Slow features between invariance and selectivity

SPATIAL COGNITION
A5-[ActionSpace]

Tobias Kluth, Sven Eberhardt, Manfred Fahle, Christoph Zetzsche

Introduction

Slow feature analysis (SFA) has been hypothesized as one possible mechanism for unsupervised learning of receptive fields of neurons in the ventral stream of primate visual cortex [1]. The algorithm has been used successfully to train multi-layer feed-forward networks capable of classification and localization tasks. However, current implementations of SFA hierarchical networks employ a fixed nonlinearity on which input weights are trained.

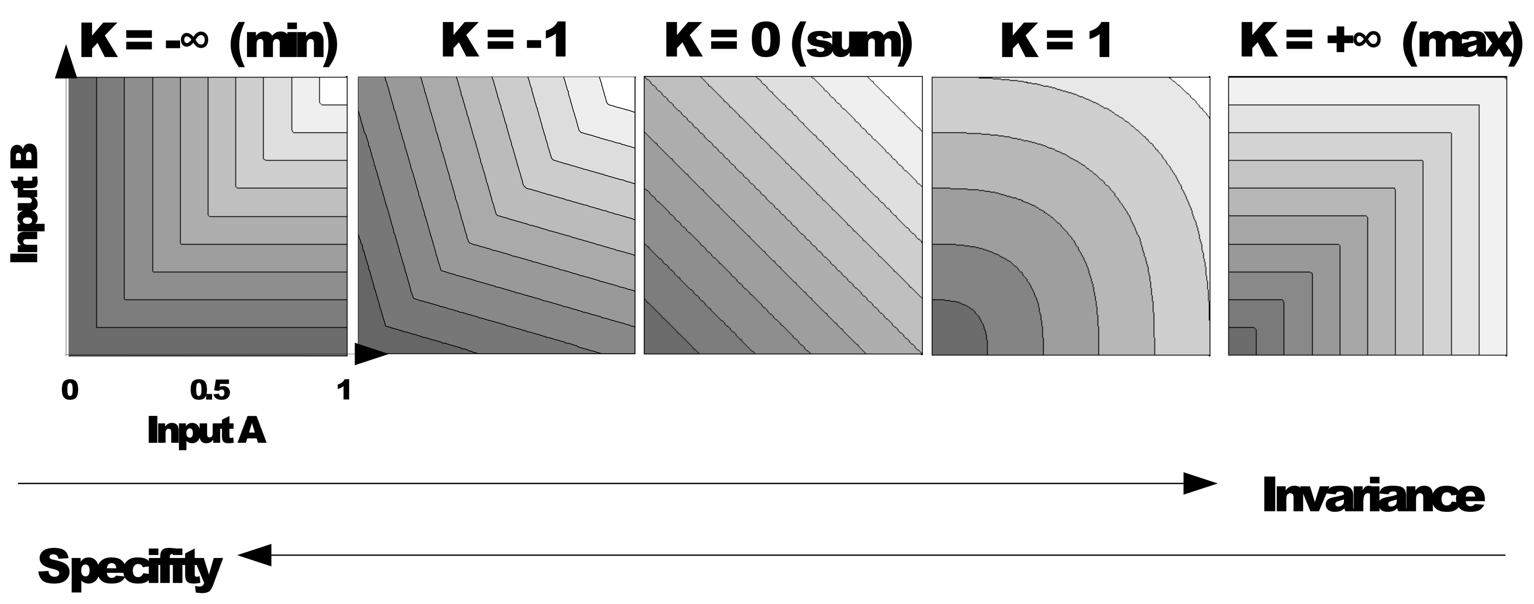
In this study, we introduce a control parameter to vary the nonlinearity between functions implementing invariance and selectivity. The invariance function acts like an OR-operation in that it reacts when any input is active. The selectivity function works like an AND-operation which requires all its input to be present.

Control parameter

Here we introduce a control parameter for the nonlinearity which effects the grade of nonlinearity in the sense of being selective or invariant with respect to the different input dimensions. The function $f_k(x)$ is defined as follows:

$$f_k(x) = \begin{cases} n^{-e^k}\left(\sum_{i=1}^{n} |x_i|^{e^k}\right)^{\frac{1}{e^k}}, & \text{if } k \geq 0 \\ \max_i(|x_i|) - n^{-e^{|k|}}\left(\sum_{i=1}^{n}(\max_i(|x_i|) - |x_i|)^{e^{|k|}}\right)^{\frac{1}{e^{|k|}}}, & \text{if } k < 0 \end{cases}$$

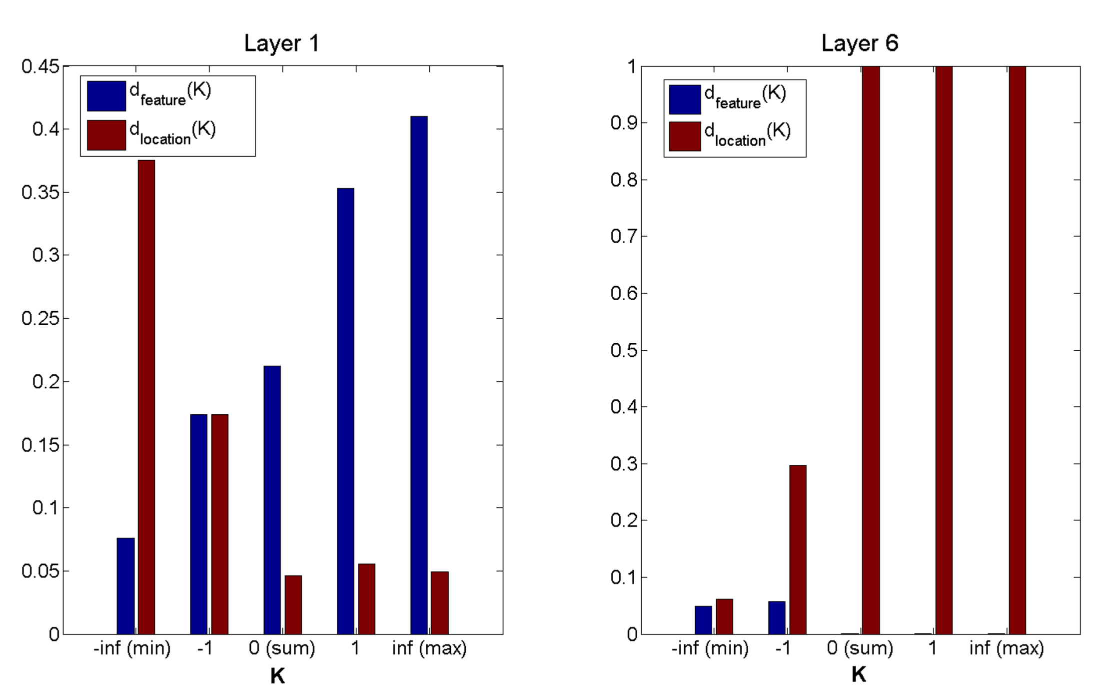
where k is the control parameter, x is the input vector including different dimensions, and n is the number of dimensions. Low values of k (left) produce specific, i.e. template matching, functions and high values prouce invariant, i.e. pooling, functions.



Training

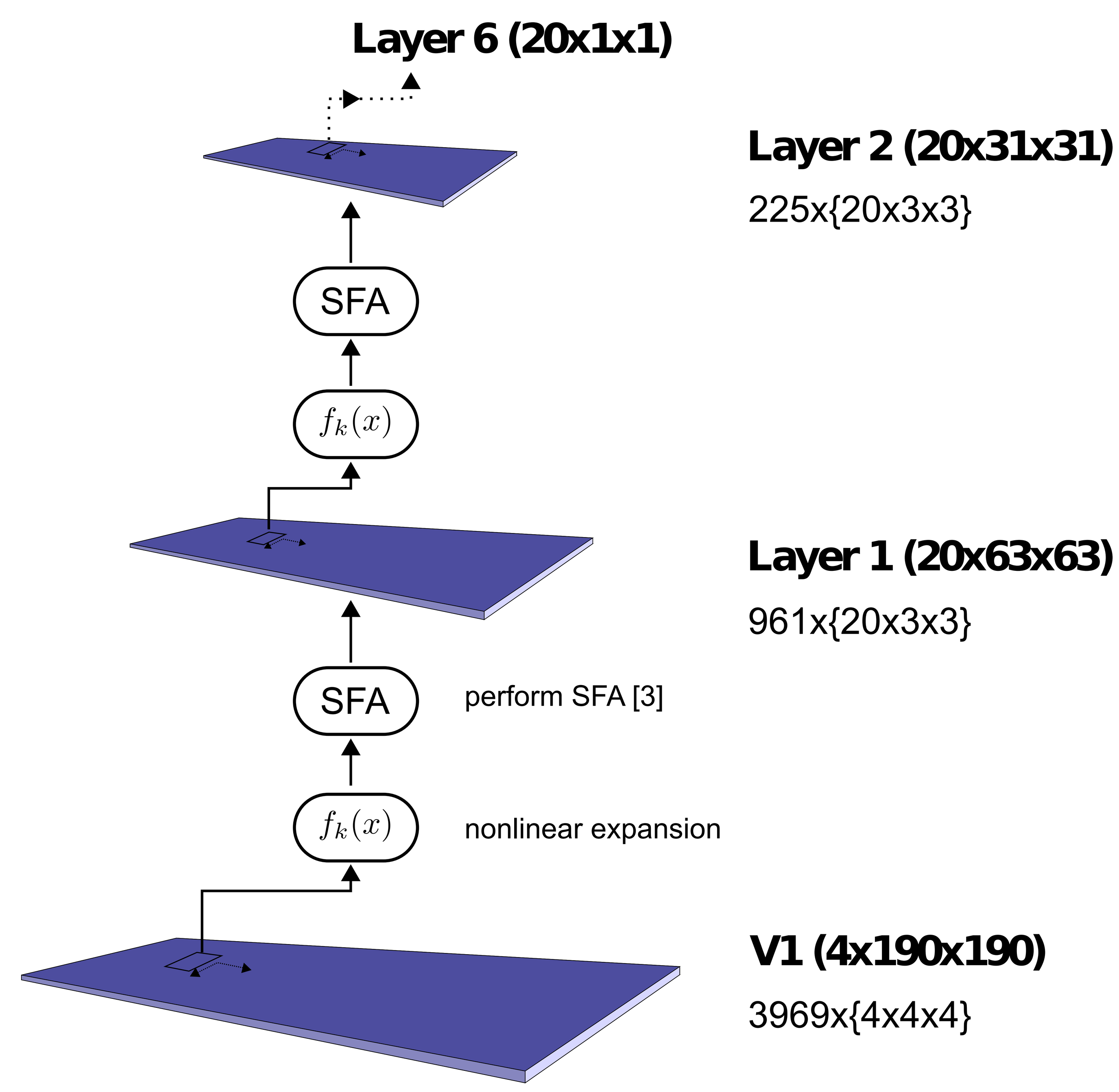
Slow features were extracted from natural videos, which were recorded with a steerable camera. We rotated the camera at one place and extracted the 20 slowest features at a central patch of each layer. Features were replicated to all other locations of the same layer. On the learned features vectors w_K, we looked at the ratio of nonlinear combinations of same features at different locations (d_{feat}) as well as different features from the same location (d_{loc}).

$$d_{feat}(K) = \frac{\sum_{f,x_1,x_2} w_K(f,f,x_1,x_2)^2}{\sum_{f_1,f_2,x_1,x_2} w_K(f_1,f_2,x_1,x_2)^2}, \quad d_{loc}(K) = \frac{\sum_{f_1,f_2,x} w_K(f_1,f_2,x,x)^2}{\sum_{f_1,f_2,x_1,x_2} w_K(f_1,f_2,x_1,x_2)^2}$$



In low layers, AND-like operations ($k < 0$) prefer combinations of different features at the same location, while pooling over locations happens for OR-like operations ($k > 0$). The situation is reversed in high layers. This result indicates that slow features could be used for training in an architecture like the HMAX model [2], where AND and OR operations implement a trade-off between location invariance and pattern selectivity.

Model



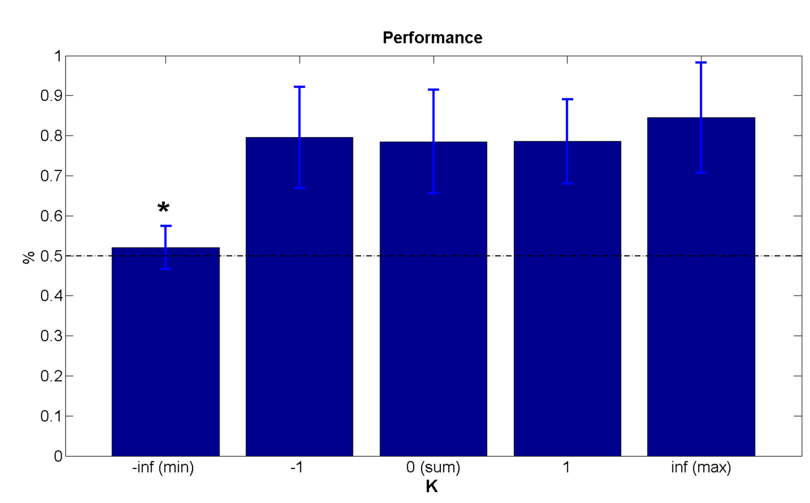
The setup of the hierarchical multi-layer system is as follows:

- **V1**: We used the responses of different Gabor filters (4 orientations at 190×190 locations) as the input of our model.
- **Layer 1-6**: We used the 20 slowest features at each position in the layer.

The receptive fields of a cell in the layer above has three dimensions. The first dimension is the feature dimension or the orientations in the V1. The second and third dimensions are describing the patch size. For each layer the receptive field is illustrated as the number of different patches times $\{n_{feature}, n_x, n_y\}$.

Classification

Patterns learned from different values for k are used to do a two-class-classification task on pictures taken from 20 locations at various rotations. The classifier was trained on one location vs all others; 10 training samples of different orientations were used.



For $k \to -\infty$, performance is almost at chance level. For all other nonlinearities, classification manages around 80% correct. This shows that if the model is too specific (pure AND like operations), features required for spatial classification cannot be extracted.

References

1. Berkes P, Wiskott L (2005). Slow feature analysis yields a rich repertoire of complex cell properties. J Vis 5(6):579-602
2. Riesenhuber M, Poggio T (1999). Hierarchical models of object recognition in cortex. Nature Neuroscience 2:1019-1025.
3. Wiskott L, Sejnowski T (2002). Slow feature analysis: Unsupervised learning of invariances. Neural Computation 14(4): 715-770.

Universität Bremen
DFG Deutsche Forschungsgemeinschaft
UNI FREIBURG