angelucci

et al., 2002; Bair, Cavanaugh, & Movshon, 2003; Cavanaugh, Bair, & Movshon, 2002; Sullivan & de Sa, 2006) or might be due to surround suppression in LGN (Ozeki et al., 2004; Naito, Sadakane, Okamoto, & Sato, 2007; Webb, Dhruv, Solomon, Tailby, & Lennie, 2005).

A previous publication (Spratling, 2010) has demonstrated that a simple functional model (PC/BC), derived from the predictive coding and biased-competition theories of cortical function, can simulate a very wide range of V1 response properties including cross-orientation and surround suppression. This article extends that work by showing that the PC/BC model of V1 can also simulate the distinct behaviours exhibited by these two forms of suppression. The PC/BC model includes two mechanisms that can give rise to suppression: a mechanism of intracortical inhibition employing divisive normalisation of the inputs to a population of competing neurons; and saturation of the LGN responses to high contrast stimuli. The latter mechanism was proposed by Priebe and Ferster (2006) and Li et al. (2006) to account for cross-orientation suppression. It is found that in the PC/BC model surround suppression is generated by the mechanism of cortical inhibition, while cross-orientation suppression is generated by a combination of cortical inhibition and LGN response saturation. Hence, the PC/BC model predicts that intracortical inhibition is essential for both forms of suppression, contrary to suggestions that completely separate mechanisms are required and to claims that cortical inhibition is not involved.

* Address: Division of Engineering, King's College London, Strand, London WC2R 2LS, UK.
   E-mail address: michael.spratling@kcl.ac.uk

2. Methods

2.1. The LGN model

The input to the PC/BC model of V1, described below, was an input image (I) pre-processed by convolution with a Laplacian-of-Gaussian (LoG) filter (l) with standard deviation equal to one. This is virtually identical to the Difference-of-Gaussians (DoG) filter that has traditionally been used to model circular receptive fields (RFs) in LGN and retina. The output from this filter was subject to a multiplicative gain (the strength of which was determined by parameter κ) followed by a saturating non-linearity, such that:

$$\mathbf{X} = \tanh\{\kappa(I * l)\} \tag{1}$$

The positive and rectified negative responses were separated into two images $\mathbf{X}_{ON}$ and $\mathbf{X}_{OFF}$ simulating the outputs of cells in retina and LGN with circular-symmetric on-centre/off-surround and off-centre/on-surround RFs respectively. This pre-processing is illustrated on the left of Fig. 1. Consistent with neurophysiological data (Reid & Alonso, 1995), the ON-centre model LGN neurons indirectly provided input to the ON sub-field of the model V1 simple cells, while the OFF-centre model LGN neurons indirectly provided input to the OFF sub-field of the model V1 neurons (see next section).

To explore the effects of the different mechanisms of suppression, in some experiments the suppression of the LGN responses was turned off. In this case Eq. (1) was replaced by:

$$\mathbf{X} = \kappa(I * l) \tag{2}$$

A value of $\kappa = 10$ was used in all experiments reported here.

2.2. The V1 model

The PC/BC model of V1 is illustrated in Fig. 1 and described by the following equations:

$$\mathbf{E}_o = \mathbf{X}_o \oslash \left(\epsilon_2 + \sum_{k=1}^{p} (\hat{w}_{ok} * \mathbf{Y}_k) \right) \tag{3}$$

$$\mathbf{Y}_k \leftarrow (\epsilon_1 + \mathbf{Y}_k) \otimes \sum_{o} (w_{ok} \star \mathbf{E}_o) \tag{4}$$

where $o \in [ON, OFF]$, $\mathbf{X}_o$ is a two-dimensional array, equal in size to the input image, that represents the input to the model of V1, $\mathbf{E}_o$ is a two-dimensional array, equal in size to the input image, that represents the error-detecting neuron responses, and $\mathbf{Y}_k$ is a two-dimensional array, equal in size to the input image, that represent the

prediction neuron responses, w_{ok} is a two-dimensional kernel representing the synaptic weights for a particular class (k) of neuron normalised so that sum of all the weights was equal to ψ, $\hat{w}_{ok}$ is a two-dimensional kernel representing the same synaptic weights as w_{ok} but normalised so that the maximum value was equal to ψ, p is the total number of kernels, ϵ_1, ϵ_2, and ψ are parameters, $\oslash$ and $\otimes$ indicate element-wise division and multiplication respectively, $\star$ represents cross-correlation (which is equivalent to convolution without the kernel being rotated 180°), and $*$ represents convolution (which is equivalent to cross-correlation with a kernel rotated by 180°). Parameter values $\psi = 5000$, $\epsilon_1 = 0.0001$ and $\epsilon_2 = 250$ were used in the simulations reported in this article.

Eq. (4) describes the updating of the prediction neuron activations. The response of each prediction neuron is a function of its activation at the previous iteration and a weighted sum of afferent inputs from the error-detecting neurons. Eq. (3) describes the calculation of the neural activity for each population of error-detecting neurons. These values are a function of the activity of the input to V1 divisively modulated by a weighted sum of the outputs of the prediction neurons in V1. The activation of the error-detecting neurons can be interpreted in two ways. Firstly, $\mathbf{E}$ can be considered to represent the residual error between the input and the reconstruction of the input generated by the prediction neurons. The values of $\mathbf{E}$ indicate the degree of mismatch between the top-down reconstruction of the input and the actual input (assuming ϵ_2 is sufficiently small to be negligible). When a value within $\mathbf{E}$ is greater than $\frac{1}{\psi}$ it indicates that a particular element of the input is under-represented in the reconstruction, a value of less than $\frac{1}{\psi}$ indicates that a particular element of the input is over-represented in the reconstruction, and a value of $\frac{1}{\psi}$ indicates that the top-down reconstruction perfectly predicts the bottom-up stimulation. A second interpretation is that $\mathbf{E}$ represents the inhibited inputs to a population of competing prediction neurons. Each prediction neuron modulates its own inputs, which helps stabilise the response of the prediction neurons, since a strongly (or weakly) active prediction neuron will suppress (magnify) its inputs, and hence, reduce (enhance) its own response. Prediction neurons that share inputs (i.e., that have overlapping RFs) will also modulate each other's inputs. This generates a form of competition between the prediction neurons, such that each neuron effectively tries to block other prediction neurons from responding to the inputs which it represents. This mechanism of competition is called Divisive Input Modulation (DIM) (Spratling, De Meyer, & Kompass, 2009).

The RF of a simple cell in primary visual cortex can be accurately modelled by a two-dimensional Gabor function (Daugman,
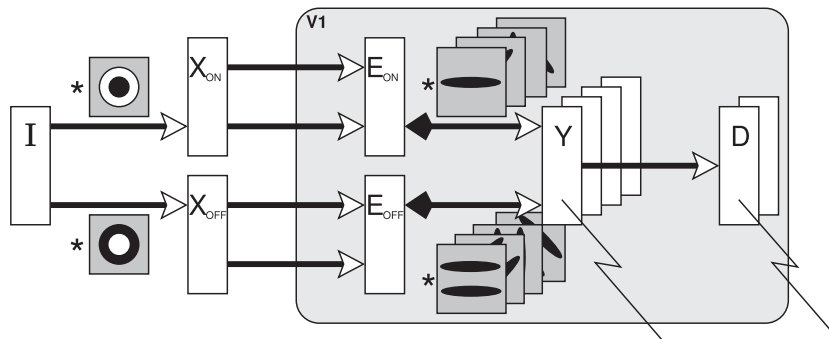

Fig. 1. The PC/BC model of V1. The input image *I* is preprocessed by convolution with a circular-symmetric on-centre/off-surround kernel (to generate the input to the ON channel of the V1 model), and a circular-symmetric off-centre/on-surround kernel (to generate the input to the OFF channel of the V1 model). The prediction neurons (labelled **Y**) represent V1 simple cells. Responses from these neurons were recorded during most experiments. These responses were generated by convolving the outputs of the (ON and OFF channels of the) error-detecting neurons (labelled **E**) with (the ON and OFF channels of) a number of kernels representing V1 RFs. This convolution process effectively reproduces the same RFs at every pixel location in the image. The responses of the error-detecting neurons are influenced by divisive feedback from the prediction neurons, which is also calculated by convolving the prediction neuron outputs with the weight kernels. Complex cell responses (labelled *D*), which were also recorded during some experiments, were calculated by taking the maximum response of a small population of simple cells representing a single orientation at a particular spatial location.