A model of surface detection and orientation tuning in primate visual cortex.

Marc-Oliver Gewaltig, Ursula Körner, Edgar Körner

Future Technology Research, Honda R&D Europe (Deutschland) GmbH,

D–63073 Offenbach/Main, Germany

Abstract

In this contribution we propose that surface detectors, responding to areas of homogeneous gray-level may be used to support orientation selective cells in V1. By inhibiting orientation detectors within their input region, surface detectors act as local feedback system, emphasizing objects with large, continuous contours and enable the fast generation of an initial hypothesis about the input. The performance of this model is demonstrated, using a set of real-world images. We demonstrate that for images with a high degree of low-frequency clutter the performance of orientation detection can be considerably improved.

1 Introduction

In early visual processing, speed matters. If an animal is confronted with a visual scene, it is important to quickly arrive at an hypothesis about its main

parts so that an appropriate reaction (e.g. escape) is possible. In order to do such a quick segmentation of the visual input, the information has to be reduced to its most salient parts. After a first object hypothesis has been established at higher processing levels, a top down signal may then enable a refined analysis of object detail (see e.g. [1]).

This poses the question, how the visual system can perform this segmentation without knowledge of the types of objects which are in the scene. We propose that information on continuous image regions may be used to confine edge detection to borders between mid-sized objects.

2 Biological background

[Fig. 1 about here.]

In primates there are two major pathways through which information flows from the retina via the LGN to the primary visual cortex. The parvocellular (P) and the magnocellular (M) pathway. M-cells in the LGN have larger receptive fields and are highly contrast sensitive. P-cells have smaller receptive fields and their contrast sensitivity is comparatively low, however, many of them are sensitive to color (wavelength). M and P cells account for 90% of all LGN relay cells. The remaining 10% belong to a third, less well investigated pathway, the so-called koniocellular (K) pathway, which is thought to be a phylogeneticly old part of the visual system.

K-cells in LGN have very large receptive fields and their projections to V1 terminate in the cytochrome-oxidase (CO) blobs of the superficial layers [2,4]. In contrast, inputs from the M and P pathway terminate in layer IV. Moreover, while M and P cells exclusively project to V1, K cell inputs have been identified in higher areas of the ventral visual pathway [5].

Because of their large receptive fields, K cells are not able to resolve fine details, however, they may provide information about large homogeneous regions. Tuning of orientation selective cells in layer IV is relatively broad, while it is considerably sharper in layers II and III. We propose that surface information, provided by the koniocellular pathway supports orientation selectivity in layer II/III and, thus, may contribute to the improved tuning.

3 The model

In this contribution, we present a computational model, which demonstrates how surface information can be used to support and enhance the response-properties of orientation selective cells in V1. Starting from the retinal image, we model LGN responses for two pathways, the koniocellular and the magnocellular pathway. The model system is derived from the schema, shown in Fig. 2 and consists of 4 topographically organized layers.

(1) a retinal input layer A,

(2) a layer of koniocellular LGN cells C,

(3) a layer of K-driven surface detectors, E

(4) a layer, representing the input to orientation selective cells F.

Here, we restrict the discussion to the properties of the surface detectors and do not explicitly model orientation selective cells of the M-path. Rather, we show how the input F to these cells is modified by the surface detection system.

[Fig. 2 about here.]

3.1 LGN activity

Receptive fields of K-cells in the LGN are very large and often have no antagonistic surround [6]. Thus, they will be fully activated, even if a stimulus is larger than their receptive field. We model the receptive field of LGN K cells by a two-dimensional Gaussian $g(x, y)$ with standard deviation σ_1:

$$h(x, y) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp\left(-\frac{x^2 + y^2}{4\sigma_1^2}\right) \tag{1}$$

The firing rate of an LGN K cell at position (x, y) is given by $\mathsf{C}[x, y] = \Theta_1\left(\mathsf{B}[x, y]\right)$ with the sigmoid activation function

$$\Theta_1(z) := \frac{1}{1 + \exp\left(-2b_1\left(z - \theta_1\right)\right)} \tag{2}$$

and $\mathsf{B}[x, y] = \left(\mathsf{A} * h\right)(x, y)$ with threshold θ_1 and slope b_1. The threshold is determined by the mean activity in the image. Note that activities are

normalized to unity.

3.2 Surface detectors

The K-cell layer C projects to a layer of surface detectors E. The receptive field of a surface detector at position (x, y) is modeled by a circular patch $\gamma(x, y)$ within C. surface detectors respond best, if their receptive field is covered by an area of *homogeneous* color or gray-level. Surface detectors effectively evaluate the variance of color or gray-level in their circular receptive field.

$$\mathsf{D}[x, y] := \sqrt{\frac{1}{P} \sum_{(m,n)\in\gamma(x,y)} \left(\mathsf{C}[i-m, j-n] - \langle \mathsf{C} \rangle_{\gamma(i,j)}\right)^2} \tag{3}$$

with

$$\langle \mathsf{C} \rangle_{\gamma(x,y)} := \frac{1}{P} \sum_{(m,n)\in\gamma(x,y)} \mathsf{C}[m-x, n-y]. \tag{4}$$

The activation of the surface detectors is given by $\mathsf{E}[x, y] = \Theta_2 \left(\mathsf{D}[x, y]\right)$, where $\Theta_2(\cdot)$ is a sigmoid activation function equivalent to (2). The threshold θ_2 is again chosen according to the mean activity of their input.

The surface detectors project one to one to an interneuron layer. Each activated surface detector will thus, trigger inhibition which will act on all edge detectors whose receptive field overlaps with that of the surface detector. Since the receptive fields of the surface detectors are considerably larger than those of the orientation selective cells, one surface detector will cause inhibition of a large number of orientation detectors. In the computational model this is achieved by masking the input image with the activity of the inhibition layer E, thus, arriving at the effective input to the edge detectors $\mathsf{F} := \mathsf{A} \cdot \mathsf{E}$.
Here, we used a multiplicative interaction to mask the input with the surface detector response. It is also possible to use an additive interaction. In this case, an additional normalization may be required, but the results are qualitatively identical.

4 Results

[Fig. 3 about here.]

Figure 3 shows how the model performs on natural images. The images are an arbitrary selection from a larger set of images on which the system was tested. Note that the images differ considerably in resolution and are scaled to fit to this display.

Figure 3 **A** shows an example from the COIL database [7]. The surface detectors emphasize the contours of the object. However, images from the COIL data-base are optimized for edge-based object recognition. Here, almost any edge-based object recognition systems performs well, since the images are high-contrast and low-noise.

The example in Figure 3 **B** is more realistic, since the main object (the monkey) is embedded in a cluttered environment. If an edge-detection system was directly applied to this image, the response would be dominated by the contributions from the rectangular grid. However, surface detectors respond best to homogeneous regions, contributed by the monkey and the tree-trunk. Thus, the image is reduced to the contours of these objects, suppressing much of the grid and the high-resolution clutter. This reduction in detail is an important guide for subsequent edge-based recognition systems.

5 Discussion

We have presented a model for surface detection which is based on local homogeneity. Surface detectors and edge detector cooperate according to the rule that surfaces and edges are mutually exclusive. In our model, this relation is hard-coded. In a natural system, however, it should be susceptible to learning. Surface detectors are activated by continuous regions of high local homogeneity. If a surface detector is activated, it inhibits all orientation detectors which

have their receptive field inside the larger receptive field of the surface detector. At the borders of such regions, edge detection is possible. Since extended homogeneous regions can in most cases be attributed to mid-sized objects, the surface detection system emphasizes the *contours* of such objects. The spatial scale is defined by the receptive field size of the surface detectors.

Because surface and edge detection work as parallel feed-forward systems, the proposed mechanism does not require additional time. It may, thus, support the fast generation of an initial hypothesis about the input. Subsequent processing steps may trigger top-down feedback to suppress the surface driven inhibition and enable access to fine detail of the input.

References

[1] Edgar Körner, Marc-Oliver Gewaltig, Ursula Körner, Andreas Richter, and Tobias Rodemann. A model of computation in neocortical architecture. *Neural Networks*, 12(7–8):989–1005, 1999.

[2] J. S Lund, T. Yoshioka, and J.B. Levitt. Substrates for interlaminar connections in area v1 of macaque monkey cerebral cortex. In A. Peters and K.S. Rockland, editors, *Cerebral Cortex: Primary Visual Cortex in Primates*, pages 37–60. Plenum Press, 1994.

[3] J.D. Boyd, J. Mavity-Hudson, and V.A. Casagrande. The connections of layer 4 subdivisions in the primary visual cortex (V1) of the owl monkey. *Cerebral Cortex*, 10(7):644–662, 2000.

[4] S. H. C. Hendry and R. C. Reid. The koniocellular pathway in primate vision. In *Annual Review of Neuroscience*, volume 23, pages 127–153. Annual Reviews, 2000.

[5] A. Hernandez-Gonzalez, C. Cavada, and F. Reinoso-Suarez. The lateral geniculate nucleus projects to the inferior temporal cortex in the macaque monkey. *Neuroreport*, 5(18):2693–2696, 1994.

[6] R. W. Rodieck. *The first steps in seeing*. Sinauer Associates, 1998.

[7] S.K. Nayar, S.A. Nene, and H. Murase. Real-time 100 object recognition system. In *Proc. of ARPA Image Understanding Workshop*, Palm Springs, 1996.

Figure Captions

Fig. 1 Sketch of the major routes of input from the dorsal lateral geniculate nucleus (LGN) of the thalamus to the monkey primary visual cortex (V1). Capital letters indicate thalamic input. Small letters indicate further relay of these signals within cortex. (adapted from [2] and [3])

Fig. 2 Schema of the proposed model.

Fig. 3 Example images, demonstrating the operation of surface detectors. Original image (left), inverse surface detector response (middle), strongest response shown as black, and input to orientation selective cells (right). For all images, the same parameter configuration of the model was used. In all cases, the surface detectors emphasize the contours of the object.. **A** Example from the COIL database [7] **B** Object embedded in a cluttered environment. A distractor (the grid) overlaps with the main object in the scene (the monkey).

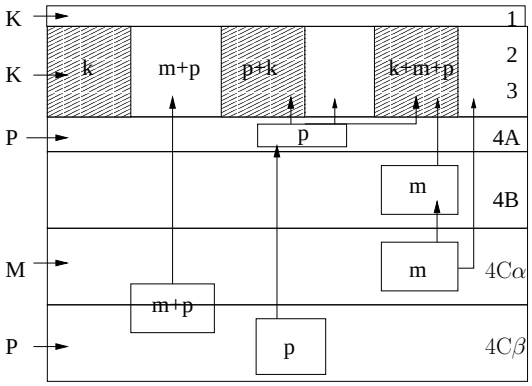

Fig. 1.

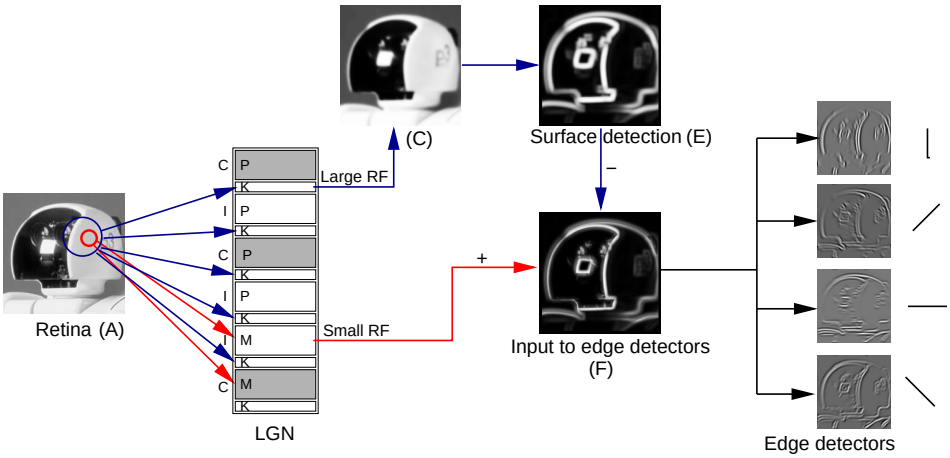
(C)
Surface detection (E)
Large RF
Retina (A)
Small RF
Input to edge detectors
(F)
LGN
Edge detectors

Fig. 2.

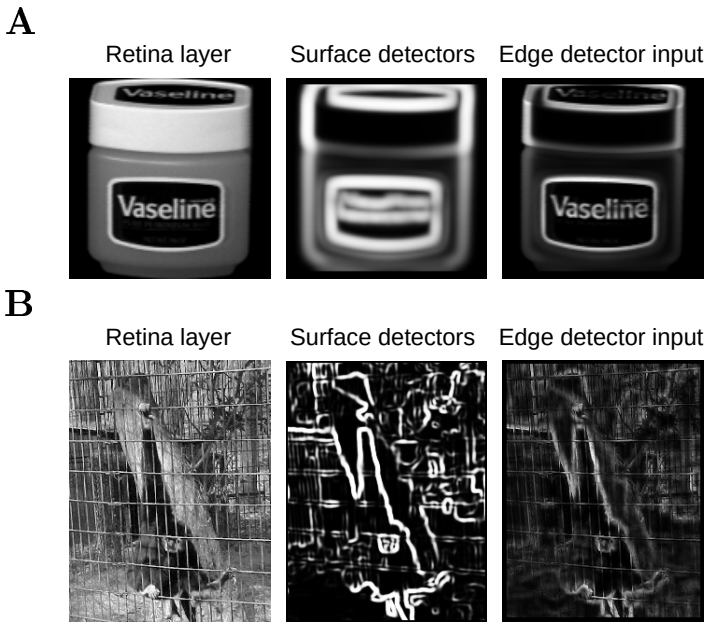
A
Retina layer
Surface detectors
Edge detector input
Vaseline
B
Retina layer
Surface detectors
Edge detector input

Fig. 3.