

Slow feature analysis yields a rich repertoire of complex cell properties

Pietro Berkes and Laurenz Wiskott

Institute for Theoretical Biology, Humboldt University Berlin
Invalidenstraße 43, D-10115 Berlin, Germany
{p.berkes,l.wiskott}@biologie.hu-berlin.de

Abstract

In this study we investigate temporal slowness as a learning principle for receptive fields using slow feature analysis on image sequences. We analyze the results with physiological methods and compare them to neurons reported in the experimental literature. We find that the learned functions have many properties found also experimentally in complex cells of primary visual cortex, such as direction selectivity, non-orthogonal inhibition, end-inhibition, and side-inhibition. Our results demonstrate that a single unsupervised learning principle can account for such a rich repertoire of receptive field properties.

1 Introduction

A possible approach to the study of the organization of the visual cortex is to assume that it is philo- or ontogenetically adapted to the statistics of its input in order to satisfy one (or possibly more) computational objectives (e.g. Field, 1994). In this work we investigate *temporal slowness* as a possible computational principle for the emergence of complex cell receptive fields. The slowness principle (see Wiskott and Sejnowski, 2002, and references therein) is based on the observation that the environment, sensory signals, and internal representations of the environment vary on different time scales. The environment changes usually on a relatively slow time scale. Sensory signals on the other hand, such as the responses of single receptors in the retina, vary on a faster time scale, because even a small eye movement or shift of a textured object may lead to a rapid variation of the light intensity received by a receptor neuron. The internal representation of the environment, finally, should vary on a time scale similar to that of the environment itself, i.e. on a slow time scale. Our working hypothesis in this paper is that the cortex organizes to extract slowly varying features from the quickly varying sensory signal in order to build a consistent internal representation of the environment. The simulation results presented here have previously been presented in (Berkes and Wiskott, 2003; Wiskott and Berkes, 2003). This paper provides a complementary analysis carried out with physiological methods and makes comparisons with neurons of primary visual cortex (V1) reported in the experimental literature.

2 Slow feature analysis

In mathematical terms the learning task we want to solve can be stated as follows (Wiskott and Sejnowski, 2002): given a multidimensional input signal $\mathbf{x}(t) = (x_1(t), \dots, x_N(t))^T$, $t \in [t_0, t_1]$, find a set of real-valued functions $g_1(\mathbf{x}), \dots, g_M(\mathbf{x})$ lying in a given function space \mathcal{F} so that for the output signals $y_j(t) := g_j(\mathbf{x}(t))$

$$\Delta(y_j) := \langle \dot{y}_j^2 \rangle_t \text{ is minimal} \quad (1)$$

$$\text{under the constraints } \langle y_j \rangle_t = 0 \text{ (zero mean),} \quad (2)$$

$$\langle y_j^2 \rangle_t = 1 \text{ (unit variance),} \quad (3)$$

$$\forall i < j, \langle y_i y_j \rangle_t = 0 \text{ (decorrelation and order),} \quad (4)$$

with $\langle \cdot \rangle_t$ and \dot{y} indicating time averaging and the time derivative of y , respectively. The function space \mathcal{F} is chosen here to be the set of all polynomials of degree 2. We refer to the j -th polynomial g_j as the j -th unit.

We solve this optimization problem with slow feature analysis (SFA) (Wiskott and Sejnowski, 2002), an unsupervised algorithm based on an eigenvector approach. SFA permits to find the optimal set of functions in a general finite dimensional function space. The exact definition of the algorithm is not necessary for understanding the remainder of the paper, and we will therefore skip it due to space reasons.

3 Methods

Our data source consisted of 36 gray-valued natural images extracted from the Natural Stimuli Collection of van Hateren and preprocessed as suggested in (van Hateren and van der Schaaf, 1998) (block averaging and taking the logarithm of the pixel intensities). We constructed image sequences by moving a window over the images by translation,

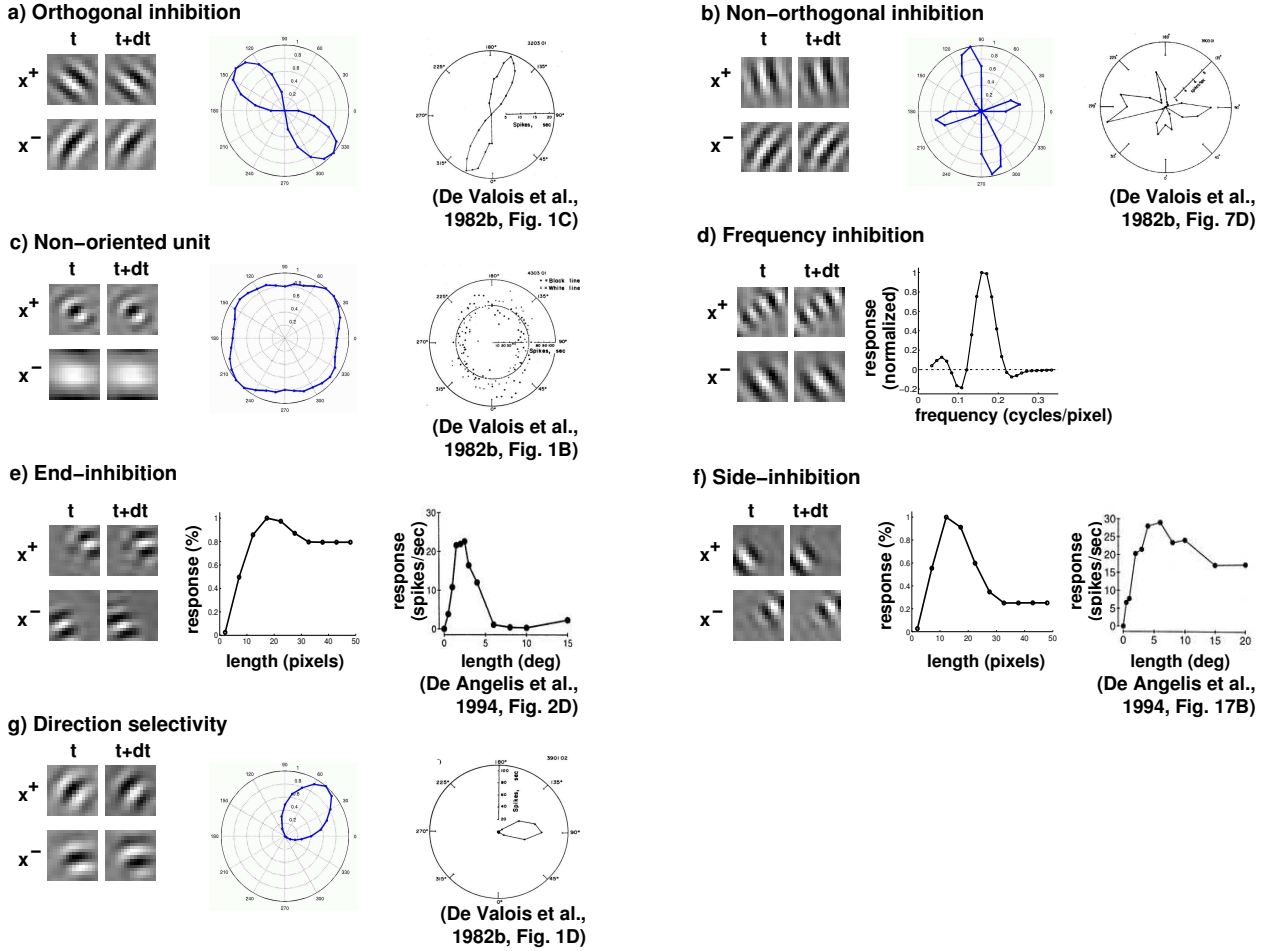


Figure 1 Complex cells behaviors. This figure illustrates the main complex cells behaviors found in the simulation. For each behavior we show from left to right the optimal stimuli, the response of the unit to sine gratings and when available for comparison the response in spikes/sec of neurons of the primary visual cortex of the cat (from (De Valois et al., 1982b) and (DeAngelis et al., 1994)). The polar plots in a–c) and g) show the response of the unit to a sine grating. The orientation varies along the angular direction, while the radial axis measures the response normalized to the maximum. The plot in d) shows the response of the unit to sine gratings of different frequencies. The plots in e) and f) show the response of the unit to a sine grating with varying length or width, respectively.

rotation, and zoom and subsequently rescaling the frames to a standard size of 16×16 pixels. The transformations were performed simultaneously, so that each frame differed from the previous one by position, orientation, and scale. We produced in this way 250,000 frames per simulation. To include temporal information, the input vectors to SFA were formed by the pixel intensities of two consecutive frames at times t and $t + dt$. Simulations showed qualitatively similar results within a reasonable range of parameters, although the distribution of some unit properties could vary. We performed a standard preprocessing step using principal component analysis (PCA) to reduce the dimensionality of the input vectors from $16 \times 16 \times 2 = 512$ to 100, capturing 93% of the total variance.

Since the input vectors are pairs of image patches, the functions g_j can be interpreted as non-linear spatio-temporal receptive fields and tested with input stimuli much like in neurophysiological experiments. The units can have a spontaneous firing rate, i.e. a non-zero response to a blank visual input. We interpret an output lower than the spontaneous one as active inhibition. In order to analyze the units, we first analytically compute for each of them the optimal excitatory stimulus x^+ and the optimal inhibitory stimulus x^- , i.e. the input that elicits the strongest and the weakest output from the unit, respectively, given a constant norm r of the input vector, i.e. given a fixed energy constraint. We choose r to be the mean norm of the training vectors, since we want x^+ and x^- to be representative of the typical input. From the two x^+ patches we compute by Fourier analysis the preferred frequency, orientation, speed, and direction of a unit and the size and position of its receptive field. Although the optimal stimuli carry much information they only

give a partial view of the behavior of a unit, since these are non-linear. We additionally perform experiments with drifting sine gratings in order to compute various unit properties and compare them with physiological results. The sine grating parameters are always set to the considered unit’s preferred ones. Unless stated otherwise, comparisons are always made with experimental data referring to complex cells only.

4 Results

In this section we analyze the first 100 units computed in a single simulation (see also Additional material).

Gabor-like optimal stimuli and phase invariance: The optimal stimuli of almost all units look like Gabor wavelets (Fig. 1), in agreement with physiological data. The response of all these units is largely invariant to phase shift, as measured by their relative modulation rate F_1/F_0 , i.e. the ratio of the amplitude of the first harmonic to the mean response (maximum modulation rate 0.16). They would thus be classified as complex cells in a physiological experiment. 98 units out of 100 had both a Gabor-like x^+ and phase-shift invariance, the missing two cells are described under the paragraph *Tonic cells*.

Active inhibition: In V1 the tuning to orientation and frequency is shaped by active inhibition and other non-linear interactions (Sillito, 1975; De Valois et al., 1982b). In our model inhibition is present in most units and typically makes an important contribution to the output. As a consequence, x^- is usually well-structured and has the form of a Gabor wavelet as well (Fig. 1). Its orientation plays an important role in determining the orientation tuning. It can be orthogonal to that of x^+ (Fig. 1a), but it is often not, which results in sharpened orientation tuning. When the orientation of x^- is very close to that of x^+ , it is possible to observe a second peak of activity at an orientation orthogonal to the optimal one (Fig. 1b) known as *secondary response lobe*, also reported in V1 (De Valois et al., 1982b). On the other hand, we also find units that would be classified as complex by their response to edges and their phase-invariance but which are not selective for a particular orientation (Fig. 1c) (four units out of 100). Similar cells are present in V1 and known as *non-oriented cells* (De Valois et al., 1982b). Figure 2a compares the distribution of the orientation bandwidth of our units with that of complex cells reported in (De Valois et al., 1982b; Gizzi et al., 1990). Our units seem to have a slightly broader tuning but the difference is not significant (one-sided Kolmogorov-Smirnov test, $P > 0.8$). Figure 2b shows the histograms of the angle between maximal excitation and inhibition in our simulation and in (Ringach et al., 2002). The more pronounced peak at orthogonal orientations in the latter case might be due to the small fraction of complex cells in the population considered (24 complex cells out of 75). Simple cells are likely to show maximal suppression at 90° since they are well described by linear functions.

Also in the frequency domain the tuning varies from a sustained response within a range of frequencies to a sharp tuning due to active inhibition (Fig. 1d). Figure 2c shows the distribution of frequency bandwidth in our simulation and in complex cells as reported by De Valois et al. (1982a). The bandwidth was computed from the units’ contrast sensitivity, like in the cited paper. The two histograms are quite different: complex cells have a rather flat distribution, while the bandwidth of our units is concentrated between 0.3 and 1.4 octaves, with a peak at 0.5. This considerable difference might be partly due to digitalization and dimensionality reduction: these two steps reduce the maximal bandwidth of our units to at most 2 octaves (somewhat more on the diagonal). The actual bandwidth would in general be much less since to reach the theoretical limit a unit would need to fire at half of its maximum exactly at 1 and 4 cycles/patch. Simulations with a higher number of input components might yield a broader distribution.

End- and Side-Inhibition: Some of the complex cells in V1 are selective for the length (*end-inhibited cells*) or width of their input (*side-inhibited cells*) (DeAngelis et al., 1994). We computed for each unit a quantitative measure of its degree of end- or side-inhibition by presenting sine gratings of different length and width. We define the end- and side-inhibition index as in (DeAngelis et al., 1994) by the decrease of the response in percent between optimal and asymptotic length and width, respectively. 10 units out of 100 had an end-inhibition index larger than 20%, and 7 units out of 100 had an side-inhibition index larger than 20%. In contrast to (DeAngelis et al., 1994) we only found 2 units that showed large ($> 20\%$) end- and side-inhibition simultaneously. End- and side-inhibited units can sometimes be identified by looking directly at the optimal stimuli: in these cases x^+ fills only one half of the window, while the missing half is covered by x^- with the same orientation and frequency (Figs. 1e,f). This receptive field organization is compatible with that observed in V1 by Walker et al. (1999) in that inhibition is asymmetric and is tuned to the same orientation and frequency as the excitatory part.

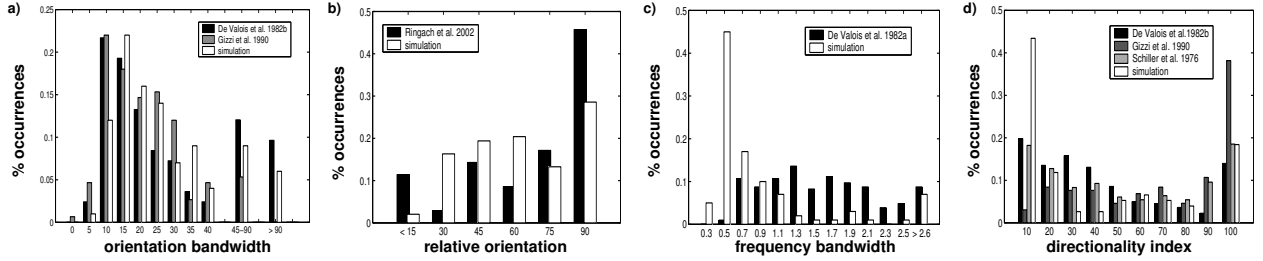


Figure 2 Population statistics. a) The distribution of half-height orientation bandwidth in complex cells and in our simulation. b) The distribution of the angle between the orientation of maximal excitation and maximal inhibition. c) The distribution of half-height frequency bandwidth. The bandwidth is measured by the units’ contrast sensitivity function. Units whose contrast sensitivity does not drop under 50% of the maximum are classified in the last bin. d) The distribution of the directionality index (the data from (De Valois et al., 1982b) also contain simple cells, but in the paper it is stated that there was no significant difference between the two populations).

Direction selectivity: Complex cells in V1 are sensitive to the motion of the presented stimuli. Some of them respond to motion in both directions while others are direction selective (Schiller et al., 1976; De Valois et al., 1982b). Similarly, in our model some units are strongly selective for direction (Fig. 1g) while others are neutral. In the latter case the optimal speed may be non-zero but the response is nearly equal for both directions. We measure direction selectivity by the *directionality index*, given by $DI = (1 - R_{np}/R_p) \cdot 100$, with R_p and R_{np} being the response in the preferred and in the non-preferred direction, respectively. Figure 2d shows the histogram of DI in our simulation compared to those reported in (De Valois et al., 1982b; Schiller et al., 1976; Gizzi et al., 1990).

Tonic cells: The first unit in every simulation codes for the mean pixel intensity, and is thus comparable to the *tonic cells* found in V1 (Schiller et al., 1976). Tonic cells respond to either bright or dark stimuli but do not need a contour in order to respond. We find in addition one unit (the second one) that responds to the squared mean pixel intensity, and therefore to both bright and dark stimuli.

5 Discussion

Other theoretical studies have successfully reproduced the basic properties of complex cells in models based on the computational principles *statistical independence* (Hyvärinen and Hoyer, 2000) or *slowness* (Körding et al., 2003). These models learned units that were equivalent to the classical energy model of complex cells and thus reproduced only the Gabor-like receptive fields and the phase shift invariance. Zetzsche and Rührbein (2001) learned by principal component analysis units with simple- and others with complex-cell characteristics. Additionally, some units of both classes showed end-inhibition. In their experiments, however, the response to collinear Gabor-wavelets was embedded in the system and was not a learned feature. To the best of our knowledge, the model presented here is the first one based directly on input images that is able to learn a population of complex cells showing effects like active inhibition, secondary response lobes, end-inhibition, side-inhibition, direction selectivity and tonic cell behavior.

In summary we have shown that slowness leads to a great variety of complex cell properties found also in physiological experiments. Our results demonstrate that such a rich repertoire of receptive field properties can be accounted for by a single unsupervised learning principle.

Additional material Additional material about the simulation described in this paper such as single-unit analysis can be found online starting at (http://itb.biologie.hu-berlin.de/~berkes/slowness/slowness_index.shtml).

Acknowledgments Supported by a grant from the Volkswagen Foundation. We thank Jack Gallant, Dario Ringach, and Reinhard Eckhorn for fruitful discussions. We also thank Andreas Herz, Tim Gollisch, and Martin Stemmler for useful comments on an earlier version of this manuscript.

References

- Berkes, P., Wiskott, L., 2003. Slow feature analysis yields a rich repertoire of complex-cell properties. *Cognitive Sciences EPrint Archive (CogPrint)* 2804, <http://cogprints.ecs.soton.ac.uk/archive/00002804/>.
- De Valois, R., Albrecht, D., Thorell, L., 1982a. Spatial frequency selectivity of cells in macaque visual cortex. *Vision Res.* 22, 545–559.
- De Valois, R., Yund, E., Hepler, N., 1982b. The orientation and direction selectivity of cells in macaque visual cortex. *Vision Res.* 22 (5), 531–44.
- DeAngelis, G., Freeman, R., Ohzawa, I., 1994. Length and width tuning of neurons in the cat's primary visual cortex. *Journal of Neurophysiology* 71 (1), 347–374.
- Field, D. J., 1994. What is the goal of sensory coding? *Neural Computation* 6 (4), 559–601.
- Gizzi, M. S., Katz, E., Schumier, R. A., Movshon, J. A., June 1990. Selectivity for orientation and direction of motion of single neurons in cat striate and extrastriate visual cortex. *Journal of Neurophysiology* 63 (6), 1529–1543.
- Hyvärinen, A., Hoyer, P., 2000. Emergence of phase and shift invariant features by decomposition of natural images into independent features subspaces. *Neural Computation* 12 (7), 1705–1720.
- Körding, K., Kayser, C., Einhäuser, W., König, P., 2003. How are complex cell properties adapted to the statistics of natural stimuli?, *Journal of Neurophysiology*, in press.
- Ringach, D. L., Bredfeldt, C. E., Shapley, R. M., Hawken, M. J., 2002. Suppression of neural responses to nonoptimal stimuli correlates with tuning selectivity in macaque V1. *Journal of Neurophysiology* 87, 1018–1027.
- Schiller, P., Finlay, B., Volman, S., 1976. Quantitative studies of single-cell properties in monkey striate cortex. I. Spatiotemporal organization of receptive fields. *J. Neurophysiol.* 39 (6), 1288–1319.
- Sillito, A., 1975. The contribution of inhibitory mechanisms to the receptive field properties of neurons in the striate cortex of the cat. *J. Physiol.* 250, 305–329.
- van Hateren, J., van der Schaaf, A., 1998. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc. R. Soc. Lond. B* 265, 359–366.
- Walker, G., Ohzawa, I., Freeman, R., 1999. Asymmetric suppression outside the classical receptive field of the visual cortex. *The Journal of Neuroscience* 19 (23), 10536–10553.
- Wiskott, L., Berkes, P., 2003. Is slowness a learning principle of the visual cortex? In: *Proc. Jahrestagung der Deutschen Zoologischen Gesellschaft 2003*, Berlin, June 9-13. Special issue of *Zoology*, 106(4):373-382.
- Wiskott, L., Sejnowski, T., 2002. Slow feature analysis: Unsupervised learning of invariances. *Neural Computation* 14 (4), 715–770.
- Zetsche, C., Röhrbein, F., 2001. Nonlinear and extra-classical receptive fields properties and the statistics of natural scenes. *Network: Computation in Neural Systems* 12, 331–350.