Learning the selectivity of V4 neurons using a nonlinear two-stage wavelet network

Ulrich Nuding[1,2] and Christoph Zetzsche[1],
[1]Kognitive Neuroinformatik, Universität Bremen
[2]Institut für Medizinische Psychologie, Ludwig-Maximilians-Universität München

{nuding, zetzsche}@informatik.uni-bremen.de

Abstract

Cells in the higher visual areas of primates (visual cortices V2-V4) are highly specialized in the processing of certain image components. The underlying neural representation is thus sparse and distributed, where single units show a high selectivity to specific classes of complex stimuli. In addition, many V4 cells are sensitive to curvature, and do only respond weakly or not at all to elongated test stimuli [1, 2]. The basic principles leading to such a highly selective behavior are not yet fully understood. Linear filters, the classical model for visual neurons, are not capable of reproducing these response properties. First, any linear filter has one (or more) maxima in the frequency domain, and has therefore to give a strong response to a high-contrast sinusoidal grating of appropriate frequency and orientation. However, responses of V2 and V4 neurons to high-contrast sinusoidal gratings are often much lower than for other high-contrast stimuli. Second, linear filters have inherently broader tuning properties in state space than required for the observed high selectivity of the actual neurons.

We have recently shown that a nonlinear AND-like combination of frequency components could provide a mechanism for increasing the selectivity in the desired direction [3]. Here we show that a hierarchical nonlinear network that is trained on the sparse representation of natural images (Fig. 1) can also achieve the desired nonlinear selectivity. The network consists of a linear transformation, a simple two-channel one-way rectification which separates the signal in an ON- and OFF-part, respectively, and a final linear pooling of the nonlinear intermediate coefficients. Each layer is trained separately using as input a large set of natural images. The optimization is based on the goal of statistical independence, using different approximations (2^{nd}- and higher-order). The initial linear transform is implemented as a fixed multi-channel wavelet decomposition, as it has already been shown that the *independent components* of natural images are wavelet-like basis functions [4]. The final linear transform of the rectified wavelet coefficients is calculated using standard Principal Component Analysis (PCA) and Independent Component Analysis (ICA).

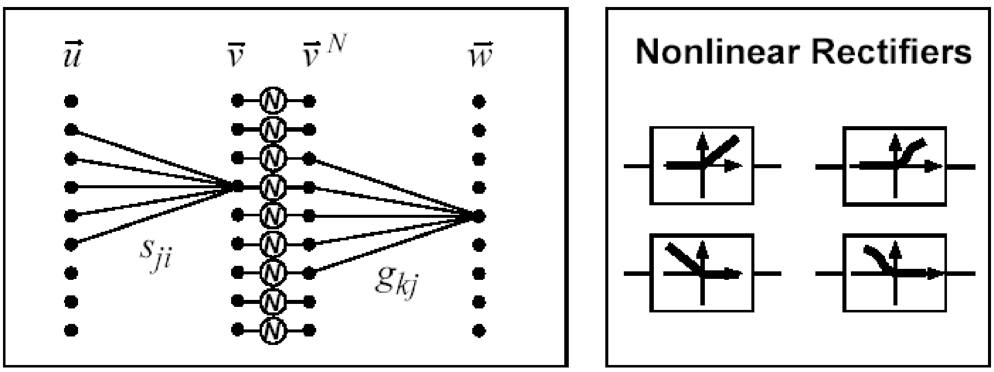

Figure 1: Schematic drawing of the two-stage network (left). Possible nonlinear transducer functions are shown on the right.

In order to analyze the response properties of the system, we investigated a reduced configuration based on an initial wavelet decomposition using two scales and four orientations. For comparison purposes, we used a set of test stimuli similar to the one used in [1]. This set consists of basis functions in Cartesian, polar and hyperbolic coordinates, that is, sinusoidal, polar and hyperbolic gratings. Although there exists no clear definition of an appropriate stimulus parameter space, the various patterns can be roughly associated with different frequencies and curvatures. The output units of our system are treated as single neurons, and their response properties are investigated by testing their responsiveness to the test patterns.

The simulation showed that many of the units are selective for curved patterns and yield only weak responses to straight sinusoidal gratings. An exemplary simulated V4 neuron is shown in Fig. 2.

In conclusion, we have demonstrated that a simple hierarchical network trained with natural images is already capable of producing the selective nonlinear response behavior of neurons in higher areas of the visual cortex. Since these neurons are believed to be substantially involved in object recognition, this indicates that the adaptation to natural scene statistics can also contribute to higher visual functions. In addition, the close relation between the present network and our other model approaches suggests that these nonlinear neural properties may not only be describable in the "opaque" fashion of neural network models but could also be understood using the more formal structural principles of Volterra-Wiener systems and nonlinear AND operations.

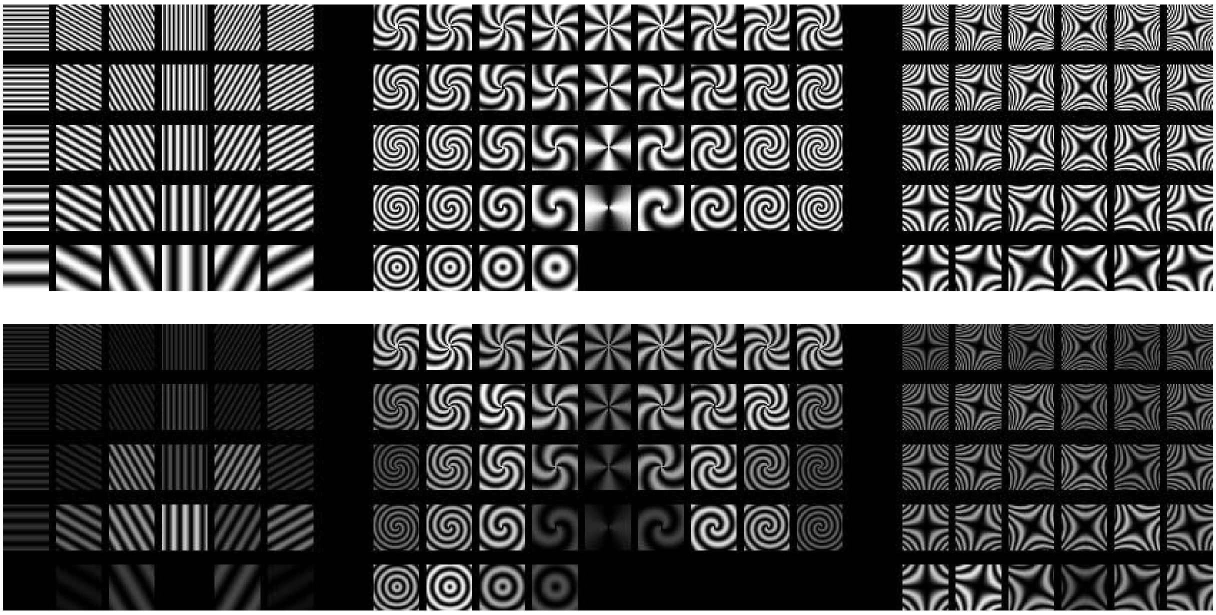

Figure 2: Responses of a nonlinear subunit of the two-stage model. Top row: Test stimulus sets as used in neurophysiological experiments. Bottom row: Responses of a typical model neuron. The contrast of the patterns corresponds to the response strength of the unit. As it can be seen, it shows a selective behavior, and reacts especially to curved test patterns.

References

Study supported by DFG Graduiertenkolleg 267 and SFB TR-8 "Spatial Cognition"

References

[1] Gallant JL, Braun J, and Van Essen DC. Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science* 259: 100—103, 1993.

[2] Pasupathy A, and Connor CE. Shape Representation in area V4: position-specific tuning for boundary conformation. *J Neurophysiol*, 86: 2505—2519, 2001.

[3] Nuding U, and Zetzsche C. Measurement of nonlinear 2nd-order kernels with polyspectra. In Ilg UJ, Bülthoff HH, and Mallot HA (Eds.), *Dynamic Perception: Workshop of the GI section "Computer Vision"*, 189—194, IOS Press, 2004

[4] Olshausen BA, and Field DJ. Sparse coding with an overcomplete basis set: a strategy employed by V1? *Vision Res*. 37(23): 3311—3325, 1997.