# A model of rapid surface detection in primate visual cortex

Michael Schmuker, Ursula Körner, Edgar Körner, Marc-Oliver Gewaltig, Thomas Wachtler

2003

**Preprint:**

This is an accepted article published in {B}eiträge zur 6. {T}übinger {W}ahrnehmungskonferenz. The final authenticated version is available online at: https://doi.org/[DOI not available]

# Modeling surface detection with integrate-and-fire neurons

*Michael Schmuker* [1,2], *Marc-Oliver Gewaltig* [2], *Ursula Körner* [2], *Edgar Körner* [2], *and Thomas Wachtler* [1]

1: Albert-Ludwigs-Universität Freiburg, Institut für Biologie III, Schänzlestr. 1, 79104 Freiburg, Germany
2: Honda Research Institute Europe GmbH, Carl-Legien-Str. 30, 63073 Offenbach/Main, Germany

## Introduction

Beginning with the work of Hubel and Wiesel, the representation of retinal input in V1 has mostly been thought of being dominated by luminance-contrast- or color-contrast edges [1]. However, recent experimental results suggest that this is not the whole story. Komatsu et al. have reported that about one third of the neurons in the perifoveal area in V1 of macaque monkeys respond to homogeneous surfaces in their receptive field[2]. Moreover, the receptive fields of these cells do not have an antagonistic surround, which also conflicts with the classical view of receptive fields in V1.

The problem of surface representation also appears in computer vision when it comes to image segmentation. Any edge-based vision system, given solely an edge-map of the input image, faces the problem of finding corresponding edges, i.e. figuring out which edges flank the same object or image region. In absence of other cues about the structure of the image, or some knowledge of the objects presented, this is a very difficult and error-prone task. This holds true especially when processing highly detailed natural images.

To overcome these limitations, several methods have been devised to increase the reliability of edge-based segmentation. Among them are contour completion with saliency maps, which is basically an application of the Gestalt laws, or top-down mechanisms (see e.g. [3], for an overview). Further, it has been speculated that contour completion can have a neural correlate in V1, of the kind that colinear edge-respondent cells with neighboring receptive fields interact in order to complete colinear line fragments [4].

However, all of these mechanisms suffer from one caveat: They are basically too slow to explain the *fast* segmentation capabilities of the primate brain. As Thorpe et al. pointed out, primates can perform image segmentation and object classification in a purely feed-forward manner, using only 10-15 ms per processing stage ([5], [6], [7]). That is, each neuron in the processing chain has time to fire at most 1-2 spikes. This especially means that there is simply not enough time for lateral interactions between neurons to become effective, let alone top-down interaction.

In this contribution, we present a network of spiking neurons performing surface detection. This is achieved by using a latency-based coding strategy for the input image and subsequent spike coincidence detection. Moreover, the network works in a purely feed-forward manner, using only 1 spike per layer, thus meeting the timing constraint mentioned above. Further we show what benefits information about homogeneous image regions has in cortical processing of visual input.

## Methods

First of all, the problem is how one defines the "surface"-property of a given image-patch. Gewaltig et al. devised a variance-based approach, which this work is based on [8]. A surface is herein defined as an image region with low grey-value variance. For each image pixel, the grey-value variance in its surround (usually a patch of 5×5 pixels) is calculated. By applying a threshold function to the resulting image, regions with high grey-value variance are masked out, available to further processing.

While the aforementioned work used a computational model, the model presented here is based on a network of *spiking neurons*. See *Fig. 1* for a description of the underlying homogeneity detection mechanism.

### Modeling homogeneity detection

To account for bright and dark surfaces, we used two surface-detection layers: The on-layer detects bright surfaces, the off-layer detects dark surfaces.

In *Fig. 1*, only one detector unit is shown. Moving from single units to image-sized layers of surface-detectors, the connection pattern for each single unit is kept, yet the receptive fields are shifted one by one to cover the whole image.

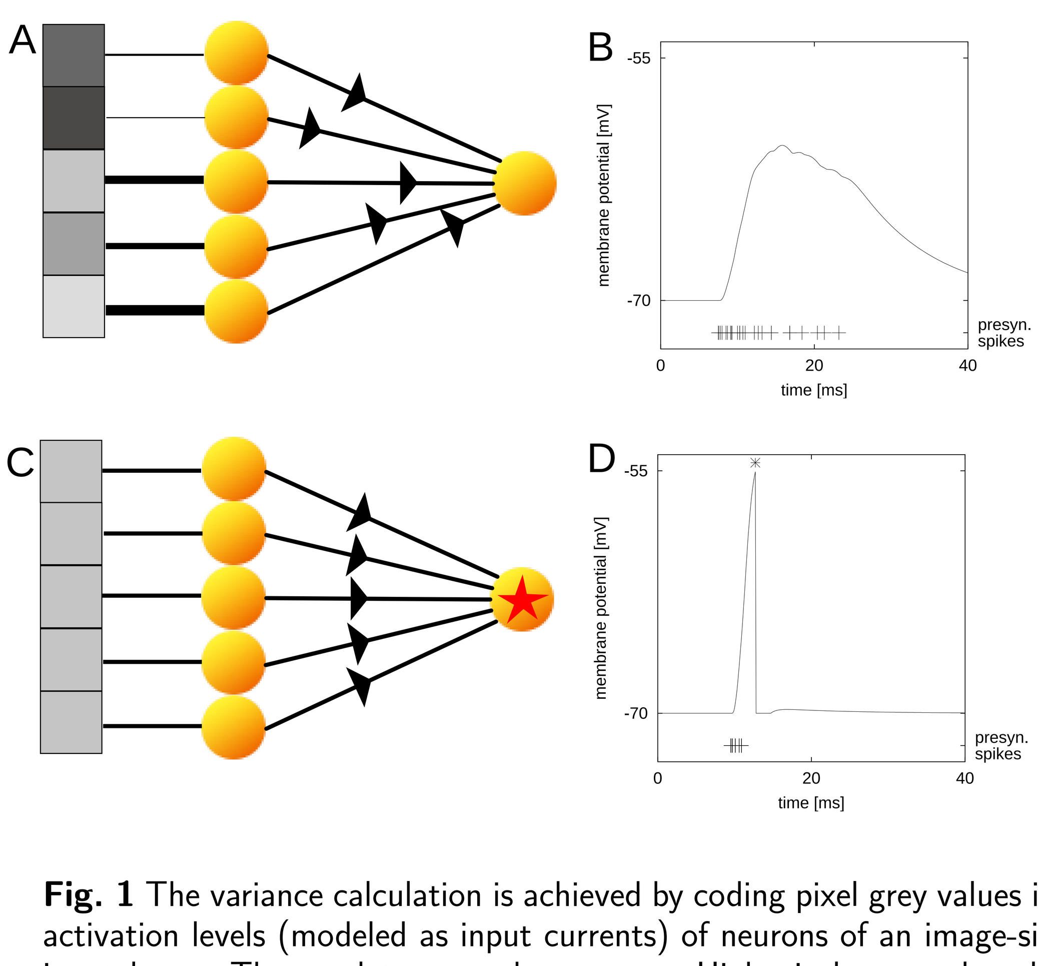


**Fig. 1** The variance calculation is achieved by coding pixel grey values into activation levels (modeled as input currents) of neurons of an image-sized input layer. Thus, a latency code emerges: High pixel grey values lead to high neuron activation levels, which in turn lead to low-latency or early spikes, and vice versa. **A)** If grey values in a given patch in the input image are highly variant, the surface-detector neuron does not respond, because the incoming spikes are too scattered in time to drive it to threshold. **B)** The time-course of the surface-detector membrane potential with non-homogeneous input. **C)** Homogenous grey-values lead to well synchronised pulse packets from the respective neurons in the input layer. These are detected by the surface-detecting neuron, which functions as a coincidence detector. **D)** Membrane-potential time-course with homogenous input. Note that only the behaviour of the on-surface detector is shown here. The off-detector works complementary (see text).

### Incorporating edge-detectors

In order to investigate the benefits of surface detection to image processing, we incorporated edge-detectors into our model. The idea that the surface-detector neurons, when activated, inhibit edge detectors with overlapping receptive fields. *Fig. 2* gives a functional overview of the system.

Interaction follows the paradigma "Where there is a surface, there is no edge": surface-detector neurons, when activated, inhibit edge detectors with overlapping receptive fields.

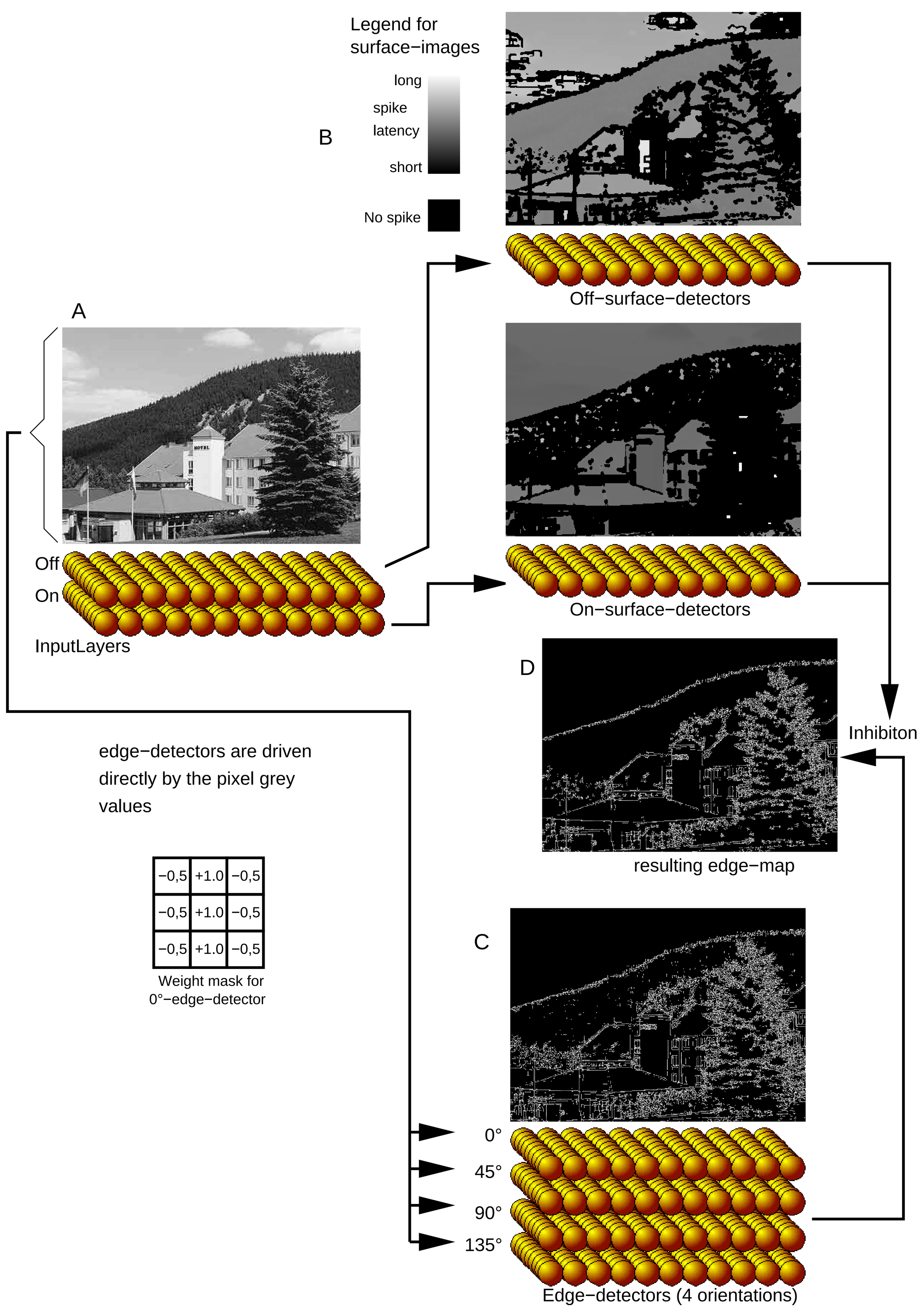


**Fig. 2** Schematic overview of the neural network. The images are to bee seen as 2-dimensional representations of neuron activity in the respective layers, except for the input image. **A)** In the two input layers, the input image (left) is encoded to spike latency. The on-layer encodes bright pixels to shorter latencies and dark pixels to longer ones. The off-layer works complementary: bright pixel results in longer, dark pixel in shorter latency. **B)** The two surface detection layers encode the surface property at each image pixel: if a pixel is black, the respective neuron did not fire a spike, if it is grey, the shade of grey indicates the latency with which the spike was fired. The legend illustrates this. **C)** edge-detection layers for 4 orientations are directly driven by the pixel grey values in the input image, that is they get injected an input current proportional to the pixel grey value. They are only driven in on-fashion, i.e. dark pixel results in low input current. For the 0-edge-detector neurons, the receptive field structure is shown in the box to the left. **D)** When surface-detecting neurons are allowed to inhibit edge-detection neurons, the edge-map gets cleared up significantly. See also *Fig. 3* for images filtered without edge processing.
Please note that in the input layers, the number of neurons is identical to the number of pixels in the input image. To avoid aliasing effects on the margin, the surface- and edge-detection layers are slightly smaller.

### Simulation environment and parameters

As our model needs at most 1 spike per neuron in each processing layer, the refractory period of each neuron was set to a value larger than the simulation time: every neuron can fire exactly 0 or 1 spikes. All neurons are modeled as standard integrate-and-fire neurons. Coincidence-detecting neurons are modeled by lowering the time constant for synaptic currents and reducing membrane capacity, in order to account for accurate coincidence detection of incoming spikes.

All simulations were carried out using the NEST-simulator [9]. Simulation time was 50 ms with a resolution of 0.1 ms. Because the number of neurons simulated is a multiple of the image pixel number, simulation duration varied and was in the range of approx. 1 Minute for 131.072 neurons to about 18 minutes for 1.809.504 neurons, using 7 processors of an 8×900 MHz SunFire V880.

## Results

First, our results show that variance-based surface detection is feasible with spiking neurons. Second, by masking image regions with surface-detector activity, only the most salient regions come through. *Fig. 3* illustrates this using natural images processed with surface-detector activation maps. The third point is that when edge-detector responses are inhibited by surface-detector activity, spurious edges are significantly reduced, leaving over only edge-detector responses along the most salient contours (*Fig. 2D*). This can facilitate edge-based image segmentation.

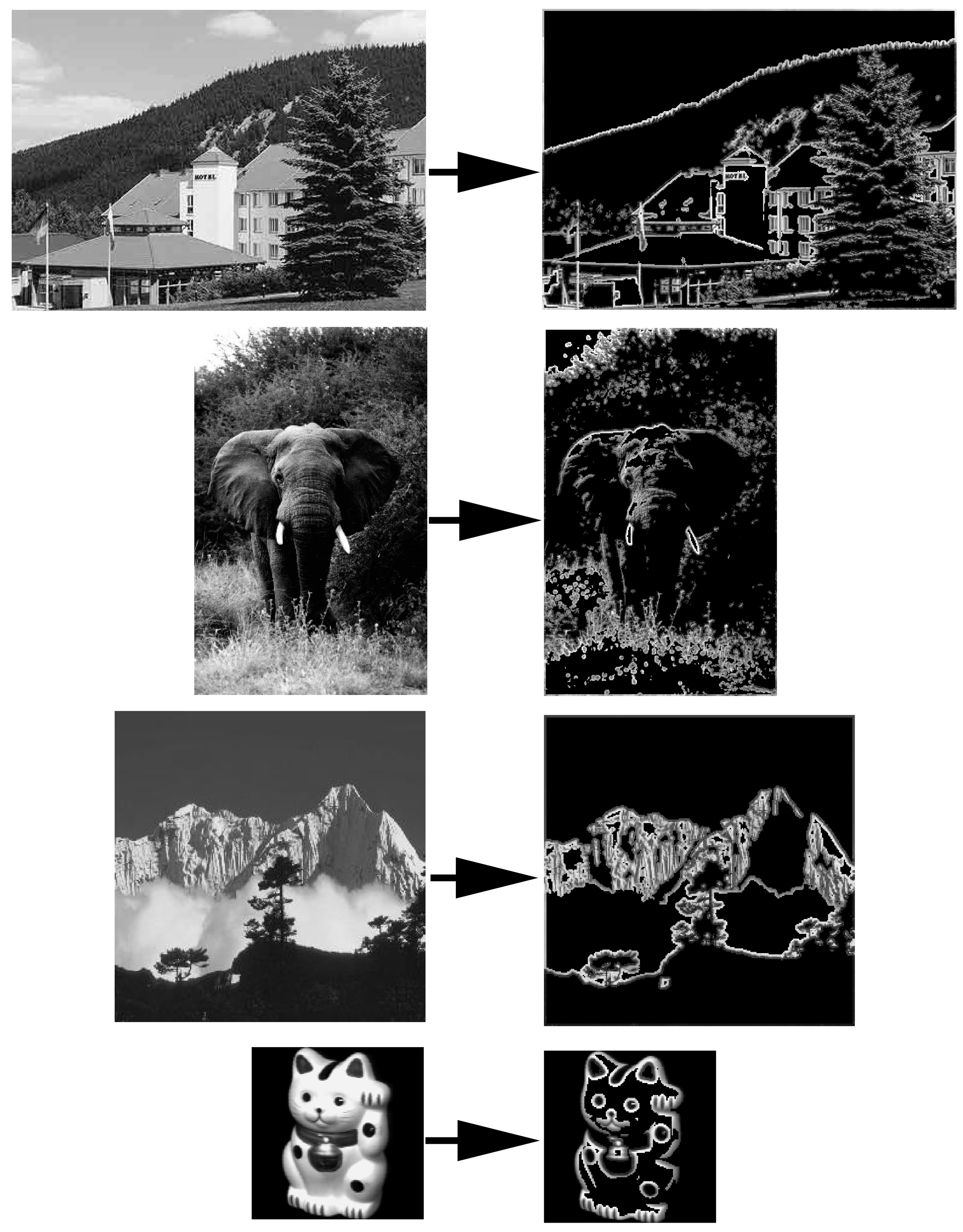

**Fig. 3** Various images and the resulting images after processing with surface-detector activation maps. Image regions with surface-detector activation have been masked out using a free image manipulation program. Only the most salient image regions are left over. Noise and low-frequency clutter have effectively been reduced. Note that, for all images, all processing parameters such as receptive field size and neuron settings were exactly the same and yielded qualitatively similar results, although image resolutions differ heavily. This infers that the mechanism is sufficiently stable against size variances. The above 2 pictures have been scaled down to 65 % (hotel scene) resp. 60 % (elephant) of their original sizes, in order to fit the page. It is also interesting to note that, in the filtered images, 3D-perception seems to be reduced.

## Discussion

- How can the required synchrony of spike responses emerge?
  - stimulus onset effect without mask as in experiments by Thorpe et al. in [6]: synchrony comes naturally, just as in our model
  - during normal vision, when segmentation has more time, synchrony can emerge by lateral interaction, or oscillations (see [10])
- Possible brain regions/paths where surface detection is performed and reasons for choosing them:
  - K-path: large receptive fields without antagonistic center-surround structure
  - M-path: fast, color-insensitive, inhibition during saccades could account for spike synchronisation upon fixation. However, M-cells are very contrast-sensitive.
  - P-path: fast color-segmentation, also grey-level sensitive, some LGN-parvo-cells have uniform receptive fields without center-surrond.

  As research on surface-responsiveness is just beginning, the above can only be speculations.
- Open questions and outlook:
  - What role can color have?
    - homogeneous color surface detection instead of grey-level (P-path)
    - but: fast segmentation works also with black and white images, see [11]
  - planned: psychophysical experiments to confirm surface influence in fast image segmentation.
  - but: maybe only perceptual 3D-reconstruction is impaired, as suggested by the sample images in *Fig. 3*
  - planned: integration into the computer vision system at Honda Research Institute to compare the performance of the surface segmentation approach to other segmentation techniques, in terms of speed and performance

## References

[1] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.*, 160:106–154, 1962.

[2] H. Komatsu et al. Surface representation in the visual system. *Cogn. Brain Res.*, 5:97–104, 1996.

[3] S. Ullman. *High-Level Vision*, chapter 8. MIT Press, Cambridge MA, 1995.

[4] W. Li and C. D. Gilbert. Global contour saliency and local colinear interactions. *J. Neurophysiol.*, 88(5):2846–2856, 2002.

[5] S. Thorpe et al. Spike-based strategies for rapid processing. *Neural Networks*, 14:715–725, 2001.

[6] S. Thorpe et al. Speed of processing in the human visual system. *Nature*, 381:520–522, 1996.

[7] M. Fabre-Thorpe et al. A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *J. Cogn. Neurosci.*, 13(2):171–180, 2001.

[8] M.-O. Gewaltig, U. Körner, and E. Körner. A model of surface detection and orientation tuning in primate visual cortex. In Eric de Schutter, editor, *Computational Neuroscience: Trends in Research 2002*. Elsevier Science, Amsterdam, 2002. in press.

[9] M. Diesmann and M.-O. Gewaltig. Nest: An environment for neural systems simulations. In V. Macho, editor, *Forschung und wissenschaftliches Rechnen, GWDG-Bericht*. Ges. für Wiss. Datenverarbeitung, Göttingen, 2002. in press.

[10] E. Körner, M. O. Gewaltig, U. Körner, A. Richter, and T. Rodemann. A model of computation in neocortical architecture. *Neural Networks*, 12(7-8):989–1005, 1999.

[11] A. Delorme, G. Richard, and M. Fabre-Thorpe. Ultra-rapid categorisation of natural scenes does not rely on color cues: a study in monkeys and humans. *Vision Res.*, 40(16):2187–2200, 2000.