

Formation of Direction Selectivity in Natural Scene Environments

Brian Blais, Leon N Cooper and Harel Shouval

The Departments of Physics and Neuroscience

The Institute for Brain and Neural Systems

Box 1843, Brown University

Providence, R. I., 02912

`bblais@cns.brown.edu, lnc@brown.edu, hzs@cns.brown.edu`

June 24, 1999

Abstract

Most simple and complex cells in the cat striate cortex are both orientation and direction selective. In this paper we use single cell learning rules to develop both orientation and direction selectivity in a natural scene environment. We show that a simple PCA rule is inadequate for developing direction selectivity, but that the BCM rule as well as similar higher order rules can. We also demonstrate that the convergence of lagged and non-lagged cells depends on the velocity of motion in the environment, and that strobe rearing disrupts this convergence resulting in a loss of direction selectivity.

^oThis work supported in part by the Charles A. Dana Foundation and the Office of Naval Research

1 Introduction

Most simple and complex cells in the cat striate cortex are both orientation(Hubel and Wiesel, 1959; Hubel and Wiesel, 1962) and direction selective(Hammond, 1978; Reid et al., 1991; Deangelis et al., 1995). At the preferred orientation, a cell which is direction selective responds to a drifting grating moving in one direction more strongly than the opposite direction. The ability of the cell to detect the direction of motion depends on the interaction of responses to at least two different points in the visual field at different times. This is to say, that it depends on the *spatiotemporal* receptive field of the cell(Reid et al., 1991).

A cell which is not direction selective (non-DS) has a maximum response to a sine grating moving in one direction equal to its maximum response to a sine grating moving in the opposite direction. In a linear approximation, the response of a cell, $c(t)$, can be written as a convolution between the spatio-temporal input pattern, $I(x, t)$, and a spatio-temporal receptive field kernel, $K(x, t)$, giving

$$c(t) = \int_{-\infty}^{+\infty} dx' \int_{-\infty}^t dt' I(x', t') K(x', t - t') \quad (1.1)$$

It is easy to show that, in this approximation, a direction selective cell must have a ST inseparable receptive field, that is, the kernel cannot be expressed as $K(x, t) = F(t)G(x)$, where $F(t)$ and $G(x)$ are functions which depend only on time and only on space, respectively.

There are many models of direction selectivity(Barlow and Levick, 1965; Adelson and Bergen, 1985; Watson and Ahumada, 1985; Burr, 1981). In all of the models, the response of the cell is determined by receptive fields that have different temporal response properties at different spatial locations (i.e. spatiotemporal inseparable). This can be realized by the appropriate spatial positioning of the receptive fields, and the introduction of temporal shifts. These temporal shifts could possibly arise from delays caused by cortical loops(Suarez et al., 1995; Maex and Orban, 1996), phase advances caused by depressing synapses(Chance et al., 1998), or by lagged responses in the LGN(Mastrorarde, 1987; Saul and Humphrey, 1990).

In this paper we present a feed-forward model of the development of direction selectivity which includes the effects of two types of LGN cells, called lagged and non-lagged cells, which differ only in their response timing. This is similar to a previous model (Feidler et al., 1997), differing significantly in the input environment used. The previous model used a neuron with three inputs each with sinusoidally varying activations governed by a single parameter. In our case, we use a natural scene environment (Law and Cooper, 1994), providing a more realistic correspondence with biology and a more direct connection to experiment. We examine two types of visual environments, differing primarily in their temporal properties. In one environment, static natural images are used with a simple model of eye movements to provide motion over the receptive cortical field. In the other environment, natural image movies are used to provide motion over the receptive field.

Using these more realistic environments, we reproduce the observation of Feidler et. al. that simple Hebb rules are incapable of producing direction selective cells, and that the BCM rule (Bienenstock et al., 1982) is capable of producing direction selective cells. We also consider several other statistically motivated rules (Blais et al., 1998), which have similar form to the BCM rule, and explore how they develop direction selectivity. In addition we provide an insight into the reason direction selectivity is, or is not, produced by these rules. Finally, we compare these simulations to experiments in strobe light environments (Cynader et al., 1973; Cynader and Chernenko, 1976; Humphrey and Saul, 1998).

2 Methods

We use as the visual environment, 13×13 circular patches from 12 images of natural scenes processed with a retinal DOG filter (Law and Cooper, 1994). The cortical cell receives input from two sets of LGN cells which view the same area of space, but differ only in their timing. The first set has a delayed response to the input (lagged cells) relative to the second set (non-lagged cells). Essentially the cortical cell has a receptive field that is composed of two different

RFs that receive input from two different times at the same spatial location. We will refer to these as the lagged and the non-lagged RFs.

There are two possible factors contributing to motion in the visual environment: movements of the eyes and head and movements of objects in the world. To model the former, input patches are chosen using a sequence of random *saccades* and *drifts*(Carpenter, 1977). A saccade is a large jump to a random part of an image, and a drift is a continuous motion within an image in a particular direction at a particular velocity. In the model, the drift velocity is kept constant, and the drifts last a random time. In between drifts are saccades to a different image or part of the same image. Although this is a simplification of both the temporal properties of lagged and non-lagged cells, and of the true input structure available to an animal in a dynamic environment composed of moving objects in addition to eye and head movements, the added complexities make no noticeable difference in the results (see Figure 3). In contrast, the simplification of the environment provides insights which may be obscured in the complexities, and it allows the velocities and delays to be more tightly controlled.

We denote the input vector by \mathbf{d} and the weight vector by \mathbf{m} . Neural activity is given by the rectified product of the inputs and the weights, $c = \sigma(\mathbf{d} \cdot \mathbf{m})$. The derivative of the sigmoidal is given simply by σ' . We consider several synaptic modification rules such as the quadratic form of BCM(Bienenstock et al., 1982; Intrator and Cooper, 1992) as well as other statistically motivated rules that share the basic properties proposed by BCM(Blais et al., 1998). We include, for comparison, a stabilized Hebb rule used for extracting the principle component of the input (PCA) (Oja, 1982). These rules have the form:

Quadratic BCM (Intrator and Cooper, 1992)

$$\frac{d\mathbf{m}}{dt} = c(c - \Theta_M)\sigma'\mathbf{d} = c(c - E[c^2])\sigma'\mathbf{d} \quad (2.2)$$

Skewness This rule is based on the statistical measure of skewness.

$$\frac{d\mathbf{m}}{dt} = c(c - E[c^3]/E[c^2])/E^{1.5}[c^2]\sigma' \mathbf{d} \quad (2.3)$$

Kurtosis This rule is based on the statistical measure of kurtosis.

$$\frac{d\mathbf{m}}{dt} = c(c^2 - E[c^4]/E[c^2])/E^2[c^2]\sigma' \mathbf{d} \quad (2.4)$$

For all of these rules, we replace the spatial average, $E[\cdot]$, by a temporal average of the form

$$E[c^n(t)] \approx \frac{1}{\tau} \int_{-\infty}^t c^n(t') e^{-(t-t')/\tau} dt'$$

PCA

$$\frac{d\mathbf{m}}{dt} = c(\mathbf{d} - c\mathbf{m}) \quad (2.5)$$

We measure direction selectivity using the DS index, defined as

$$DS \equiv \frac{R_{(\text{preferred})} - R_{(\text{non-preferred})}}{R_{(\text{preferred})} + R_{(\text{non-preferred})}} \quad (2.6)$$

where $R_{(\text{preferred})}$ and $R_{(\text{non-preferred})}$ are the responses to a sine grating, at optimum orientation and spatial frequency, moving in the preferred direction and non-preferred direction, respectively.

3 Results

Example receptive fields and their orientation tuning, for a drift velocity of 2 pixels per iteration, are shown in Figure 1. The orientation tuning was obtained using drifting oriented sine gratings. It is clear that the PCA rule did not develop direction selectivity whereas the other rules did. A spatiotemporal *separable* receptive field would be attained if the lagged RF is a simple scalar multiple of the non-lagged RF, such as that observed to occur using the PCA learning rule.

The other learning rules developed different lagged and non-lagged RFs. We observe that the lagged RF is a shifted version of the non-lagged RF, which yields a spatiotemporal *inseparable* receptive field. Figure 2 shows the direction selectivity index as a function of eye drift velocity, for a constant LGN lag of 1 iteration. Example receptive fields from the BCM learning rule for each velocity are shown. Other than the PCA rule, all rules show some tuning to eye drift velocity: they develop direction selectivity for some velocities, but all lose it for either velocities which are too high or too low. At very low velocities, the lagged and non-lagged RFs are identical (ST separable), and at very high velocities only the lagged or the non-lagged develops while the other RF is small and random (also ST separable).

Using a more dynamic environment (movies from (van Hateren and Ruderman, 1998)) and a more realistic temporal filtering (Saul and Humphrey, 1990; Wimbauer et al., 1997a) doesn't alter the main result (see Figure 3). The neurons develop a spatiotemporal *inseparable* receptive field, and are thus direction selective.

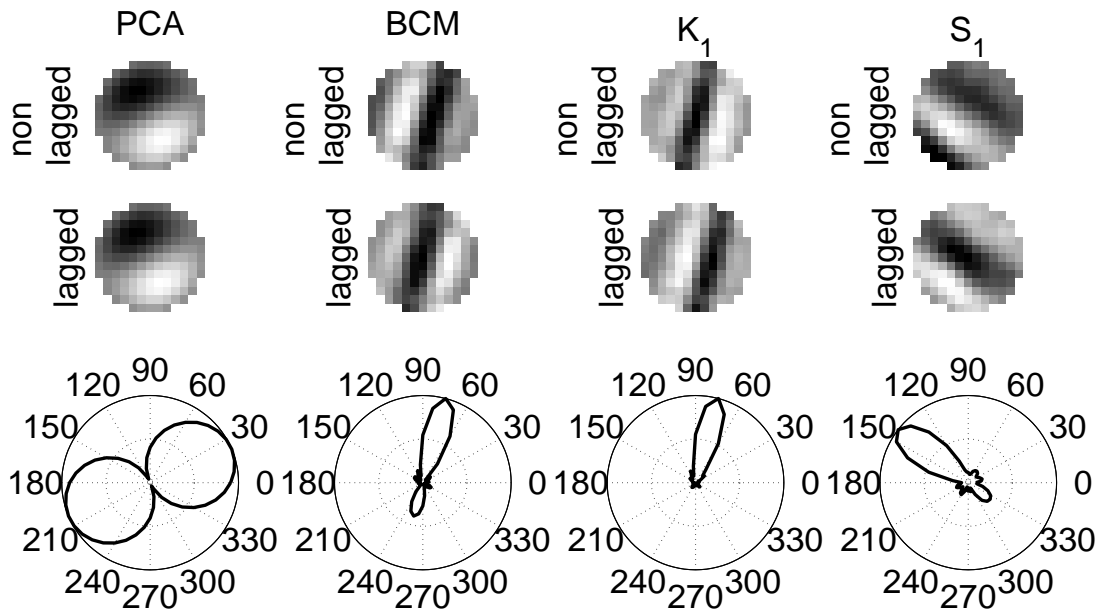


Figure 1: Sample receptive fields and their orientation tuning, for a velocity of 2 pixels per iteration. The orientation tuning was obtained using drifting oriented sine gratings. Orientations larger than 180 degrees denotes motion in the opposite direction. Tuning curves which have a larger response for one direction than another are for direction selective cells. The PCA neuron is the only one which did not achieve direction selectivity.

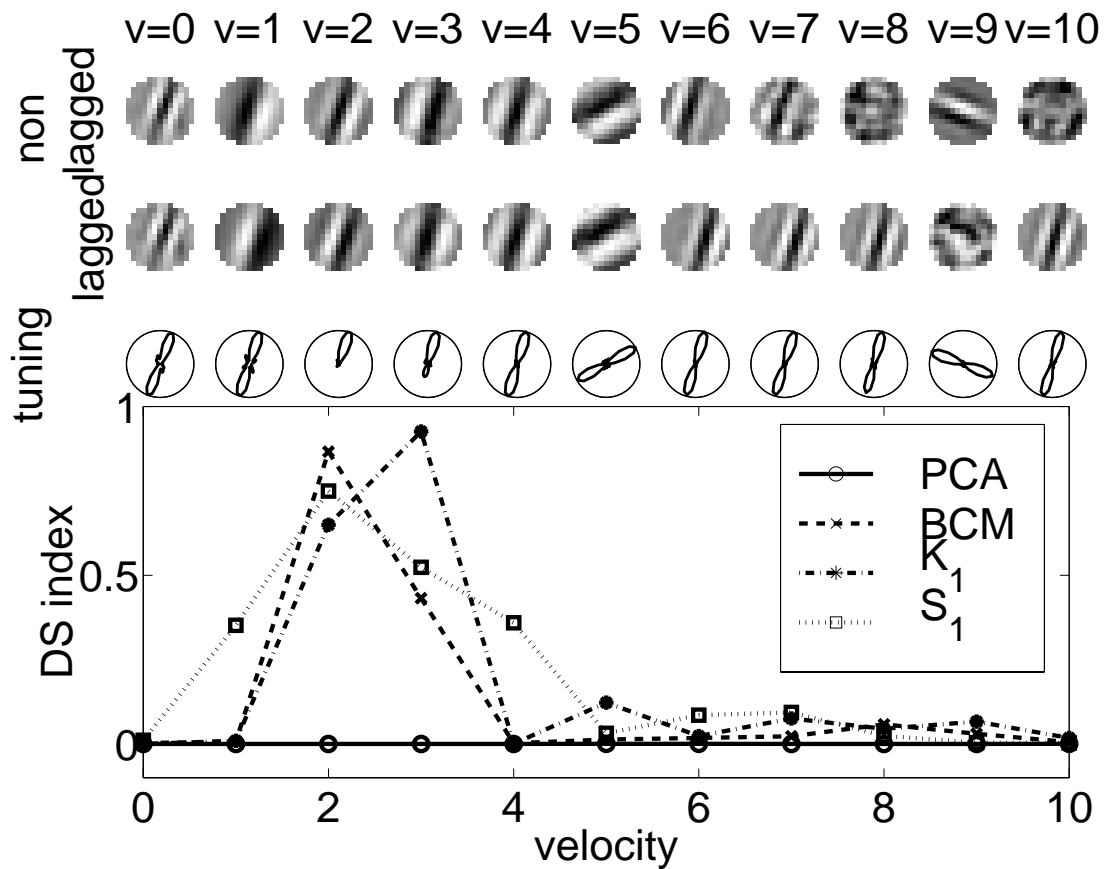


Figure 2: Sample receptive fields with polar tuning plots (above) for BCM, for several eye drift velocities. The direction selectivity index as a function of drift velocity (below) for four different learning rules. The PCA learning rule did not develop direction selectivity. The other rules show some tuning to eye drift velocity: they all lose direction selectivity for either velocities which are either too high (no overlap) or too low (near complete overlap). The LGN lagged cells had a constant 1 iteration lag.

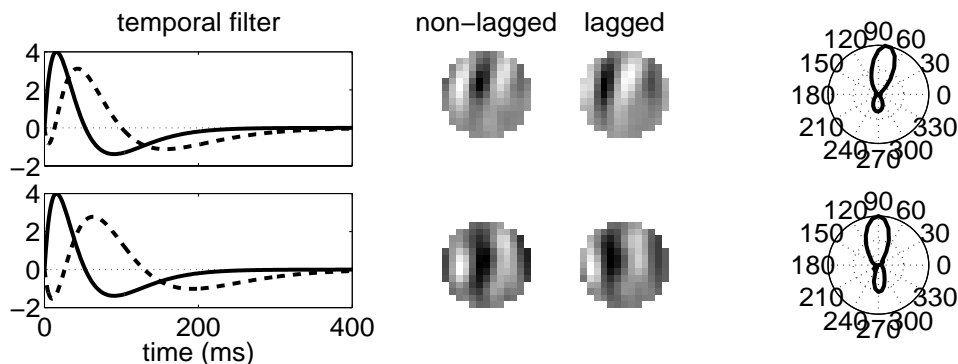


Figure 3: Direction selectivity in a movie environment. Neurons are trained with sequential patches from movies preprocessed with a spatial DOG and temporal filters for lagged and non-lagged cells, as specified in (Wimbauer et. al. 1997a). The BCM rule was used in these examples. Shown are the temporal filters (left), the receptive fields for the lagged and non-lagged channels (middle), and the orientation tuning of the cells. The results are comparable to those from the drift-saccade input model.

4 Summary and Conclusions

This work extends previous work(Feidler et al., 1997), in which a cell attained direction selectivity in a simplified environment, and their conclusion, that a simple Hebb rule is inadequate to develop direction selectivity, is reproduced. Other than the PCA rule, all the rules examined develop direction selectivity for some velocities, but all lose it for either velocities which are either too high or too low. There is a parallel between these results and the results on binocular cortical misalignment(Shouval et al., 1996), which we explain presently.

In the misalignment work, it was shown that the BCM rule in a natural scene environment, with varying degrees of binocular overlap, developed either identical receptive fields (for complete overlap), *monocular* receptive fields (for no overlap), or receptive fields formed only in the overlap region (for intermediate overlap). In the current work, if the lag time of the LGN lagged cells is kept constant, then a constant velocity would imply a constant amount of overlap of input patterns during eye drift. Thus, zero velocity would yield identical lagged and the non-lagged

RFs, yielding no direction selectivity. Likewise, high velocity would give no overlap of input patterns during eye drift, and would yield either a completely lagged or a completely non-lagged receptive field, and again no direction selectivity. The high velocity case is analogous to the strobe light environment (Cynader et al., 1973; Cynader and Chernenko, 1976; Humphrey and Saul, 1998), because in both situations the temporal correlations are lost. In the experiment, direction selectivity was lost but orientation selectivity remained, which is reproduced by the simulations for the high velocity case. In addition, there is evidence (Humphrey et al., 1998) that strobe rearing prevents the convergence of the lagged and non-lagged inputs onto the cortical cell, i.e. the cortical receptive field is affected by either lagged or non-lagged, but not both. This result is reproduced completely in the simulations.

It is straightforward to show that a PCA neuron in a spatially isotropic input environment develops either two identical receptive fields or two receptive fields differing only by a sign.¹ These two cases are, in the context of lagged and non-lagged receptive fields, ST separable and *not* direction selective. A recent model of direction selectivity used a correlation based learning paradigm, with lagged and non-lagged LGN responses (Wimbauer et al., 1997a; Wimbauer et al., 1997b). The results were that a) direction selectivity could not form unless the lagged and non-lagged RFs developed independently (i.e. very weak correlations), or b) if the environment contains anisotropic motion (i.e. objects move in some directions more often than others). Despite the added complication of network effects and continuous response properties, these results can be understood from the results presented here using the single cell PCA neuron. The BCM rule, as well as higher order rules as we have seen, require neither of these assumptions to develop direction selectivity.

Recent comparisons between ICA or efficient coding of natural image sequences and the ST receptive fields of visual cortical neurons has been performed (van Hateren and Ruderman,

¹Simply stated, if the two-channel correlation function is symmetric, and has the form $\mathbf{C} = \begin{pmatrix} \mathbf{C}_1 & \mathbf{C}_2 \\ \mathbf{C}_2 & \mathbf{C}_1 \end{pmatrix}$, then the two channels of the principal components, $\mathbf{v} = \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix}$, are identical up to a sign, $\mathbf{v}_1 = \pm \mathbf{v}_2$.

1998; Rao and Ballard, 1997). The assumptions used there were different than the ones used in the current work, however some of the results are similar. The primary differences are that many more temporal delays are included in the ICA work and the environment is required to be “whitened”. There are two possible interpretations to that work. One is that there are many different type of LGN neurons each with a different time lag, and each converging through independent synaptic weights on a single cortical cell. This interpretation is similar to our work differing in our use of only two populations of LGN cells, as indicated experimentally, and in that we have investigated several different learning rules. Another interpretation is that there is one population of LGN cells but that synaptic weights have independent values at different time points. This interpretation is very different than our suggestion, and is unlikely biologically.

In addition, the use of “whitening” can be problematic. Although LGN responses could be approximately decorrelated (Atick and Redlich, 1992; Dan et al., 1996), some of the rules used for ICA are particularly sensitive to small correlations remaining in nearly whitened environments (Blais et al., 1998). Also, for the same interpretations described above, the spatio-temporal whitening performed could be questionable biologically. The learning rules in the current work do not require whitening, nor are they affected adversely by it, and thus provide a more robust system for achieving direction selectivity. The interpretation of these other works and the comparison to biological quantities is, therefore, significantly different. The fact that they achieve direction selectivity in the natural scene environment is the same as the current work.

Although we have focussed on a geniculocortical single-cell model, there is some indication that network effects, especially cortical-cortical interaction, play a part in direction selectivity (Livingstone, 1998). This work can be seen as a first step towards a network model, and as a straightforward way of addressing some of the issues regarding direction selectivity.

References

Adelson, E. H. and Bergen, J. R. (1985). Spatiotemporal energy models for the perception of

- motion. *Science*, 275:220–222.
- Atick, J. J. and Redlich, A. (1992). What does the retina know about natural scenes. *Neural Computation*, 4:196–211.
- Barlow, H. and Levick, R. (1965). The mechanism of direction selectivity in the rabbit’s retina. *Journal of Physiology*, 173:477–504.
- Bienenstock, E. L., Cooper, L. N., and Munro, P. W. (1982). Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience*, 2:32–48.
- Blais, B. S., Intrator, N., Shouval, H., and Cooper, L. N. (1998). Receptive field formation in natural scene environments: comparison of single cell learning rules. *Neural Computation*, 10(7).
- Burr, D. (1981). A dynamic model for image registration. *Computer Graphics and Image Processing*, 15:102–112.
- Carpenter, R. (1977). *Movements of the Eyes*. Pion-London.
- Chance, F. S., Nelson, S. B., and Abbott, L. F. (1998). Synaptic depression and the temporal response characteristics of V1 cells. *Journal of Neuroscience*, 18:4785–4799.
- Cynader, M., Berman, N., and Hein, A. (1973). Cats reared in stroboscopic illumination: effects on receptive fields in visual cortex. *Proceedings of the National Academy of Sciences*, 70:1353–1354.
- Cynader, M. and Chernenko, G. (1976). Abolition of direction selectivity in the visual cortex of the cat. *Science*, 193:504–505.
- Dan, Y., Atick, J. J., and Reid, R. C. (1996). Efficient coding of natural scenes in lateral geniculate nucleus: Experimental test of a computational theory. *J. Neurosci.*, 16(1):3351–3362.

- Deangelis, G. C., Ohzawa, I., and Freeman, R. C. (1995). Receptive field dynamics in the central visual pathway. *Trends in Neuroscience*, 18:451:458.
- Feidler, J. C., Saul, A. B., Murthy, A., and Humphrey, A. L. (1997). Hebbian learning and the development of direction selectivity: the role of geniculate response timings. *Network: Computational Neural Systems*, 8:195–214.
- Hammond, P. (1978). Directional tuning of complex cells in area 17 of the feline visual cortex. *J. Physiol. (Lond)*, 275:479–491.
- Hubel, D. H. and Wiesel, T. N. (1959). Receptive fields of single neurons in the cat striate cortex. *J. Physiol. (London)*, 148:509–591.
- Hubel, D. H. and Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *J. Physiol*, 160:106–154.
- Humphrey, A. L. and Saul, A. B. (1998). Strobe rearing reduces direction selectivity in area 17 by altering spatiotemporal receptive-field structure. *Journal of Neurophysiology*, 80:2991–3004.
- Humphrey, A. L., Saul, A. B., and Feidler, J. C. (1998). Strobe rearing prevents the convergence of inputs with different response timings onto area 17 simple cells. *Journal of Neurophysiology*, 80:3005–3020.
- Intrator, N. and Cooper, L. N. (1992). Objective function formulation of the BCM theory of visual cortical plasticity: Statistical connections, stability conditions. *Neural Networks*, 5:3–17.
- Law, C. and Cooper, L. (1994). Formation of receptive fields according to the BCM theory in realistic visual environments. *Proceedings National Academy of Sciences*, 91:7797–7801.
- Livingstone, M. S. (1998). Mechanisms of direction selectivity in V1. *Neuron*, 20:509–526.
- Maex, R. and Orban, G. A. (1996). Model circuit of spiking neurons generating directional selectivity in simple cells. *Journal of Neurophysiology*, 75:1515–1545.

- Mastronarde, D. N. (1987). Two classes of single input X-cells in cat lateral geniculate nucleus. Cat lateral geniculate nucleus. I. Receptive field properties and classification of cells. *Journal of Neurophysiology*, 57:357–380.
- Oja, E. (1982). A simplified neuron model as a principal component analyzer. *Journal of Mathematical Biology*, 15:267–273.
- Rao, R. P. N. and Ballard, D. H. (1997). Efficient encoding of natural time varying images produces oriented space-time receptive fields. Technical report, Department of Computer Science, University of Rochester.
- Reid, R. C., Soodak, R. E., and Shapley, R. M. (1991). Directional selectivity and spatiotemporal structure of receptive fields of simple cells in cat striate cortex. *Journal of Neurophysiology*, 66(2).
- Saul, A. B. and Humphrey, A. L. (1990). Spatial and temporal properties of lagged and nonlagged cells in the cat lateral geniculate nucleus. *Journal of Neurophysiology*, 68:1190–1208.
- Shouval, H., Intrator, N., Law, C. C., and N Cooper, L. (1996). Effect of binocular cortical misalignment on ocular dominance and orientation selectivity. *Neural Computation*, 8(5):1021–1040.
- Suarez, H., Koch, C., and Douglas, R. (1995). Modeling direction selectivity of simple cells in striate visual cortex within the framework of the canonical microcircuit. *Journal of Neuroscience*, 15:6700–6719.
- van Hateren, J. and Ruderman, D. L. (1998). Independent component analysis of natural image sequences yields spatiotemporal filters similar to simple cells in primary visual cortex. *Proc.R.Soc.Lond. B*, 265:2315–2320.
- Watson, A. B. and Ahumada, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optics Society of America*, A2:322–342.

Wimbauer, S., Wensch, O. G., Miller, K. D., and von Hemmen, J. L. (1997a). Development of spatiotemporal receptive fields of simple cells: I. Model formulation. *Biological Cybernetics*, 77:453–461.

Wimbauer, S., Wensch, O. G., Miller, K. D., and von Hemmen, J. L. (1997b). Development of spatiotemporal receptive fields of simple cells: II. simulation and analysis. *Biological Cybernetics*, 77:463–477.