# Modeling Homogeneity Detection in Primate Visual Cortex with Spiking Neurons

---

Diplomarbeit

vorgelegt von

Michael Schmuker

im September 2003


angefertigt am
Institut für Biologie III
Fakultät für Biologie
Albert-Ludwigs-Universität
Freiburg im Breisgau

betreut von

Dr. Marc-Oliver Gewaltig
Honda Research Institute Europe, Offenbach

und

Dr. Thomas Wachtler
Prof. Dr. Ad Aertsen
Neurobiologie und Biophysik, Universität Freiburg

# Danksagungen

Mein Dank gilt natürlich zu aller erst meinen Eltern, die mir alle Chancen eröffneten.

Außerdem danke ich meinen Betreuern:

Marc-Oliver Gewaltig, für die kompetente Betreuung, sowohl in wissenschaftlichen als auch in stilistischen Fragen,

Thomas Wachtler, für die hilfreichen Kommentare während der Entstehung der Arbeit und zum Manuskript,

Ad Aertsen, der immer mit Rat und Tat zu helfen wußte, wenn es nötig war.

Weiterhin geht mein Dank an

Frau Körner, die zu jeder meiner Detailfragen rund um die Neurowissenschaften eine Antwort oder zumindest eine Referenz wußte,

Renate Czech und Claudia Schäfer, ohne die vieles komplizierter gewesen wäre,

Tobias, Rüdiger, Lars und Julian, für anregende Diskussionen, hilfreiche Tips und nette Gespräche,

Volker, Inna, Henning, Sven und Stephan, für erlösende Kaffeepausen wenn der Streß überhand zu nehmen drohte,

Martin, Daniel, Franz, Uli, Evi, Jan, Hilmar, Anna, Henning, Andreas, Uli und Giso, für alles, was nichts mit Neurowissenschaft zu tun hat.

Diese Liste ist sicherlich nicht vollständig; deswegen möchte ich auch allen, die ich zu erwähnen vergessen habe, meinen herzlichsten Dank aussprechen: Ihr wißt wer Ihr seid!

Zu guter Letzt gilt mein herzlichster Dank meiner Freundin Caroline, die in allen Höhen und Tiefen stets zu mir gehalten hat.

# Zusammenfassung

Die Frage, wie das Gehirn visuelle Information verarbeitet, beschäftigt Neurowissenschaftler schon seit mehreren Generationen. Ein Meilenstein dieser Forschung sind die Ergebnisse von Hubel und Wiesel (Hubel & Wiesel, 1962), die nachweisen konnten, daß Zellen in der primären Sehrinde der Katze bevorzugt auf orientierte Kontrastkanten antworten. Es stellte sich heraus, daß dieses Verarbeitungskonzept auch in Primaten Anwendung findet, sowie in den meisten anderen Spezies, die über eine primäre Sehrinde verfügen. Seither geht man im allgemeinen davon aus, daß bei der Verarbeitung visueller Information vor allem solche orientierten Kontrastkanten zählen, und daß diffuse Flächen kaum eine Rolle spielen.

Neuere experimentelle Ergebnisse zeigen jedoch, daß es in der primären Sehrinde sowohl von Primaten als auch von Katzen Zellen gibt, die antworten, wenn sich ihr rezeptives Feld innerhalb einer homogenen Fläche befindet (Komatsu et al., 1996; Tani et al., 2003). Dies steht im Widerspruch zu der Annahme, daß homogene Flächen keinen Beitrag zur Verarbeitung visueller Information leisten.

Ausgehend von diesen Befunden stellt sich die Frage, wie die Erkennung homogener Flächen im Cortex funktioniert. Gewaltig et al. (2002) haben ein hierfür Modell vorgestellt, in dem homogene Flächen als Bildareale mit niedriger Varianz der Intensität definiert sind. Homogenitätselektive Zellen berechnen die Varianz der Intensität in ihrem rezeptiven Feld, und antworten wenn diese unter einem bestimmten Schwellwert liegt.

Die Annahme, daß Neuronen Varianz berechnen, ist zunächst nur schwer vorstellbar. Allerdings ändert sich diese Situation, wenn die Intensitäten in einen Aktionspotential-Latenzcode übersetzt werden. Hohe Varianz der Intensitäten ergibt Aktionspotentiale mit hoher zeitlicher Streuung, während bei niedriger Varianz die Aktionspotentiale innerhalb eines

kürzeren Zeitraums erzeugt werden. Dies kann von einem koinzidenzdetektierenden Neuron ausgewertet werden, das bevorzugt reagiert, wenn viele Aktionspotentiale innerhalb eines kurzen Zeitraums eintreffen. Diese Eigenschaft ist nach aktuellem Stand der Wissenschaft Neuronen inherent. Die Realisierung eines Modells für Homogenitätsdetektion aufgrund dieser biologisch realistischeren Annahmen steht im Mittelpunkt des praktischen Teils dieser Arbeit.

Weiterhin ist zu beantworten, welchen Nutzen das Gehirn aus der Information über homogene Flächen ziehen kann. Gewaltig et al. (2002) führen hierzu an, daß durch Unterdrückung von orientierungsselektiven Zellen innerhalb homogener Bildregionen eine Verbesserung der Bildrepräsentation im Cortex erreicht werden kann. Die Idee hierbei ist, daß sowohl orientierungsselektive als auch homogenitätsselektive Zellen keine perfekten Detektoren sind, sondern durch ihre komplementäre Funktionsweise im Zusammenspiel eine verbesserte, genauere Antwort liefern können. Die vorliegende Arbeit greift diese Idee auf und untersucht den Effekt, den Information über homogene Bildanteile auf die Verarbeitung visueller Information haben kann.

Darüber hinaus erlaubt das vorgestellte Modell eine sehr schnelle Verarbeitungsgeschwindigkeit. Thorpe et al. (2001, 1996) haben gezeigt, daß Primaten dazu fähig sind, Objekte ca. 150 ms nach Präsentation zu klassifizieren. Daraus folgt, daß jedes Neuron in der Verarbeitungskette maximal ein bis zwei Aktionspotentiale produzieren kann. Das hier vorgeschlagene Modell erfüllt diese Voraussetzung.

## Gliederungsübersicht

In der Einleitung wird zuerst eine Übersicht über das visuelle System von Primaten gegeben. Dabei wird genauer auf den koniozellulären Pfad eingegangen, da diesem eine besondere Rolle in dieser Arbeit zugeschrieben wird. Weiterhin wird die oben erwähnte Thorpe'sche Hypothese genauer dargelegt. Schließlich wird das oben genannte Modell von Gewaltig et al. vorgestellt, auf dem diese Arbeit basiert.

Im zweiten Teil werden die theoretischen Grundlagen der Modellierung dargelegt. Hierbei handelt es sich zum einen um das Neuronenmodell, das in den Simulationen verwendet wurde, und zum anderen um die Theorie

der Synfire Chains. Für diese Arbeit sind im besonderen die Konzepte der Puls-Packete und der Koinzidenzdetektion von Interesse.

Das eigentliche Modell wird im dritten Teil vorgestellt. Nach einer Übersicht wird das Arbeitsprinzip erläutert und danach auf die Implementierung eingegangen, sowie auf die Lösung zweier Detailfragen, die für das Verständnis der Funktionsweise des Modells wichtig sind.

Im vierten Teil wird gezeigt, was das Modell zu leisten im Stande ist, wenn natürliche Bilder damit analysiert werden. Ebenfalls wird gezeigt, welche Ergebnisse die Interaktion von homogenitäts- und orientierungsselektiven Zellen liefert.

Der fünfte Teil dient der Diskussion des Modells. Zuerst wird ein Vergleich mit dem vorhergehenden Modell, auf dem das vorgestellte basiert, gezeigt. Danach wird die biologische Realisierbarkeit des Modells kritisch untersucht. Schließlich werden noch mögliche Folgeexperimente zur Verfeinerung des Modells vorgeschlagen.

# Abstract

Beginning with the work of Hubel & Wiesel (1962), the representation of retinal input in V1 has mostly been thought of being dominated by luminance-contrast- or color-contrast-edges. However, Komatsu et al. (1996) show that some cells in V1 also respond to diffuse illumination of their receptive field, which conflicts with the classical view of receptive fields in V1.

Based on these findings, Gewaltig et al. (2002) presented a model for surface detection, which the present work is based on. Surfaces, or homogeneous image areas, are defined as areas with low intensity variance. While the model by Gewaltig et al. inherently uses a rate code, the model presented in this work uses a network of spiking neurons to detect homogeneous areas. Moreover, it uses only one spike per neuron and processing stage, thus fulfilling the timing constraints of rapid feed-forward processing, as postulated by Thorpe et al. (1996).

Given the evidence that information about homogeneous image regions is present in the visual cortex, the question arises how the cortex might achieve homogeneity detection, and which thalamic and cortical areas could be involved. The present work provides an hypothesis addressing this question. Further, it is demonstrated what benefits for visual processing can be drawn from information about homogeneous image areas.

# Contents


**1 Introduction** **1**

1.1 Overview of the visual system . . . . . 2
1.2 The role of the koniocellular pathway . . . . . 7
1.3 Processing speed in the visual system . . . . . 11
1.4 Homogeneity detection . . . . . 12
1.5 Scope of this work . . . . . 14

**2 Methods** **17**

2.1 The Lapicque neuron model . . . . . 17
2.2 Pulse packets and coincidence detection . . . . . 21

**3 The Model** **25**

3.1 Overview . . . . . 25
3.2 Working principle . . . . . 26
3.3 Implementation . . . . . 28
3.4 The variance threshold . . . . . 31
3.5 Brightness dependency . . . . . 32

**4 Results** **35**

4.1 Analyzing natural images . . . . . 35
4.2 Interaction with edge-detectors . . . . . 37
4.3 Effect of large retinal receptive fields . . . . . 39
4.4 Homogeneity-detector receptive field size . . . . . 41

**5 Discussion** **43**

5.1 Comparison with the rate code model . . . . . 43
5.2 Biological context . . . . . 44
5.3 Benefits of homogeneity detection . . . . . 48
5.4 Outlook . . . . . 49

**6 Appendix** **53**
A Implementation of the rate code model . . . . . 53
B Simulation environment . . . . . 54
C Parameters of the spike-based model . . . . . 55
D Used images . . . . . 56
E Some words on modeling . . . . . 57

**References** **64**

# List of Figures


1.1 Overview of the visual pathway . . . . . 2
1.2 Simplified overview of cells in the retina . . . . . 3
1.3 The lateral geniculate nucleus (LGN): schematic . . . . . 5
1.4 Information flow in primary visual cortex (V1) based on anatomical connections . . . . . 6
1.5 Telling homogeneous from non-homogeneous image areas by gray-level variance . . . . . 13

2.1 Input current vs. spike latency . . . . . 19
2.2 Spike response . . . . . 20
2.3 Obtaining neuron response by convolution . . . . . 22
2.4 Effect of $\mathbf{n_{spikes}}$ . . . . . 22
2.5 Effect of $\sigma_{\mathbf{t}}$ . . . . . 23
2.6 Effect of $\tau_{\mathbf{syn}}$ . . . . . 24
2.7 Integration window gets shorter when $\tau_{syn}$ gets shorter . . . . 24

3.1 Illustration of the hypothesis . . . . . 25
3.2 Architecture of one homogeneity-sensitive unit . . . . . 26
3.3 Working principle . . . . . 27
3.4 Response to increasing levels of gray-value variance . . . . . 31
3.5 Brightness dependency: Sketch of principle . . . . . 32
3.6 Brightness dependency: on- and off-detector . . . . . 33

4.1 Obtaining the homogeneity map . . . . . 36
4.2 Complementary results from On- and Off-detectors . . . . . 36
4.3 Inhibition of edge-responses by homogeneity responses: Principle . . . . . 37
4.4 Inhibition of edge-responses by homogeneity-responses: Whole image . . . . . 37

4.5 Another example . . . . . 38
4.6 Effect of retinal preprocessing . . . . . 39
4.7 Retinal preprocessing: principle . . . . . 40
4.8 Retinal preprocessing: dark and bright homogeneous stimuli . 41
4.9 Effect of homogeneity-detector receptive field size . . . . . 41

5.1 Comparison of results from rate code model and spike-based model . . . . . 43

1 Relative sizes of the pictures used in the simulations . . . . . 57

# Part 1: Introduction

## What we see

Humans, like primates in general, are heavily depending on their visual capabilities. *Vision* enables us to explore our environment without physically moving around and instead scan the scene from a safe place. It also enables us to react in situations, where a possible threat or reward can not be detected by our auditory, somatosensory or olfactory senses.

Imagine the following scenario: An early *homo sapiens* is on the hunt in the savannah. Suddenly, a stone's throw away, he *sees* an animal *moving*. In this situation, it is vital to *recognize* the moving object as prey, which can ensure his family's dinner, or as a hungry predator, which wants him for lunch. His other senses could not cope with this task, since he might not smell the difference in unfavorable wind conditions, or feel it unless it is already too late[1].

Most of the time, seeing is an unconscious process which happens "all by itself", without the need for us to think about it. This is rather astonishing, if one considers the effort necessary to construct a consistent view of the outside world, provided just the rain of photons onto the retina. Generations of neuroscientists have dedicated their research to unraveling the process of seeing, and most of them unraveled only bigger secrets. Nevertheless, today we have already gained a lot of knowledge about how the brain processes what the eye sees. This work might just add a small piece to the giant puzzle of how the brain works.

[1] Besides, the task of hunting itself would be very difficult if he were blind.

## 1.1 Overview of the visual system

In order to discuss neural processing in the primate visual system, it is essential to know its anatomy. Since the anatomic details of each stage in the visual system can fill books by themselves, this section can only serve as a quick reference. I mainly provide facts which are necessary in the scope of this work, especially in the *discussion* (part 5). Most of the facts can be found in more detail in any textbook on neuroscience, such as Nicholls et al. (2001) or Kandel et al. (2000). Especially Rodieck (1998) is a comprehensive book on the physiological and functional properties of the first stages in the visual system.

Figure 1.1 illustrates the flow of information in the visual system. Starting with the retina, visual information flows from there to the lateral geniculate nucleus (*LGN*), which relays it to the primary visual cortex, or V1. Higher processing stages are V2, and afterwards the various areas in the so-called ventral and dorsal pathways. In the scope of this work, the main interest lies on processing from the retina to V1, which I present in more detail in the following.

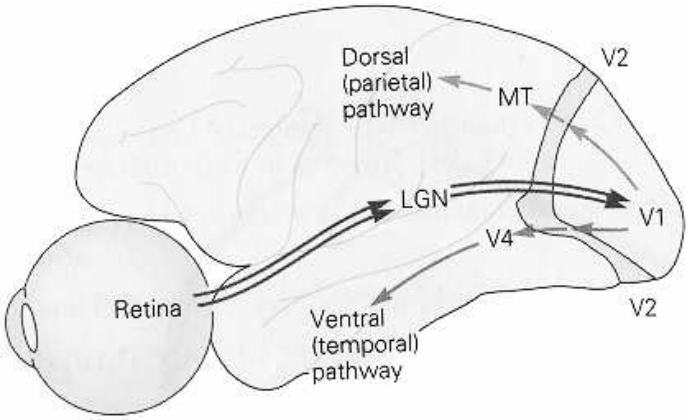


**Figure 1.1: Overview of the visual pathway.**
From Kandel et al. (2000).

### 1.1.1 Retina: transforming photons to spikes

Seeing begins with the retina. Here, light is transformed to spike patterns, and there is already some computation going on, such as the construction of center-surround receptive fields. As figure 1.2 illustrates, we find different types of cells in the retina, which I will describe shortly in the following.

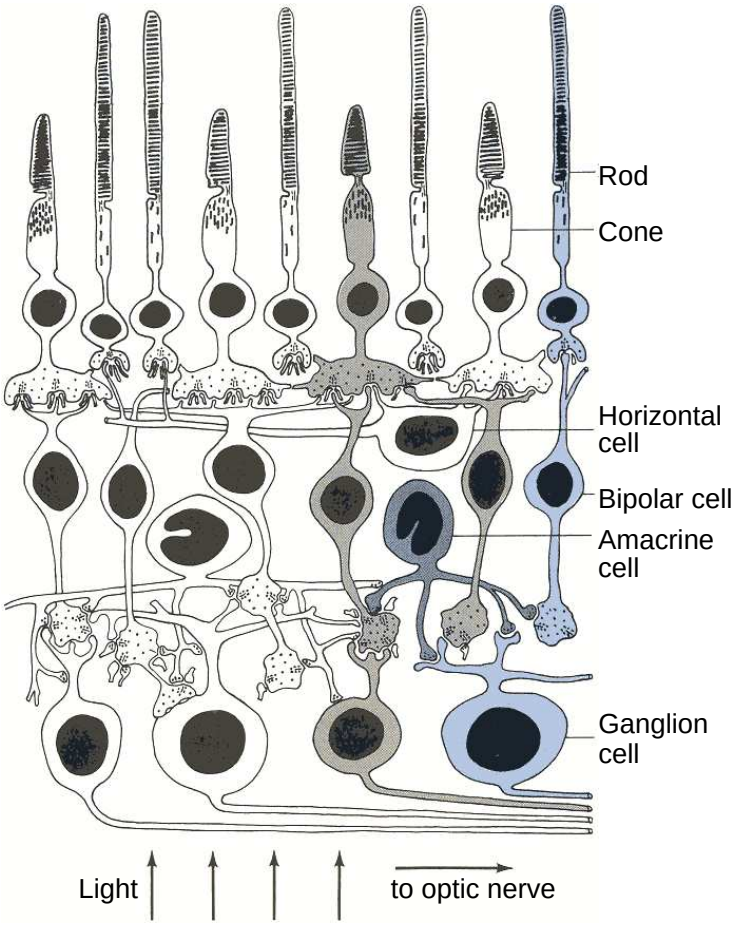


**Figure 1.2: Simplified overview of cells in the retina.** Adapted from Nicholls et al. (2001).

### Photoreceptors

Photons falling on the retina are being absorbed by *photoreceptors*, transformed into chemical energy, starting a signal cascade, which in the end leads to hyperpolarization of the photoreceptor membrane.

There are different types of photoreceptors: *rods*, which are very sensitive to light. They only contribute to vision under conditions where very few light is available.

If there is more light available, the *cones* come into play. They are less sensitive to light, but different wavelength sensitivity makes color vision possible. The three types of cones are *L-*, *M-* and *S-cones*, which are sensitive for light of long, medium and short wavelength, respectively. There are no S-cones in the *fovea*, the central part of the retina where visual acuity is highest.

#### Horizontal, bipolar and amacrine cells

Photoreceptors are laterally interconnected by *horizontal cells*. In first approximation, these cells sum up membrane potentials across several photoreceptors and feed the summed potential back on them. They contribute to constructing receptive field characteristics (see section 1.1.2). Nicholls et al. (2001, p. 399) provide a demonstration of the principle. *Bipolar cells* basically relay the membrane potential from photoreceptors and horizontal cells to the ganglion cells. *Amakrine cells*, which receive input from bipolar cells, send synapses back to the bipolar cells as well as on ganglion cells.

#### Ganglion cells

These are the first cells in the signal chain to actually produce spikes. Most of them show distinct center-surround receptive field properties—i.e. a stimulus in the center of the receptive field has a different effect on the activity of the neuron than the same stimulus would have in the surround.

### 1.1.2 The concept of receptive fields

The notion "Receptive Field" is a widely used term in neuroscience. However, neuroscientists in different research fields will associate slightly different meanings with this concept, which has often led to fierce discussions. In this thesis, the following definitions are applied:

**Classical receptive field (CRF):** The area on the retina from which the activity of a neuron can be influenced by light (Nicholls et al., 2001).

**On- and off-regions:** The classical receptive field of a neuron in V1 can be divided into a *center* and an antagonistic *surround*. In on-center cells, the optimal stimulus consists of a small spot of light (on) with a dark surround (off) (Hubel & Wiesel, 1962).

### 1.1.3 Lateral Geniculate Nucleus (LGN): Relaying the retinal signals to the brain

In the primate brain, retinal ganglion cells project their axons to a variety of brain regions. The most prominent among these is the *lateral geniculate*

*nucleus* (LGN). In the following, I summarize its most important aspects with respect to this work. Figure 1.3 illustrates the anatomy of the LGN.

In the primate brain, the LGN consists of 6 major layers of neurons. Layers 1, 3 and 5 get input from the ipsilateral eye, 2, 4 and 6 from the contralateral one. Cells in layers 6 – 3 are relatively small in size, and therefore these layers are called *parvocellular* (or P) layers (lat. *parvus* = small). Their receptive fields are small, with high selectivity for color-stimuli and low contrast-sensitivity, their responses are sustained. Parvo-cells mainly project to layer 4C$\beta$ in V1.

Layers 1 and 2 are called *magnocellular* (or M) layers, because the cells therein are relatively large (lat. *magnus* = large). Their receptive fields are rather large, with high contrast-sensitivity, albeit virtually no color-selectivity. Since these cells respond transiently, they are especially sensitive to moving stimuli. Their axons terminate mainly in layer 4C$\alpha$ in V1.

Intercalated between these 6 layers lie the *koniocellular* (or K) layers (greek *konis* = dust). This phylogenetically old class of cells, rather inhomogeneous in itself, is physiochemically distinct from magno- and parvo-cells, and has comparably large receptive fields. Most important for this work is that the receptive fields of some konio-cells have no antagonistic surround, thus they can respond to diffuse illumination (Rodieck, 1998). Another specialty of K-cells is that their axons project directly to layers 2 and 3 of V1, in contrast to M and P cells, which project to layer 4. Moreover, while M and P cells project exclusively to V1, K cell inputs have been identified in higher areas of the ventral visual pathway (see Hernandez-Gonzales et al., 1994).

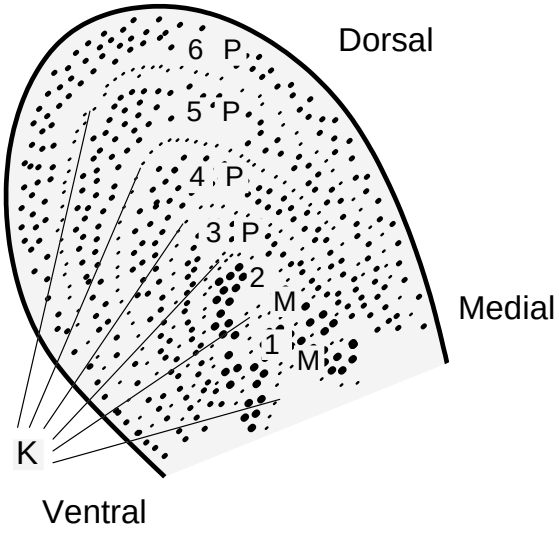


**Figure 1.3: The lateral geniculate nucleus (LGN): schematic.** After Nicholls et al. (2001). P cells form the upper 4 layers, M cells the lower 2. In between the layers, the K cells are situated.

### 1.1.4 V1: The first cortical processing area

V1, as all cortical areas, has a layered structure, however with a different functional organization. Most incoming axons from the M- and P-layers in the LGN terminate in layer 4 (see above). From there, neurons project into layers 2 and 3, and from there connections to higher areals in the visual system, such as V4 and MT, are made. See figure 1.4 for a schematic overview.

Neurons in layer 4C mainly respond to oriented contrast edges. Depending on their receptive field characteristics (size, color- and contrast-sensitivity), cells in layer 4C can be characterized as belonging to the P- or M-channel. Cells which preferably respond to luminance contrast edges are assigned to the M-channel, while color-contrast-sensitive cells belong to the P-channel, in analogy to the terminology in the LGN.

Besides the anatomically distinct layers, V1 can be functionally divided into *columns*, perpendicular to the layers. Neurons inside one such column share their preferred stimulus, such as the orientation of a contrast edge.

Another fundamental characteristic of V1 is that the spatial arrangement of stimuli on the retina is preserved in V1. In other words, two neighboring stimuli on the retina excite neighboring cells in V1. This feature is called *retinotopy*, and V1 is therefore said to contain a *retinotopic map*. A similar kind of spatial mapping is observed in primary somatosensory cortex (*somatotopy*), or in primary auditory cortex (*tonotopy*).

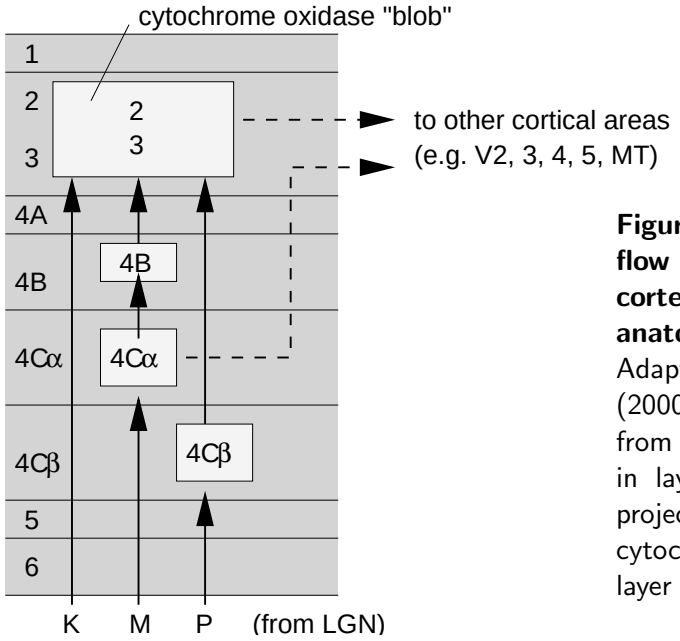


**Figure 1.4: Information flow in primary visual cortex (V1) based on anatomical connections.** Adapted from Kandel et al. (2000). Most afferent fibers from the LGN terminate in layer 4, however K cells project their axons to the cytochrome oxidase blobs in layer 2/3.

#### Enhanced orientation tuning in Layer 2/3 of V1

In layers 2 and 3 of V1, the tuning of orientation-selective cells improves significantly, i.e. the orientation range the cells respond to is smaller than in layer 4. It has been speculated that this improvement may be achieved by lateral connections between neurons responding to neighboring colinearly oriented contrast edges (Li & Gilbert, 2002). However, as this work shows, there are also other putative mechanisms which can explain the enhanced tuning.

## 1.2 The role of the koniocellular pathway

### 1.2.1 Overview of the M- P- and K-pathway

Electrophysiology and psychophysics tell us that form, motion and color are processed in parallel in the visual system (Nicholls et al., 2001). At first sight, it is tempting to think that the three classes of cells in the LGN are the neuronal substrate for the independent processing of different visual information. If we look at the characteristics of each of the M, P and K channels, or *pathways*, we find some interesting correlations to the psychophysically claimed pathways:

The *P-channel* is color-sensitive and could therefore account for the color-pathway; the *M-channel*, especially sensitive to moving stimuli, may state the motion pathway; both could account for form.

There is yet little known about the *K-channel*. Formerly said to have a rather modulatory role, recent experimental results suggest a more fundamental role for the K-channel in vision (see section 1.2.2 below).

In the LGN, the M, P and K channels are rather separated, without lateral interconnections. This changes in V1: looking at the cytochrome oxidase (CO) blobs in layers 2 and 3, we see all three channels intertwined. Here, the channels seem to affect each other's processing.

Further, cells in each of the M, P and K channels also respond to stimuli which psychophysics would classify in another pathway. So one must take care to distinguish the *cellular* M-, P- and K-classification from the *functional* channel classification for form, motion and colour. When I refer to the K-pathway or the K-channel in this work, I am using the cellular classification.

### 1.2.2 The koniocellular pathway in more detail

The K pathway plays a special role in this work. Unfortunately, we are just beginning to understand the purpose of the K-Pathway; research up to now mainly focused on the more popular M- and P-pathways. On account of that I will summarize some of the recent research on the K-Path (adapted from Körner, 2001).

#### Input to the LGN

Input to the K-cells in the LGN is segregated from the M- and P-cells, in such that LGN K-cells get input only from those retinal ganglion cells which do *not* project to M- and P-layers of the LGN. As M- and P-projecting ganglion cells constitute about 90 % of all ganglion cells, K-cell input can arise from less than 10 % of the ganglion cells, all with quite large receptive fields. These 10 % contain at least 15 – 18 different types of ganglion cells, most of them "phylogenetically old" (Rodieck, 1998). Unfortunately, it is very difficult to record from these cells: their rareness makes them hard to find, and they have small cell bodies.

Among these ancient cells are the bistratified ganglion cells, which mainly relay the information from S-cones (blue-on) to the middle pair of K-cell layers in the LGN. S-cones are highly conserved across mammalian species, and therefore considered as a phylogenetically ancient color system (Silveira et al., 1999). In human and monkey retina, S-cones represent 5 – 10 % of the cone mosaic and distribute in a quasi-regular fashion over most of the retina with the exception of the central fovea (Calkins, 2001). As the tonically firing bistratified ganglion cells compare the input from one to three S-cones with input from 20 to 30 other cones, their output conveys very rough information about wavelength and intensity in their large receptive fields (Rodieck, 1998).

#### In the LGN

The six koniocellular layers in the LGN seem to be functionally distinct. As the M- and P-layers, K-layers group in pairs receiving input from either the ipsi- or contralateral eye. In macaques and owl monkeys, the middle pair of K-layers relays information from small bistratified ganglion cells to

the CO blobs in V1, whereas the the most dorsal pair relays low-acuity visual information (from nondefined ganglion cells) to layer 1 in V1 (Ding & Casagrande, 1998).

### From the LGN to V1

Inside the CO-blobs, K-cell axons form quite large excitatory terminals, which synapse closer to the pyramidal cell bodies than layer 4 axons do (Ding & Casagrande, 1998). Thus, inside the blobs, K-input may be more effective for pyramidal cells than intracortical input.

### Latencies in the K-pathway

There is only few information available about latencies to visual stimulation in the koniocellular pathway. The small size of K-cells in the LGN suggests that they have small, slowly conducting axons. Indeed, for the prosimian bush baby (*Galago crassicaudatus*), relatively long latencies of about 80 ms to the onset of a visual stimulus have been reported (Irvin et al., 1986). However, the retina of bush babies has only a single type of M/L-cones and no S-cones at all (Jacobs et al., 1996). Thus, there is no S-cone input to the LGN K-layers (via the bistratified ganglion cells) as in macaques and humans, and the delay times may not be comparable.

It has been speculated that the K-pathway may be homologous to the slowly conducting W-pathway in cats, but generalising from cats to other mammals may be inappropriate. For instance, W-like cells in rats have latencies of about 3.18 ms after stimulation at the optic chiasm, while the fastest responding Y-cell responded after 1.98 ms (Fukada et al., 1979). This small difference would not argue for a very slow propagation.

Interestingly, in humans, electric field potentials evoked in V1 by S-cone isolating stimuli appear earlier in time, namely with latencies of about 40 ms, than the common luminance-defined motion-specific potentials. This indicates a very fast activation of V1 specific to K-pathway input (Morand et al., 2000).

### Possible functions of the K-pathway

The CO-blobs in layer 2/3 of V1 have often been linked to color vision (Nicholls et al., 2001). Because of their projections to those blobs, and be-

cause of the input they get from different cone types, K-cells have been proposed to contribute to color vision (Martin et al., 1997). However, the owl monkey and the bushbaby (both are nocturnal primates), have only one type of cone, and therefore lack color vision (Jacobs et al., 1996). Surprisingly, these species *do* posses CO-blobs in V1, suggesting that blobs are not exclusively involved in color vision.

Moreover, lower mammals like rats *do* have color vision, but do not possess CO-blobs. Phylogenetically, blobs are a relatively new structure. We may speculate that CO-blobs have become necessary in evolution as retinal receptive fields have become smaller to achieve higher vision acuity. Because of the fine spatial resolution of the information conveyed by the P-system, a system providing more coarse information about the extent of objects and surfaces may be of use in order to "keep the overview" and make *segmentation* of the image into separate objects possible. Because of their comparably large receptive fields, K-cells could transmit such information to the CO-blobs.

The fact that color processing seems to play an important role in the CO-blobs of color-seeing primates is well in congruence with this assumption, since it also has been suggested that color can support object segmentation in early vision (Gegenfurtner & Rieger, 2000).

#### Summary

The growing interest in the K-pathway has already led to some interesting conclusions. However, there is no common theory about its purpose or function. Moreover, many researchers do not even try to propose a function for it. Certainly, this is at least partly due to its inhomogeneous nature, which may make it inappropriate to speak of *the* K-system. Regarding its ubiquitous presence in primate visual systems, it surely *does* have an important function; otherwise it would not have made it through the evolutionary optimization process.

## 1.3 Processing speed in the visual system

### 1.3.1 Bottom-up versus top-down processing

The term *bottom-up processing* in the visual system describes the information flow from the "lower" parts of the visual system (e.g. the retina) to higher ones (e.g. V1). The term *top-down* describes the opposite direction: from higher vision areas (e.g. IT) to lower ones (as V1).

Ullman (1995) proposed that object recognition works in a combined bottom-up/top-down manner, because pure bottom-up processing is not believed to be robust enough against noise and distractors. Other proposed mechanisms to stabilize a noisy percept (e.g. the enhanced orientation tuning in layer 2/3 in V1) involve lateral connections between neurons responding to neighboring colinear edge fragments (Li & Gilbert, 2002; Ullman, 1995). Ullman (1995, chapter 8) provides a technically oriented overview on more mechanisms. These mechanisms work fine. However, iterative processing as well as lateral interaction requires *time*, which runs contrary to fast processing.

### 1.3.2 Timing constraints from psychophysics

Thorpe & Imbert (1989) (cited in Thorpe et al., 1996) pointed out that primates can reliably classify objects in an image approx. 150 ms after presentation (see also Fabre-Thorpe et al., 2001), using only 10-15 ms per processing stage. That is, each neuron in the processing chain has time to fire at most 1 or 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 other words, processing must be performed in an essentially *feed-forward* way. Moreover, given only one spike per neuron, it is impossible to use some kind of rate code to transmit information. Consequently, mechanisms for object recognition which involve lateral or top-down interaction are inherently *too slow* to account for the rapid object classification capabilities of primates.

To overcome the caveats of rate-coding, other coding strategies have been proposed, involving spike latency or spike timing (see Thorpe et al., 2001). It has also been shown that such coding strategies, which use only

one spike per neuron, can be very effective, both in terms of time as in terms of the amount of information transmitted (Van Rullen & Thorpe, 2001).

For such temporal coding stategies to work, it is necessary that neurons are capable of producing their action potentials with high temporal accuracy. There is an ongoing debate on the variability of spike timing (see Gewaltig (2000, chapter 1.3) for an overview), with strong evidence that neurons actually are capable of producing spikes with a temporal accuracy of 1 ms. I do not want to go into detail here, but instead refer to the discussion of spike timing accuracy in part 5.

## 1.4 Homogeneity detection

### 1.4.1 Beyond Hubel and Wiesel: Homogeneous receptive fields

As said above, the representation of retinal input in V1 is mostly dominated by luminance-contrast- or color-contrast-edges, while diffuse illumination does not elicit a response (Hubel & Wiesel, 1962). However, recent experimental results suggest that this is not always true. Komatsu et al. (1996) 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; moreover, the receptive fields of these cells do not have an antagonistic surround. Similarly, Friedman et al. (2003) found that 20% of the neurons in V1 and V2 respond to uniform color and show no edge- or orientation-selectivity. Tani et al. (2003) showed that uniform surfaces are also represented in cat visual cortex. These findings conflict with the classical center-surround view of receptive field properties in V1.

### 1.4.2 A rate code model for homogeneity detection

Based on the above findings, Gewaltig et al. (2002) presented a computational model for surface detection, on which the present work is based on. Surfaces are herein defined as homogeneous image areas, with homogeneity being defined as low gray-level variance. Figure 1.5 illustrates this.

The authors discuss how surface information can be used to support and enhance the response properties of orientation-selective cells in V1. Homo-

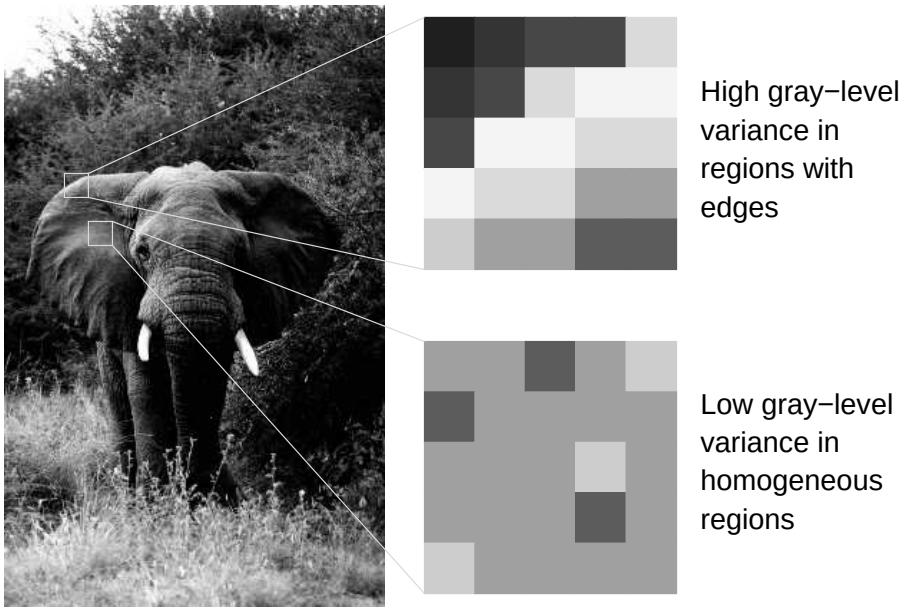


**Figure 1.5: Telling homogeneous from non-homogeneous image areas by gray-level variance.**

geneity and "edgy-ness" (oriented contrast edges) are mutually exclusive at any image point (*"Where there is a (homogeneous) surface, there is no edge"*). By inhibiting responses from orientation-selective cells in homogeneous areas, it should be possible to enhance the response properties of orientation-selective cells in V1, reducing responses to spurious edges. In particular, it is proposed that Konio-cells in the LGN contribute to the improved orientation tuning of orientation-selective cells in layer 2/3 of V1 compared to the tuning in layer 4.

### Working principle

The is divided into three parts: a *retina*, which is equivalent to the input image, a layer of *LGN K cells*, with large receptive fields without antagonistic surround, and a layer of *homogeneity-detectors* calculating the variance of intensity in their receptive fields.

LGN K cells respond gradually to illumination of their receptive fields. The effect of large receptive fields is mimicked by spatially smoothing the retinal image (in more technical terms, the original image is low-pass filtered). Additionally, a sigmoid activation function is applied pointwise to model the neuron's response properties.

Each homogeneity-detector cell evaluates the input it gets from a fixed

number of LGN K cells. Effectively, it calculates gray-value variance inside its receptive field. Low gray-value variance indicates that the neuron's receptive field is inside an homogeneous image area.

Since the result of the variance calculation is continuous-valued, it can be represented best by a rate code. Therefore, I will refer to this model as *rate code model*. A more formal description of the mechanism is given in appendix A.

Edge-detectors were not part of the former work. Instead, the main focus was on demonstrating the effect of homogeneity information on the *input* to orientation-selective cells.

## 1.5 Scope of this work

In the rate code model, it is assumed that homogeneity-detecting neurons are able to estimate gray-value variance in their receptive field. In the context of rate code neurons, this assumption is quite artificial. However, Gewaltig et al. (2002) point out that the situation may change when introducing a spike-based latency code, in such a way that a distribution of input densities will result in a similar distribution of spike latencies, which can be evaluated by a coincidence-detecting neuron.

The realization of such a network of spiking neurons performing homogeneity detection dominates the practical part of this work (see part 3, "The Model"). In contrast to the rate code based model, this is achieved by using a spike-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 one spike per layer, thus meeting the timing constraint set up by Thorpe & Imbert (1989).

In part 4 ("Results"), I show how the model performs with natural images as input. I provide a comparison between the results from the rate code model and the spike-based model in the discussion in part 5.

Further, given the evidence for homogeneity detection in primate visual cortex, two questions arise:

*How?* — Which thalamic and cortical structures could be involved in homogeneity detection?

*Why?* — What benefits can be drawn from homogeneity information?

The first question is addressed by the hypothesis provided at the beginning of part 3, extending the hypothesis given by Gewaltig et al. (2002), and discussed in part 5. This is also where I address the second question: I show what benefits information about homogeneous image regions can have in cortical processing of visual input. In addition, I lay out possible experiments to test the hypothesis.

Preliminary results from this work have been presented in poster form at the Perception Conference at Tübingen (Schmuker et al., 2003a) and at the Göttingen Neurobiology Conference (Schmuker et al., 2003b).

# Part 2: Methods

This part outlines the theoretical framework for the present work. First, I will provide a formal discussion of the neuron model used in the simulations. Then, I shortly present the concept of pulse packets, which emerged from synfire chain theory. On this theoretical background, I will present the concept of coincidence detection, an inherent feature of neurons, and what parameters of the model neuron affect this property.

## 2.1 The Lapicque neuron model

This work uses the *Lapicque* neuron model (Lapicque, 1907), with the extension of so-called $\alpha$-shape post-synaptic currents ($\alpha$-PSCs, see below). I provide some analytic descriptions for the Lapicque model neuron which are relevant for this work. The focus shall be on exposing relevant parameters instead of giving a complete mathematical analysis, since this has already been done elsewhere. Most of this section is based on the book by Tuckwell (1988), where a more extensive analytical and functional description is available.

### 2.1.1 Response to currents and single spikes

#### Subthreshold response to injected current

The subthreshold response to injected current $I(t)$ is described by the first order linear differential equation

$$C\frac{dV}{dt} + \frac{V}{R} = I(t)\,,\ t > 0\,. \tag{2.1}$$

$C$ is the membrane capacitance, $R$ is the membrane resistance (both assumed constant), $V$ is the membrane potential.

The solution to this differential equation is

$$V(t) = V_0 + e^{-\frac{t}{RC}} \int_0^t \frac{I(t')}{C} e^{\frac{t'}{RC}} \, dt' \ . \tag{2.2}$$

$V_0$ is the membrane resting potential. In general, nerve cells have resting potentials in the order of $-70 \ mV$. Since the value of $V_0$ is not significant for the formal description of the model, it will for simplicity be omitted in the following equations. This is only allowed because the model neuron is linear in its aubthreshold behavior.

The product of $R \cdot C$ shall in the following be referred to as $\tau$, the membrane's *time constant*:

$$R \cdot C = \tau \ . \tag{2.3}$$

### Threshold

As the membrane potential reaches a fixed threshold $\theta$, it produces an action potential, or *spike*. This is why the Lapicque neuron is often referred to as *integrate-and-fire* (IAF) model.

After having produced a spike, the neuron enters a *refractory period* of length $t_R$, during which it cannot produce another spike. Let the sequence of times at which spikes are produced be $\{t_i, \ i = 1, 2, \ldots\}$. Then

$$\theta(t) = \begin{cases} \infty, & t_i < t < t_i + t_R \\ \theta, & \text{otherwise.} \end{cases} \tag{2.4}$$

Further, during the refractory period, the membrane potential is set to $V_0$:

$$V(t) = V_0, \ t_i < t < t_i + t_R \ . \tag{2.5}$$

### Time to threshold: Spike latency

If the current input $I(t)$ is constant and maintained indefinitely, then as $t \rightarrow \infty$, the membrane potential reaches the *steady-state value* $IR$, given that it does not exceed threshold:

$$V_{steadystate} = I \cdot R \ . \tag{2.6}$$

However, if threshold is reached a spike will be produced. The critical

input current $I_{crit}$ can easily be obtained by setting (2.6) equal the threshold $\theta$. This yields

$$I_{crit} = \frac{\theta}{R} \; . \tag{2.7}$$

We call the time at which the spike is produced the *spike latency*, relative to the time $t_0$ when the current is switched on. To calculate the spike latency, I first introduce the *Heaviside step function*:

$$H(t) = \begin{cases} 0, & t < t_0 \\ 1, & t \geq t_0 \end{cases} \; . \tag{2.8}$$

Some constant current $I$ switched on at $t = t_0$ can then be described as

$$I(t) = I \cdot H(t - t_0) \; . \tag{2.9}$$

Now let $t_0 = 0$. Inserting (2.9) into (2.2) yields

$$V(t) = IR \cdot \left(1 - e^{\frac{-t}{\tau}}\right) \; , t \geq 0 \; . \tag{2.10}$$

Given that the input current $I$ is large enough to drive the membrane potential to threshold, we can calculate spike latency as a function of $I$. So let $V(t) = \theta$ and solve for $t$:

$$\begin{aligned} \theta &= I \cdot R \left(1 - e^{\frac{-t}{\tau}}\right) \\ \text{finally yields} \quad t &= -\tau \cdot ln \left(1 - \frac{\theta}{IR}\right) \end{aligned} \tag{2.11}$$

with $t$ the spike latency, $\tau$ the membrane time constant ($R \cdot C$), $\theta$ the threshold value for spike generation, and $I$ a constant input current. Figure 2.1 illustrates eq. (2.11). Note that as $I \rightarrow I_{crit}^{+} \Rightarrow t_{spike} \rightarrow \infty$.

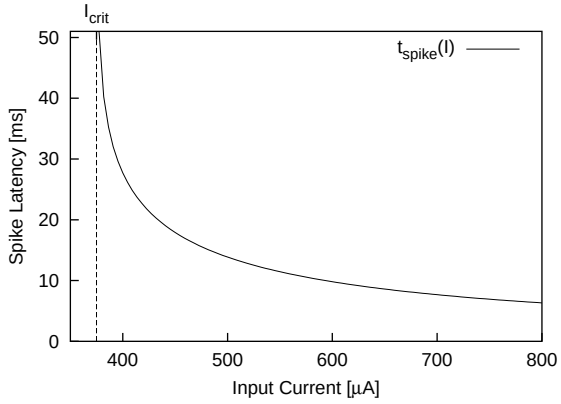


**Figure 2.1: Input current vs. spike latency.** Spike latency decreases in an inverse logarithmic manner with increasing input current (see eq. (2.11)). Neuron parameters: $\tau = 10ms$, C = $250pF$, $\theta = 15mV$. $I_{crit}$ denotes the critical input current below which no spike is triggered. The value is $375\mu A$, as calculated with eq. (2.7).

### Spike response

When a spike arrives at the post-synaptic neuron, it elicits a post-synaptic current (*PSC*) through the membrane. We model the PSC using a so-called $\alpha$-function PSC (Jack et al., 1985):

$$I_{PSC}(t) = kt \cdot e^{-\alpha t} \quad , t \geq 0,\, \alpha > 0\ , \tag{2.12}$$

with $k$ being an amplitude factor, $\alpha = \frac{1}{\tau_{syn}}$, where $\tau_{syn}$ is the time-constant of the synaptic current, and $t_0$ the time at which the spike arrives. The total charge delivered to the neuron depends on $\tau_{syn}$:

$$k/\alpha^2 \text{ resp. } k \cdot \tau_{syn}. \tag{2.13}$$

Inserting (2.12) into (2.2) yields the course of the post-synaptic membrane potential after an incoming spike:

$$V(t) = \frac{k}{C} e^{\frac{-t}{\tau}} \int_0^t t' e^{\left(\frac{-1}{\tau} - \alpha\right)t'} dt'\ . \tag{2.14}$$

We integrate (2.14) by substituting $\beta = \frac{1}{\tau} - \alpha$ , and obtain the post-synaptic potential *(PSP)* in response to a spike

$$PSP(t) = \begin{cases} \dfrac{k \cdot e^{-t/\tau}}{\beta C} \left( t \cdot e^{\beta t} - \dfrac{e^{\beta t} - 1}{\beta} \right) & , \beta \neq 0 \\ \dfrac{kt^2}{2C}\, e^{-t/\tau} & , \beta = 0\ . \end{cases} \tag{2.15}$$

In this work, $\tau$ is never equal to $\tau_{syn}$, so $\beta$ is never zero. Thus, only the first case is relevant. Figure 2.2 shows how the synaptic current precedes the resulting membrane potential in response to one input spike.

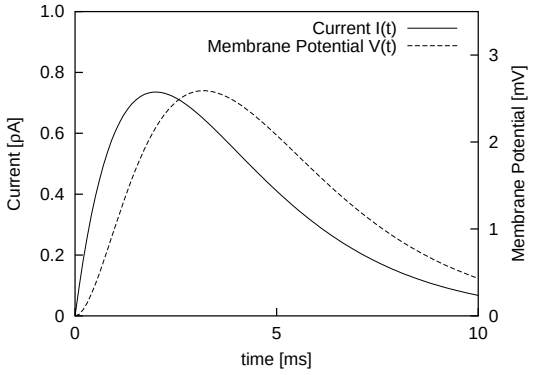


**Figure 2.2: Spike response.** The input spike arrives at t=0. Solid curve: the synaptic current, dashed curve: the resulting membrane potential. Note that scales are different for current and potential; current is depicted on the left, potential on the right hand side. Neuron parameters were $V_0 = 0$V, $k = 1$, $C = 250$pF, $\tau = 10$ms, $\tau_{syn} = 2$ms.

### 2.1.2 Impulse response and linear systems theory

In linear systems theory, an *impulse* is a function which is zero everywhere except in one singular point, where it is infinite, in such a way that the integral from $-\infty$ to $+\infty$ is 1.[1] If the impulse response $h(t)$ of a linear system is known, one can compute the response $y(t)$ of this system to an input signal $x(t)$ by *convolution* of the input with the system's impulse response (see also Papoulis (1977)):

$$y(t) = x(t) * h(t) = \int_{-\infty}^{+\infty} x(t-\tau)\, h(\tau)\, d\tau \tag{2.16}$$

A system $f$ can be called linear if and only if

$$f(a \cdot x) \stackrel{!}{=} a \cdot f(x) \quad \text{and} \quad f(x+y) \stackrel{!}{=} f(x) + f(y). \tag{2.17}$$

The differential equation (2.1) describing the subthreshold membrane voltage response to current is linear and therefore fulfills this condition. This means that we can consider the model neuron as a linear system (except for the threshold). Further, if we consider an incoming spike as an incoming *impulse* in the sense of linear systems theory, it follows that the neuron's response to a single spike can be interpreted as the neuron's impulse response. This allows us to compute the subthreshold neuron's response for any input.

## 2.2 Pulse packets and coincidence detection

### 2.2.1 Synfire chains

Moshe Abeles postulated the theory of *synfire chains* (Abeles, 1982a, 1991). These are groups of neurons which forward synchronous spike volleys depending on the width and amplitude of the packet, thus forming a functional cell assembly. Abeles et al. (1993) show that such highly synchronized discharge patterns of groups of neurons can actually be observed in the brain of behaving monkeys. Gewaltig (2000) and Diesmann (2002) provide a thorough analysis of properties and critical parameters of synfire chains.

[1] The so-called *Dirac delta function* fulfills this condition, and is therefore often used in theoretical analysis of linear systems.

### 2.2.2 Response to pulse packets

Diesmann et al. (1996, 1999) presented the concept of *pulse packets* to describe synchronized spike volleys. Using a stochastic approach, pulse packets can be characterized as Gaussian densities with two parameters, namely *standard deviation of spike times* $\sigma_t$ and the *number of spikes* $n_{spikes}$ in the pulse packet (Diesmann et al., 1996; Gewaltig, 2000). One can obtain the neuronal response to such an input density by *convolving* the input density with the impulse response. This is illustrated in figure 2.3.

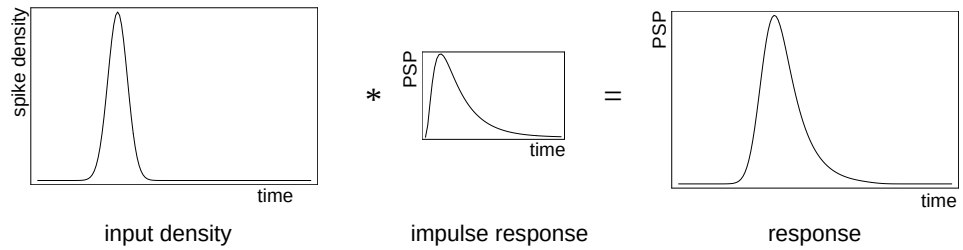


**Figure 2.3: Obtaining neuron response by convolution.** The input density is convolved with the system's impulse response. This yields the neuron's response to the given input density.

Decreasing the number of spikes $n_{spikes}$ leads to a smaller maximum PSP. Changing $n_{spikes}$ means changing the area (or *integral*) under the density curve (figure 2.4).

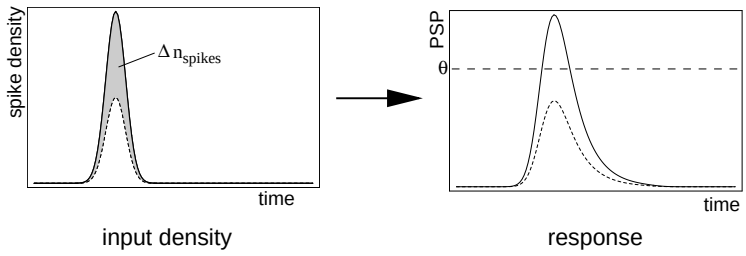


**Figure 2.4: Effect of $\mathbf{n_{spikes}}$.** *Left:* Two input densities differing in $n_{spikes}$. The difference is in the *area* (or *integral*) under the density curves (gray area). *Right:* The response evoked by the input density with smaller $n_{spikes}$ (dotted line) is smaller in amplitude and fails to reach threshold $\theta$.

A decrease in PSP amplitude can also be obtained by increasing spike time standard deviation $\sigma_T$ while keeping the spike number $n_{spikes}$ constant (see figure 2.5).

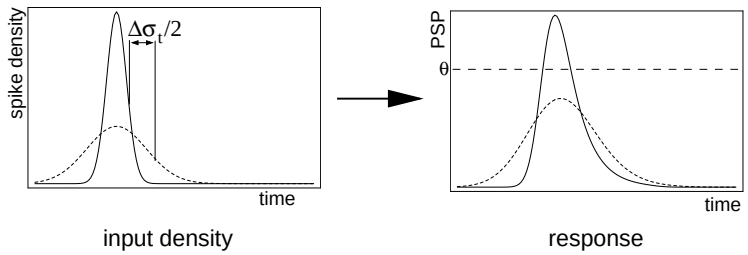


**Figure 2.5: Effect of $\sigma_t$.** *Left:* Two input densities differing in $\sigma_t$. *Right:* The input density with larger $\sigma_t$ (dotted line) evokes a PSP smaller in amplitude, failing to reach the threshold $\theta$, but wider in time.

### 2.2.3 Coincidence detection

The fact that increasing $\sigma_t$ leads to a smaller maximum PSP implies that the same number of spikes is more likely to trigger an action potential in the target neuron if the spikes arrive almost synchronously than if they arrive in a more isolated fashion. This has already been argued by Abeles (1982a). He also concluded that a cortical neuron effectively operates as a *coincidence-detector* rather than as an integrator (Abeles, 1982b).

From this follows that, if $n_{spikes}$ is fixed, there is a threshold in the degree of synchronization of the incoming spikes, or *coincidence threshold*, below which they will fail to elicit a spike in the target neuron. The degree of synchronization could be expressed either in terms of pulse packet width $\sigma_t$, or, more qualitatively, in terms of interspike intervals (*ISI*), i.e. the time between the arrival of two successive spikes.

#### Adjusting the "coincidence threshold"

The coincidence threshold depends on the *integration window*, i.e. the maximum time interval in which two spikes have to arrive in order to interact, i.e. sum up in terms of the resulting PSP. In the model neuron, the width of this integration window can effectively be adjusted using $\tau_{syn}$, the time-constant of the synaptic current (eq. (2.12)).

Figure 2.6 (next page) shows that smaller values of $\tau_{syn}$ lead to shorter PSPs. Further, according to eq. (2.13), less total charge is delivered to the neuron, which in turn would lead to a smaller PSP amplitude. In the figure, current amplitude has been adjusted to yield the same maximum potential.

If we look at the PSP elicited by two subsequently arriving spikes, its maximum amplitude can be taken as a qualitative measure for the width of

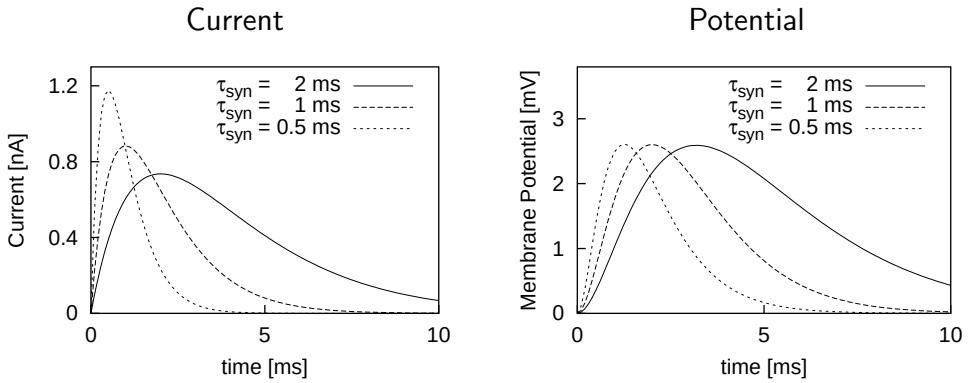


**Figure 2.6: Effect of $\tau_{\mathbf{syn}}$.** Smaller $\tau_{syn}$ lead to shorter PSPs. *Left:* Postsynaptic current course for three $\tau_{syn}$. *Right:* Resulting PSPs.

the integration window. Figure 2.7 shows that for smaller $\tau_{syn}$, two spikes with the same interspike interval elicit a smaller maximum PSP.

In other words, the integration window gets smaller with smaller values of $\tau_{syn}$. In turn, lower settings of $\tau_{syn}$ require smaller interspike intervals to drive the detector neuron to threshold, therefore "sharpening" its tuning to coincident spikes.

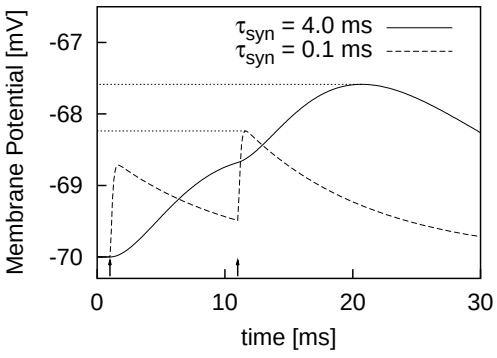


**Figure 2.7: Integration window gets shorter when $\tau_{syn}$ gets shorter.** PSP for $\tau_{syn} = 4.0$ ms (solid curve) and $\tau_{syn} = 0.1$ ms (dashed curve). Dotted lines: maximum PSP. Small arrows near the abscissa depict incoming spikes. Interspike interval is 10 ms. Data was obtained using the simulator (see B.1), with $V_0 = -70mV$, in contrast to the data shown in figure 2.6, which was obtained analytically ($V_0 = 0mV$).

For the simulations in this study, a value of 0.63 ms for $\tau_{syn}$ yields the best results (see also table 2 in appendix C.2). It is difficult to find biologically realistic values for the synaptic time constant in the literature. Smith & Sherman (2002) use a value of 1 ms for $\tau_{syn}$ in their biologically oriented simulations of thalamic neurons, which is in the same order of magnitude as the value used here. However, it must be said the neuron model used in that study is conductance-based, and the equation which models synaptic input is slightly different from the one used in this work. Nevertheless, the parameters are comparable on a qualitative basis.

# Part 3: The Model

## 3.1 Overview

The model concept is based on the work of Gewaltig et al. (2002) (see section 1.4.2). It realizes the following hypothesis (depicted by fig. 3.1):

Homogeneity detection is performed in the K-Path. Konio-cells in the LGN, with large receptive fields and no antagonistic surround, project to coincidence detecting inhibitory interneurons in the cytochrome-oxidase (CO) blobs in layer 2/3 of V1. These interneurons respond to homogeneous illumination of their receptive fields. Contrast-edge responding M- and P-cells in layer 4, with smaller receptive fields, also project to the CO-blobs, and from there to higher processing areas. The interneurons in the blobs can block the relay of edge-information in homogeneous image regions, thus reducing the amount of edge-information to the most salient compounds, allowing faster analysis of the scene by higher processing areas.

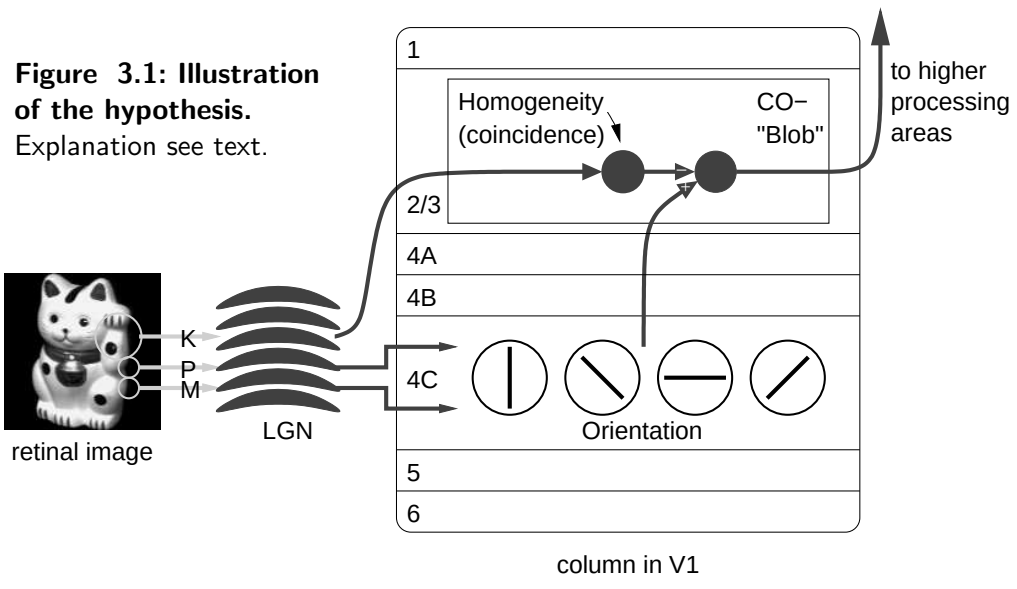


**Figure 3.1: Illustration of the hypothesis.** Explanation see text.

In contrast to the previous work, which implicitely uses a continuous rate code, the present work uses a network of *spiking neurons* to perform homogeneity detection. Section 3.2 explains the working principle: By translating gray-value variance into spike-timing variance, the homogeneity-sensitive neuron is realized by a coincidence-detecting neuron, allowing a more realistic way of calculating gray-value variance.

Section 3.3 gives an overview on the implementation of the model.

The *variance threshold*, i.e. the amount of gray-value variance below which the detector shall spike, is difficult to tackle analytically. I use an empirical approach, which I present in section 3.4.

Further, as I explain in section 3.5, the variance threshold also depends on the mean gray-value, i.e. the *brightness* of the stimulus in the receptive field. In order to reduce this effect, *on-* and *off*-detectors are introduced, sensitive to bright, respective dark homogeneous regions.

## 3.2 Working principle

The basic unit of this network consists of three parts (see fig. 3.2): a retina (fig. 3.2A), a layer of konio cells in LGN (fig. 3.2B) and a K-driven coincidence detecting neuron in V1 (fig. 3.2C)

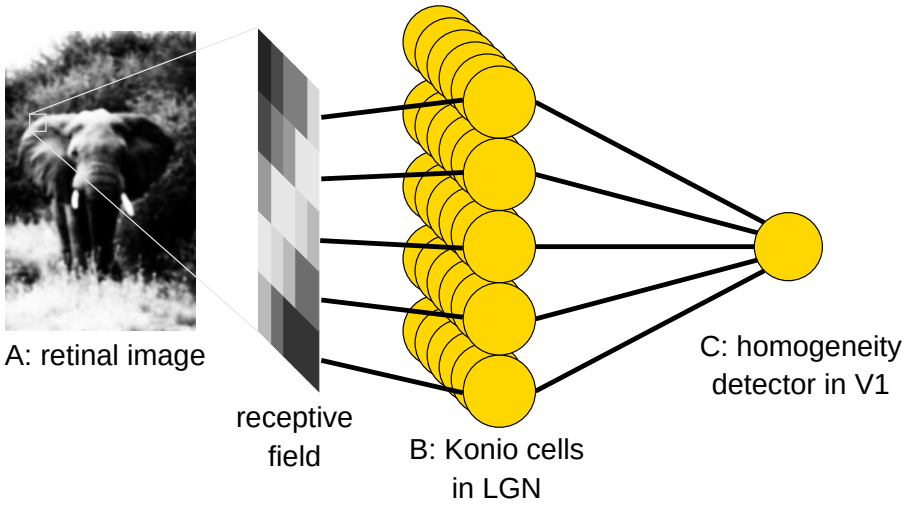


**Figure 3.2: Architecture of one homogeneity-sensitive unit.** Each konio (or K) cell receives input from one retinal ganglion cell. K cell projections converge to a coincidence detecting neuron. The image area the K cells cover defines the homogeneity detector's receptive field.

The retina directly feeds to the LGN konio cells. This part of the model is identical to the rate code model, mimicking characteristics of large receptive fields. High gray value leads to high K cell activation, while low gray-value leads to lower activation.

Thus, K-cell spike-timing depends on the activation by the retina; i.e. the higher the activation, the earlier (relative to stimulus onset) a spike is produced, and vice versa. The retinal activation of K-cells is adjusted such that they always produce a spike, even if a pixel is completely dark (black). In addition, K-cells in the model can fire only one spike during the simulation time interval. This ensures that the number of spikes $n_{spikes}$ is constant.

Figure 3.3 illustrates the working principle. If gray-value variance in the receptive field of the homogeneity detector neuron is high, K cells will produce spikes at different times, leading to a scattered pulse packet, which fails to drive the detector neuron to threshold (see figure 3.3A). However, low gray-value variance leads to a focused pulse packet, eliciting a spike in the detector neuron (figure 3.3B).

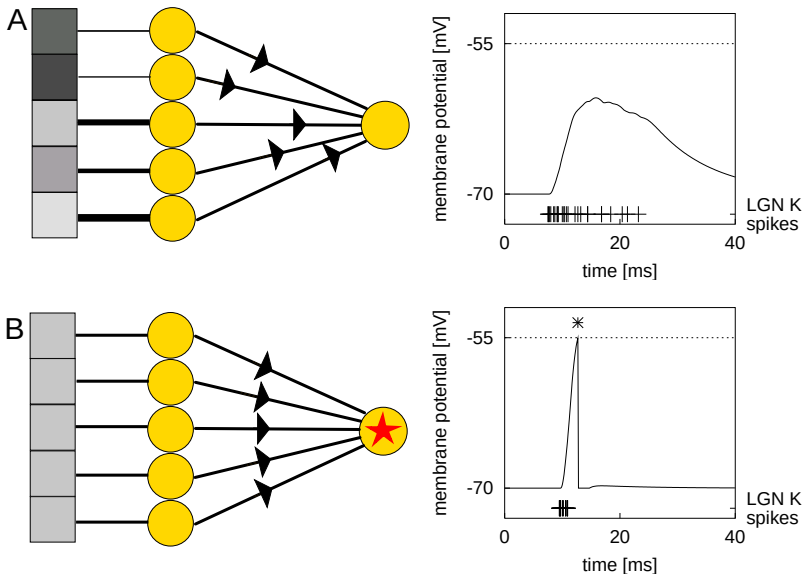


**Figure 3.3: Working principle.** Network units as depicted in figure 3.2. Only 5 of 25 input pixels are shown. **A)** *Left:* High gray-value variance in the receptive field. Line thickness reflects K cell activation. Triangles depict action potentials traveling along the axon. *Right:* Time-course of the detector's PSP. Plus-signs near the abscissa depict incoming spikes. **B)** *Left:* Low gray-value variance in the receptive field. The homogeneity-detecting neuron fires (illustrated by the star). *Right:* PSP with homogeneous input. As threshold (-55mV, dashed line) is reached, a spike is produced.

In terms of pulse packet theory, the gray-value distribution of an image patch is translated into a spike-latency distribution. The width $\sigma_t$ of the resulting pulse packet depends on the gray-level variance in the receptive field. Since $n_{spikes}$ is fixed, it depends only on $\sigma_t$ if a spike in the target neuron is triggered (the *coincidence threshold*, see section 2.2.3).

In consequence, there is a threshold in gray-value variance, below which the homogeneity-sensitive neuron will produce a spike.

## 3.3 Implementation

The model is realized by computer simulation of a network of integrate-and-fire neurons, i.e. Lapicque neurons with $\alpha$-shape synaptic currents (see sec. 2.1). I use the NEST-simulator (see appendix B) for the simulations. The core simulation is implemented as a SLI[1]-script.The input to this simulation script is an image file, which has already been preprocessed by the model retina.

### 3.3.1 Retina

K-projecting ganglion cells have large receptive fields. In order to mimick the effect of such large receptive fields, the input image is convolved with a Gaussian kernel. Subsequently, a sigmoidal activation function is applied pointwise, to emulate a threshold in neuronal activation. This part of the model is identical to the rate code model, as described in appendix A.1.

### 3.3.2 LGN K-cells

In the model LGN, pixel gray values from the stimulus image are translated into spike-latencies. This is achieved by translating the pixel's gray value $W$ into a current $I_{in}$ using the formula

$$I_{in} = I_{min} + \frac{(I_{max} - I_{min})}{W_{max} - W_{min}} \cdot W \tag{3.1}$$

with $W_{min/max}$ being the minimal/maximal gray value in the image, and $I_{min/max}$ the minimal/maximal input current.

[1]Synod Language Interpreter — an interpreter language for the NEST-simulator

The resulting current is injected into an LGN neuron. $I_{min}$ was set to 400$\mu A$, $I_{max}$ to 750$\mu A$. Using eq. (2.11) it follows that spike latency was in the range from 6.3 to 27.7 ms. Thus, 28 ms after stimulus onset, *all* LGN K-cells have produced a spike. These values were chosen to approximate spike latencies in primate visual cortex: After light has fallen on the primate retina, there is a delay of approximately 40 ms until all first spike responses related to that stimulus have arrived in V1.

Modeling retinal input to LGN K-cells as current input may not be very realistic, since axons from retinal ganglion cells transmit spikes, not currents, to the LGN. I use current stimulation to mimick the effect of a biologically plausible firing rate, in the sense that high activation, be it through firing rate or current stimulation, leads to a low-latency response spike.

By setting the refractory period for LGN neurons to a larger value than the simulation time, it is ensured that the number of spikes produced in response to a stimulus is constant, since every LGN neuron produces exactly one spike. Thus, LGN cells in the model have extremely transient responses.

### 3.3.3 Coincidence detecting neurons

The homogeneity detecting neuron in V1, as the LGN K cells in the model, is implemented with an integrate-and-fire neuron of the Lapique type. As described in section 2.2.3, the coincidence threshold can be adjusted by modifying $\tau_{syn}$, the time-constant of the post-synaptic current in response to an action potential.

I chose a value of $0.63$ ms for $\tau_{syn}$ and 0.75 $\rho F$ for the membrane capacitance C. These parameters yield best results for most stimulus images. In appendix C.1, I discuss why C must also be adjusted when $\tau_{syn}$ is changed.

### 3.3.4 Orientation-selective cells

Modeling orientation-selective cells can almost be called a research field on its own (see e.g. the review article by Shapley et al., 2003). The goal of this work is to demonstrate the benefit of homogeneity information to efficient processing in the visual system. Since orientation-selective cells are no major element of this work, they are realized using a rather simple approach.

By definition, an orientation-selective cell shall respond if there is a luminance-contrast edge inside its receptive field, and the orientation of this

edge is inside a certain range[2]. In the model, this is achieved by connecting retinal ganglion cells to the orientation-selective cells with an asymmetric weight matrix reflecting the cell's receptive field. As above, gray-values of the stimulus image are translated into activation currents using eq. (3.1). However, since orientation-selective cells have smaller receptive fields, the preprocessing stage mimicking large receptive fields is skipped, and the stimulus image is fed to the edge-sensitive cells without preprocessing. The model comprises orientation-selective cells for four orientations. The activation currents are fed to the orientation-selective cells using the weight matrices for the respective orientations:

$$0^\circ: \begin{bmatrix} -0.5 & 1 & -0.5 \\ -0.5 & 1 & -0.5 \\ -0.5 & 1 & -0.5 \end{bmatrix}, \quad 45^\circ: \begin{bmatrix} 1 & -0.5 & -0.5 \\ -0.5 & 1 & -0.5 \\ -0.5 & -0.5 & 1 \end{bmatrix},$$

$$90^\circ: \begin{bmatrix} -0.5 & -0.5 & -0.5 \\ 1 & 1 & 1 \\ -0.5 & -0.5 & -0.5 \end{bmatrix}, \quad 135^\circ: \begin{bmatrix} -0.5 & -0.5 & 1 \\ -0.5 & 1 & -0.5 \\ 1 & -0.5 & -0.5 \end{bmatrix}.$$

For the $0^\circ$ orientation-selective cell, the three vertical pixels in the center column of the receptive field activate the orientation-selective cell with weight factor $1$, while the six pixels vertically flanking the center to the left and to the right have a negative influence on cell activation, with weight factor $-0.5$ respectively. In consequence, the activation current the cell receives will be greatest if there is a bright vertical line in the center of its receptive field, and it will still be large if there is a vertical contrast-edge. The current will be zero for homogeneous illumination, and also if the orientation of the contrast-edge is perpendicular to the neuron's preferred orientation. The mechanism works similarly for the other orientations.

In addition to skipping the preprocessing stage mimicking large receptive fields, receptive fields of orientation-selective cells are considerably smaller than those for homogeneity-detectors, namely $3\times3$ pixels, compared to $5\times5$ pixel for the homogeneity-detectors.

Finally, the interaction with the homogeneity-detecting cells is modeled after the paradigm *Where there is a surface, there is no edge*, i.e. edges and homogeneous areas mutually exclude each other, with homogeneous ar-

[2]This is why such cells are often called "edge-detectors".

eas being dominant. Thus, retino-topographically spoken, everywhere a homogeneity-detecting neuron fires a spike, any response from orientation-selective cells is suppressed.

## 3.4 The variance threshold

In the scope of this model, homogeneity is defined as low gray-value variance (see section 1.4.2). However, in contrast to the rate code model, where variance is explicitely calculated, the variance threshold in the spike-based model cannot easily be described in analytical terms. Thus, we take an empirical approach to investigating the variance threshold (see figure 3.4): stimulus images with increasing levels of gray-value variance, measured in terms of grey-value standard deviation $\sigma_{grayval}$, are created by adding Gaussian noise to an homogeneous patch of $5 \times 5$ pixels. Here, the variance threshold lies in the range between $\sigma_{grayval}$ of 42.3 and 59.6.

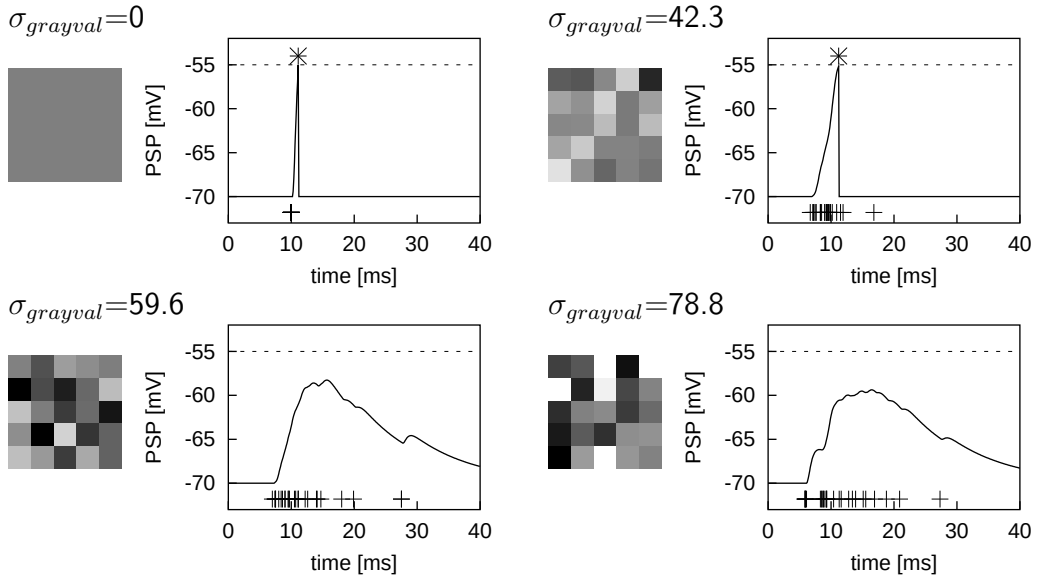


**Figure 3.4: Response to increasing levels of gray-value variance.** Stimulus images with different levels of gray-value standard deviation $\sigma_{grayval}$ and the homogeneity-detector's membrane potential trace. Dashed line: threshold (-55 mV). Stimulus is presented at $t$=0. Crosses at the bottom indicate LGN K cell spikes. Stars at the top indicate detector spikes.

## 3.5 Brightness dependency

The pulse-packet width $\sigma_t$ does not only depend on the gray-value variance in the stimulus image, but also on its average gray-level.

Activation of LGN K cells by retinal ganglion cells is modeled as current input (see section 3.3.2). The function for translating gray-value into current is linear (see eq. (3.1)). Thus, the gray-values of an image patch will result in current values in a fixed interval, even if the *brightness* of the image patch is shifted down (i.e. an offset is subtracted from the gray-values, making the image patch darker).

However, the relationship between current stimulation and spike latency is not strictly linear, but decays in an inverse logarithmic manner (see figure 2.1) and eq. (2.11). In consequence, a fixed current interval will result in a larger spike time interval, or *pulse packet width*, when an offset is subtracted. Thus, a given image patch will result in a larger spike-time interval when its brightness is shifted down.

Figure 3.5 sketches the principle: A bright image patch results in current values in an interval of 50 $\rho$A, from 600 to 650 $\rho$A, which yield spike times between approx. 9 and 10 ms, i.e. a pulse packet width of 1 ms.

Another image patch with a similar degree of grey-level variance, but lower average grey-value (i.e. lower brightness), also results in a current interval of 50 ms, however between 450 and 500 $\rho$A. The width of the resulting pulse packet is 4 ms, which is considerably larger.

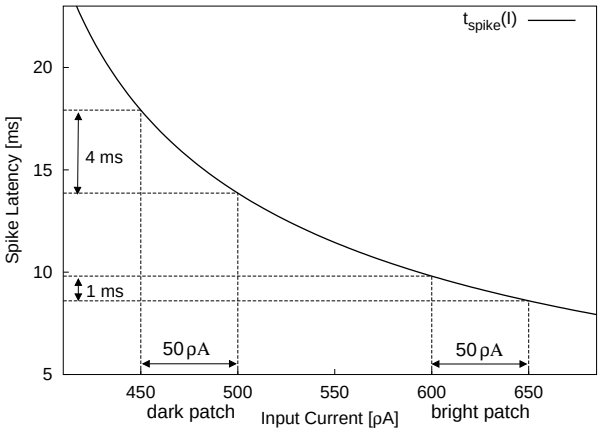


**Figure 3.5: Brightness dependency: Sketch of principle.** Solid curve: Spike latency (ordinate) for current stimulation (abscissa). An (imaginary) bright patch yields currents between 600 and 650 $\rho$A, a dark one between 450 and 500 $\rho$A. Dashed lines symbolize translation of current to spike latency. Resulting pulse packet widths are shown near the ordinate. More details see text.

Thus, the pulse-packet width $\sigma_t$ depends on the brightness, or *average gray-value* ($\mu_{grayval}$), of the stimulus image. In consequence, the variance threshold for the homogeneity-detector depends on the stimulus' average

gray-value. This means that dark homogeneous areas cannot be detected with the same reliability as bright ones, in the sense that the detector will be more sensitive to bright homogeneous areas.

To overcome this limitation, I use two detector units working in parallel: One "on"-detector with the default characteristics, more sensitive bright homogeneous areas, and one "off"-detector with inverted characteristics, to detect dark homogeneous areas. In fact, the implementation of both detector types is the same, but the off-detector's stimulus image is inverted, using the formula

$$W_{off} = W_{max} - W_{on} \ ,$$

where $W_{off}$ is the pixel gray-value in the stimulus image for the off-detector, $W_{on}$ the gray-value of the original pixel and $W_{max}$ the maximal gray-value.

Figure 3.6 shows the complementary behavior of the parallel on- and off-detectors: Two stimulus images with similar gray-level variance $\sigma_{grayval}$ but different average gray-value $\mu_{grayval}$ are presented to the on- and off-detectors. The on-detector responds to the brighter stimulus image, but does not respond to the darker one, because the resulting pulse packet from the darker image is too wide. The off-detector shows complementary behaviour: The pulse packet resulting from the brighter stimulus fails to evoke

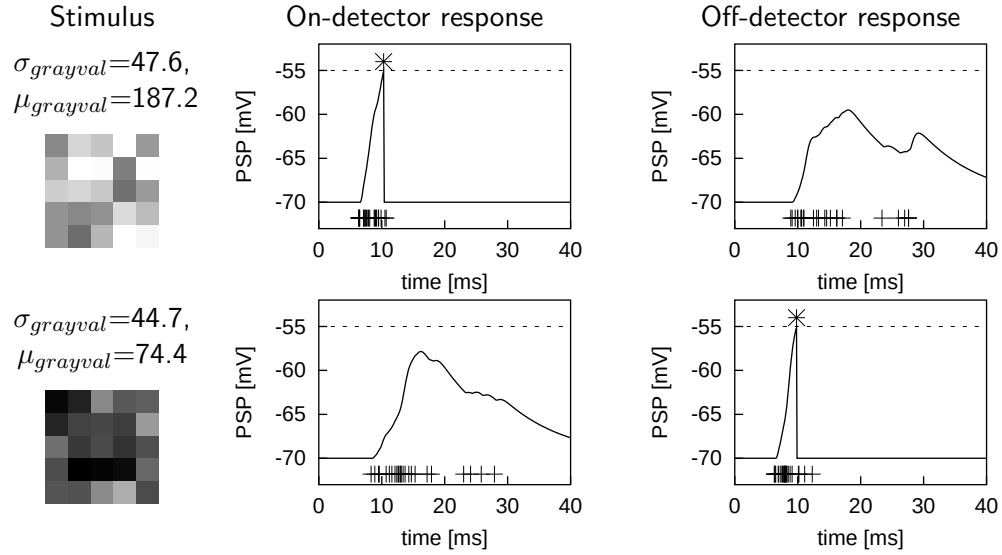


**Figure 3.6: Brightness dependency: on- and off-detector.** *Left:* Stimulus images. Both images have approximately the same amount of gray-level variance, but differ in brightness (*top:* bright, *bottom:* dark). *Middle:* On-detector PSP in response to the stimulus images. Lines and points have the same meaning as in figure 3.4. *Right:* Off-detector response.

a spike in the detector, while the darker stimulus yields a focused pulse packet and triggers an action potential.

Thus, the stimulus image is analyzed in parallel by on- and off-detectors, where on-detectors respond to bright homogeneous stimuli, and off-detectors respond to dark homogeneous stimuli. The responses of both detectors are combined in the way the logical OR operation would do: If one or both on- and off-detectors has produced a spike, the stimulus inside the receptive field is considered as homogeneous. If none of both detectors responds, the stimulus is considered inhomogeneous.

By combining the responses from on- and off-detectors, brightness dependency is reduced, because bright homogeneous stimuli can be detected as well as dark ones. In section 4.1.1, I show the results from parallel on- and off-detectors "in practice", i.e. when analyzing natural images.

# Part 4: Results

In this section, I demonstrate how the model performs with natural images as stimuli. First, I give an overview how the stimuli are applied to the model, and how to interpret the resulting images (section 4.1). Then I present the results for interaction with edge-detectors (section 4.2).

The process of emulating large receptive fields has quite an impact on the result. This is elucidated in section 4.3. Finally, I examine the effect of changing the receptive field size of the homogeneity-detecting neuron, i.e. changing the number of K-cells that project to one coincidence-detecting neuron.

## 4.1 Analyzing natural images

In order to cover the whole stimulus image, the network is extended, such that at every image point there are two homogeneity-detecting neurons, one on- and one off-detector, each with a receptive field of 5×5 pixels. If the gray-value variance of the stimulus inside their receptive fields is low, one or both of the on- and off-detector neurons will produce a spike, indicating that their receptive fields are inside a homogeneous image area.

Homogeneity-detector neurons are arranged in a retinotopic manner. Thus, by displaying the spikes produced by homogeneity-detector neurons in the same topography, we obtain a *homogeneity map* of the input image: a spike at a given image position indicates that the image area in the receptive field of the respective homogeneity-detector neuron is homogeneous. Figure 4.1 (next page) sketches how the homogeneity map is obtained from an image: The original retinal stimulus image is first processed to account for large receptive fields of retinal ganglion cells. The resulting input to the LGN is then processed by the homogeneity-detecting network, and the homogeneity map is obtained.

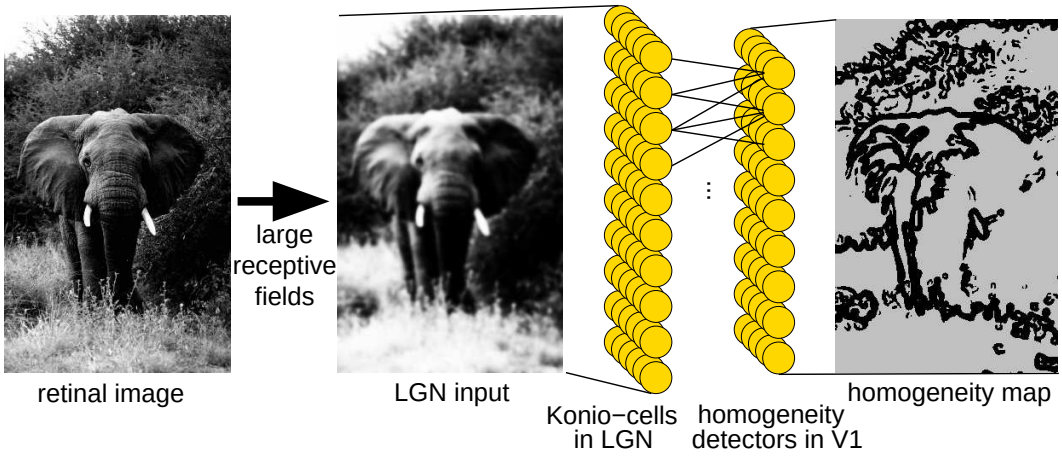


**Figure 4.1: Obtaining the homogeneity map.** Gray pixels in the resulting spike map mean that the homogeneity-detector unit at this image location has produced a spike, while a black pixel indicates that no spike was produced.

### 4.1.1 Complementary results from on- and off-detectors

As stated in section 3.5, the results from on- and off-detectors are complementary. Figure 4.2 shows, how on-detectors respond in bright homogeneous areas where off-detectors fail, and vice versa. Certainly, in rare cases where gray-level inside the receptive field is around the mean gray value ("halfway" between black and white), both detectors respond. The "threshold" operation applied in retinal processing however, renders this case very unlikely.

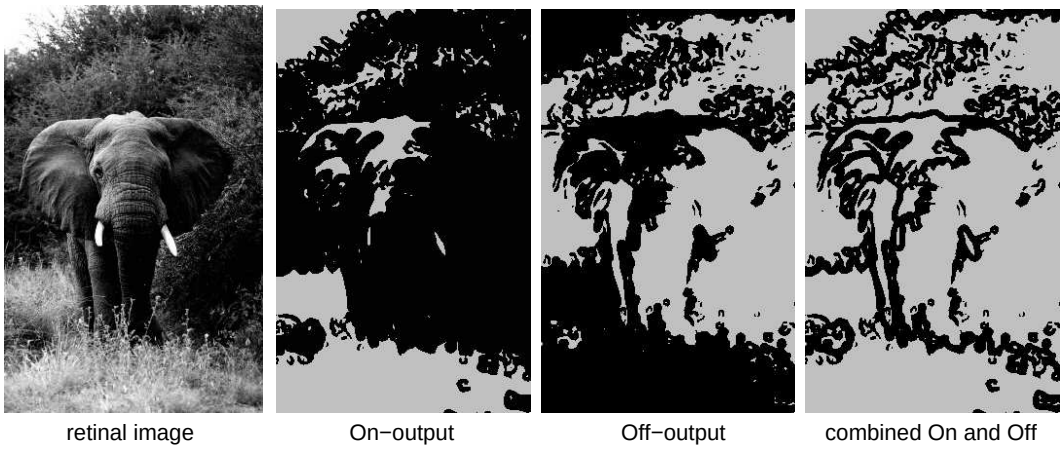


**Figure 4.2: Complementary results from On- and Off-detectors.**

## 4.2 Interaction with edge-detectors

Spikes from orientation-selective cells (or edge-detectors), which are also retinotopically arranged, give an *edge-map* of the image. Figure 4.3 illustrates the mechanism: homogeneity-detecting cells inhibit the relay of responses from orientation-selective cells in homogeneous image regions, yielding a clearer edge-map.

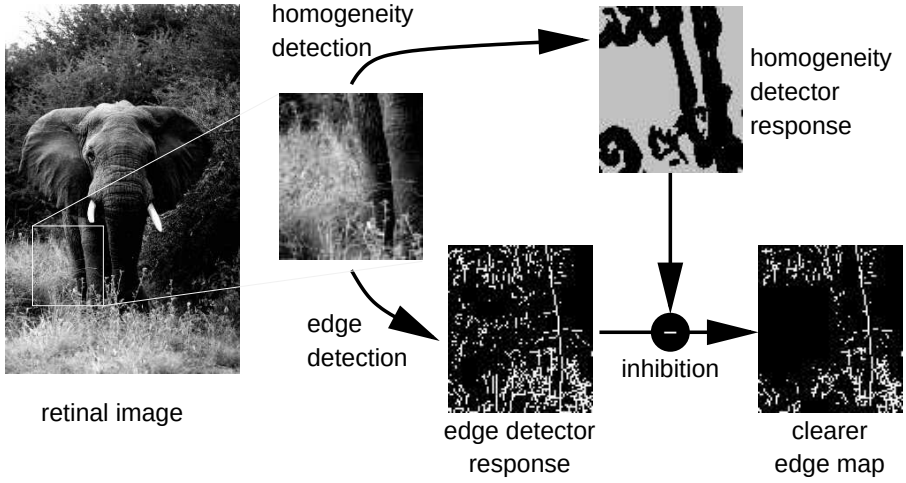


**Figure 4.3: Inhibition of edge-responses by homogeneity responses: Principle.** Gray pixels indicate spikes in the homogeneity-detector response. In edge-detector responses, spikes are represented by white pixels.

Figure 4.4 shows the result for the entire image: responses to spurious edges in homogeneous image areas are suppressed while salient edges persist. Figure 4.5 (next page) shows results for another stimulus image.

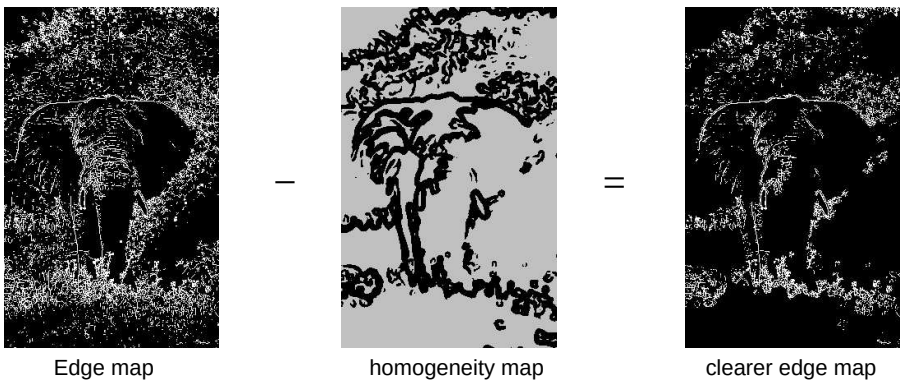


**Figure 4.4: Inhibition of edge-responses by homogeneity-responses: Whole image.** Inhibition of relay of edge-detector spikes is symbolized by the minus sign.

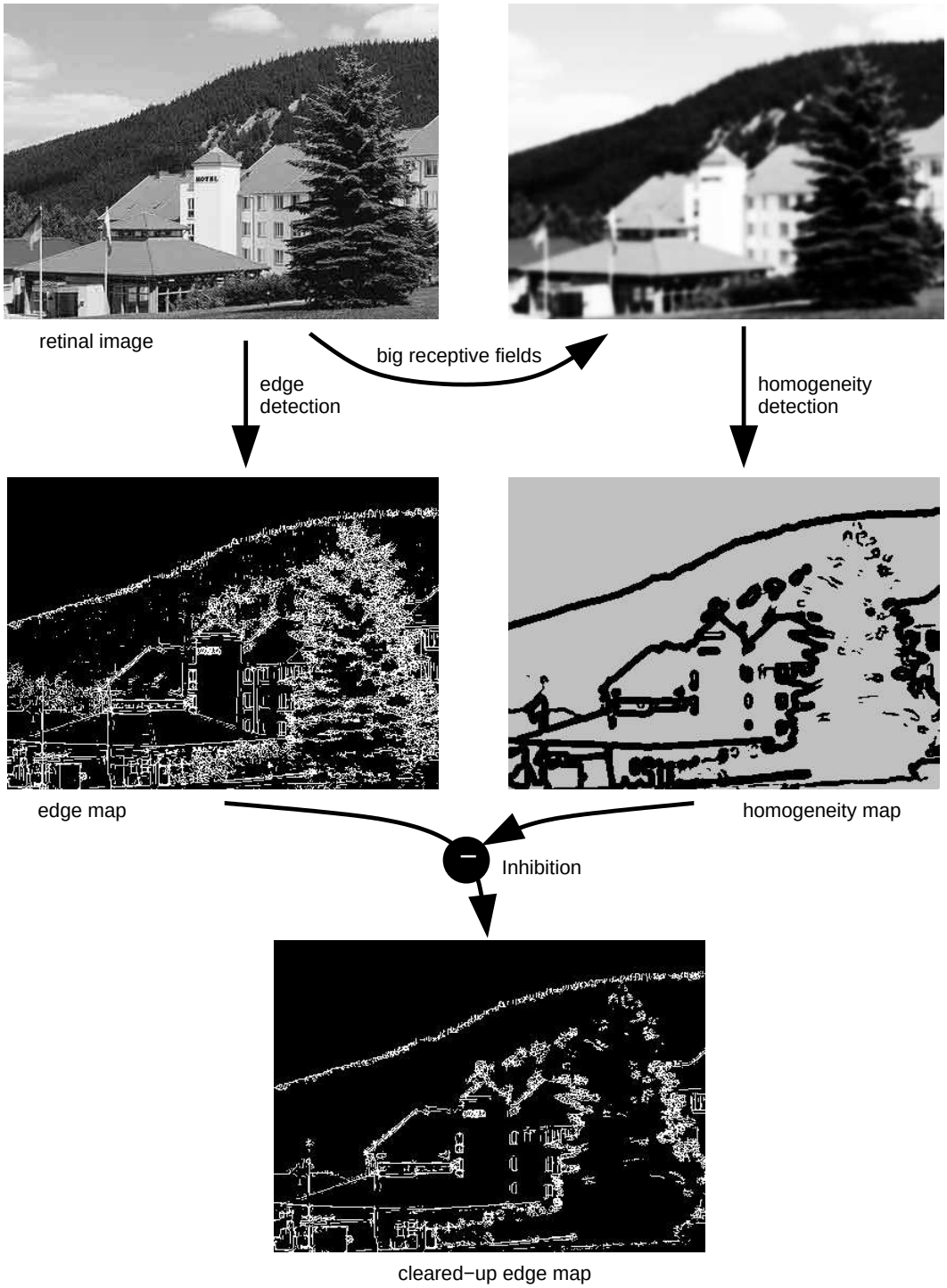


**Figure 4.5: Another example.**

## 4.3 Effect of large retinal receptive fields

The retinal processing stage mimicking large receptive fields (see section 3.3.1) has great impact on the quality of homogeneity detection, yet the effect becomes visible preferably in noisy images. Figure 4.6 shows two examples: In high-contrast images with clear object contours, such as the robot head, there is no qualitative difference. However, in the "Elephant"-image below, we see a clear difference in the spatial pattern of homogeneity-detector spikes. With retinal processing, the detected homogeneous regions are much more coherent. This gets clear when looking at the process itself, which effectively reduces noise and facilitates the distinction between edges and homogeneous areas.

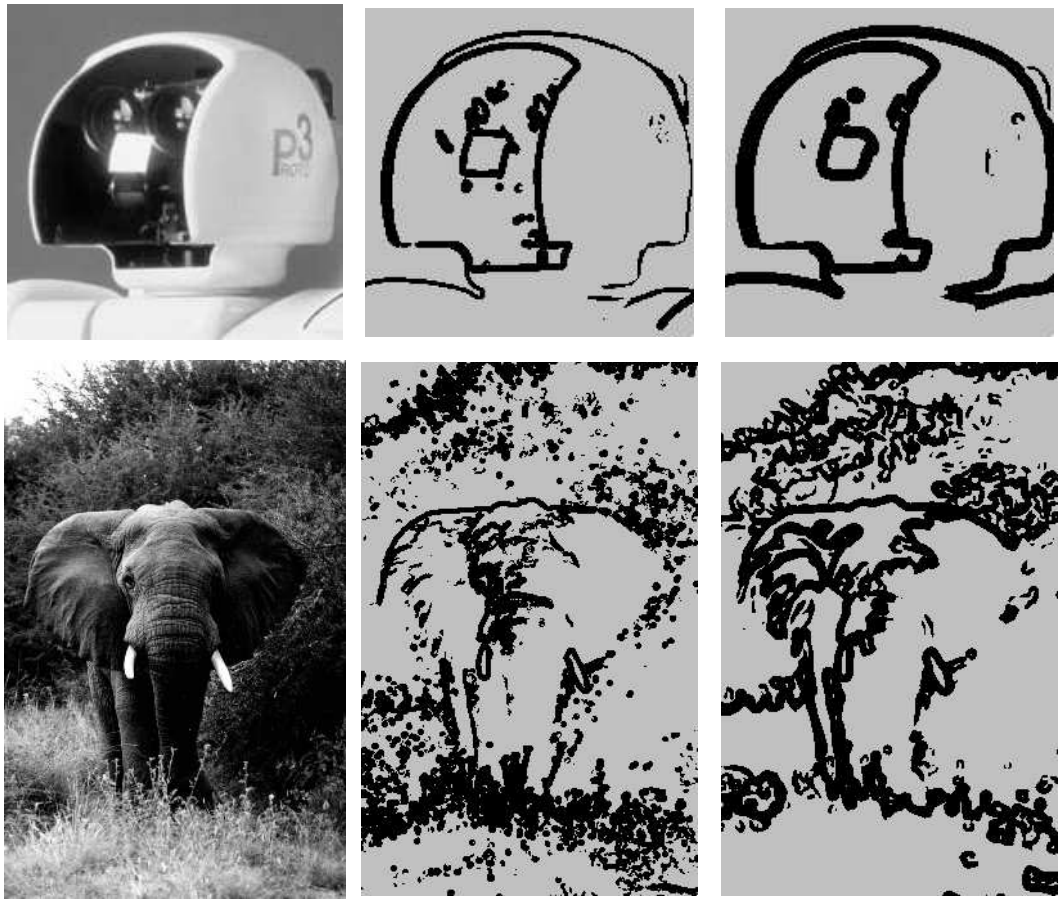


**Figure 4.6: Effect of retinal preprocessing.** In the above image (robot head), there's no qualitative difference. The reduction of spurious surface responses in noisy areas in the "elephant"-image is more striking.

Figure 4.7 illustrates what is going on in the retina. Retinal preprocessing is implemented in a two-step procedure: First, the image is convolved with a Gaussian. This low-pass filtering supresses high-frequency clutter in the image. The subsequent pointwise application of a sigmoid transfer function enhances contrast.

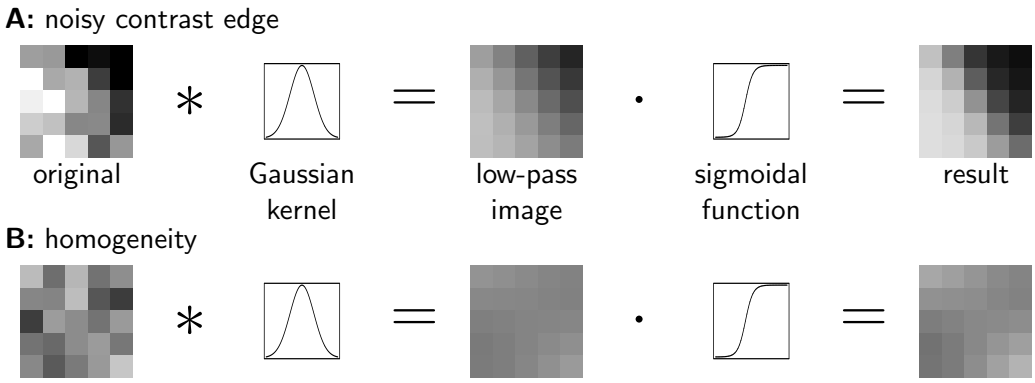


**Figure 4.7: Retinal preprocessing: principle.** *A:* noisy contrast edge stimulus. The stimulus image is convolved with a Gaussian kernel, which yields a low-pass filtered image, to which a sigmoidal (or *threshold*) function is applied pointwise. *B:* same procedure with noisy homogeneous stimulus.

Without the application of the sigmoid, this would lead to a more homogeneous image patch, however if contrast is enhanced, the edge-character is emphasized. The benefit of this is that either "edgy-ness" or homogeneity of image areas are enhanced, making a clear separation between the two cases easier. For an example what the result of retinal preprocessing looks like for whole images, please refer to the two leftmost images in figure 4.1, page 36 and the upper two images in figure 4.5, page 38.

In the example above, the stimulus has an average gray-value very close to the "mean gray", i.e. a gray-level of 50% (in the middle between black and white on the luminance scale). As figure 4.8 on the next page shows, the effect of the two-step "preprocessing" is even more pronounced when the stimulus has a mean gray-value well above or below the mean gray, since the threshold operation (application of a sigmoidal function) will "push" the stimulus gray-values versus black if the stimulus is darker than 50% gray, or versus white in the other case. In consequence, not only the stimulus' "edgy-ness" or homogeneity are emphasized, but also its "darkness" or "brightness", accommodating the parallel on- and off-detector architecture.

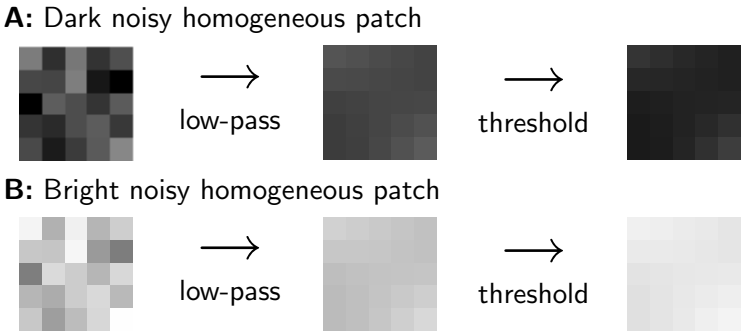


**Figure 4.8: Retinal preprocessing: dark and bright homogeneous stimuli.** *A*: dark stimulus, *B*: bright stimulus. For both stimuli, the gray values are equalized to their average by the low-pass step. The threshold-step then "pushes" the stimulus versus the bright respective dark extreme.

## 4.4 Homogeneity-detector receptive field size

In contrast to the previous section, which deals with the effect of large receptive fields in the retina, this section deals with receptive field size for the homogeneity-detecting neuron, i.e. how many LGN K cells project to one homogeneity-detecting cell.

However, at least in images with high contrast and low amounts of high-frequency clutter, increasing the receptive field size of the homogeneity-detector neuron yields qualitatively the same results as retinal preprocessing, as figure 4.9 illustrates.

Instead of low-pass filtering and contrast-enhancing the image, the original (retinal) image is not modified, but the number of K cells projecting to one homogeneity-detector neuron is increased.

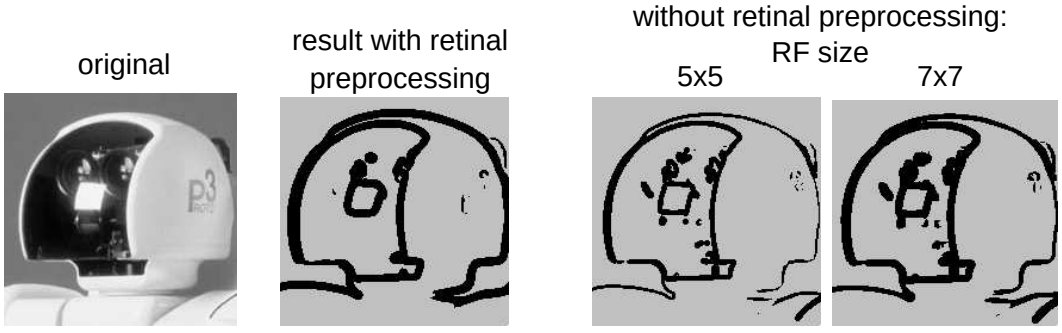


**Figure 4.9: Effect of homogeneity-detector receptive field size.** "RF" is abbreviated for "receptive field".

# Part 5: Discussion

## 5.1 Comparison with the rate code model

Figure 5.1 shows results from homogeneity detection by the spike-based model and the rate code model. The results are qualitatively the same, in the sense that the same regions have been recognized as homogeneous. Note that the rate code model has a graded distinction between homogeneous and non-homogeneous regions, while in the spike-based model, it's "all or nothing", i.e. spike or no spike.

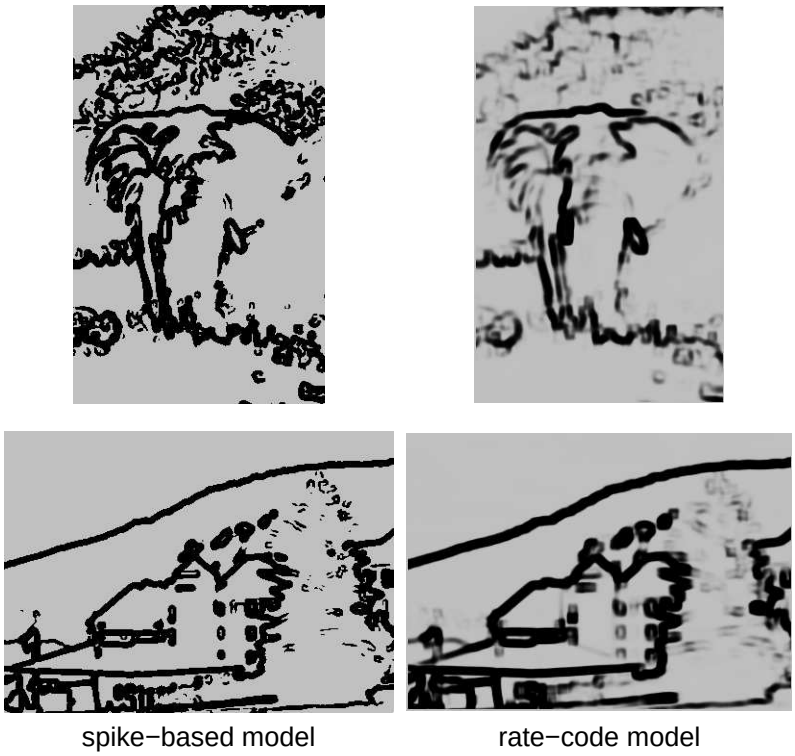


**Figure 5.1: Comparison of results from rate code model and spike-based model.** Homogeneous areas obtained from the spike based model (*left*) and from the rate code model (*right*).

By using spiking neurons, the spike-based model is biologically more realistic. Moreover, unlike other mechanisms to reduce spurious edges (see section 1.4.1), the mechanism is purely feed-forward and uses only one spike per neuron (except between retina and LGN). Thus, it fulfills the requirements for rapid processing postulated by Thorpe et al. (1996).

### 5.1.1 Parameter stability

It must be emphasized that the model yields remarkably stable results. Unless otherwise noted, all results presented have been achieved with the same parameter set for neurons and receptive fields; only the total number of neurons is adjusted according to the size of the stimulus image (in pixels). Regarding the different resolutions and scales of the images (see appendix D), this argues in favor of the model's robustness.

## 5.2 Biological context

### 5.2.1 Homogeneity detection in the Cortex

In the hypothesis (see section 3.1), it is assumed that the homogeneity detection is performed in the K-Path. In the first place, it is not clear if homogeneity detection is actually *happening* in the cortex. However, looking at the properties of cells in the K-pathway let this hypothesis appear more plausible.

- K-cells have large receptive fields, often with no antagonistic surround. In consequence, they should not selectively respond to oriented contrast edges, or contrast stimuli at all, but rather to diffuse illumination.

- Cells in the K-pathway are phylogenetically very old. Homogeneity detection could be a basic kind of vision; it does not convey detailed information about the visible environment, but it gives information of the type *"there is something"*. In this context it is also interesting to note that K-cell receptive fields are evenly distributed across the retina (except in the fovea), which also supports the rudimentary character of "homogeneity vision".

- Homogeneity detection and edge detection (achieved by the phylogenetically younger M- and P-pathways) are complementary mechanisms; They can enhance each other in a parallel, feed-forward way in order to yield a more robust percept of the outside world.

The hypothesis has also some caveats which remain to be investigated. First of all, as I pointed out in the introduction to this work, konio-cells in the LGN are a very inhomogeneous group of cells. It may, thus, be inappropriate to generalise them to *the* K-Path, implying that cells in this path are functionally similar. Some might be involved in homogeneity detection, some might have other roles.

Moreover, it is commonly assumed that small bistratified ganglion cells respond to color changes in the medium to short wavelength range (yellow - blue), and hardly to pure luminace changes (black - white) (Rodieck, 1998). These cells provide the input to the middle two K cell layers in the LGN, which in turn project to the blobs in layer 2/3 in V1, thus being a part of the putative biological realization of the model. Color information is not included in the model; all stimulus images are gray-scale, in disregard of the ganglion cells' color-selectivity.

Thus, introducing color in the model will be necessary in future studies. As I also explain below (section 5.3), introducing color may yield interesting results, in particular to the benefit of object segmentation.

Another possibility is that the actual homogeneity detection does already take place in the LGN instead of layer 2/3 of V1. Since retinal ganglion cells are able to produce synchronized spikes, especially if they have overlapping receptive fields or share common input (Greschner et al., 2002), the model would also work if we exchanged the LGN cells with ganglion cells; cells in the LGN would then do the actual coincidence detection.

But, homogeneity detection could also be performed in other pathways than the K-pathway, e.g. by cells in the M-pathway, which are not color-selective. Their receptive fields are also relatively large and convey visual information of relatively low spatial resolution. M-cells respond transiently, which is in favor of the processing concept, since it needs only one spike. The fact that M-cells have receptive fields with an antagonistic surround does not necessarily exclude that some also respond to diffuse illumination, for instance if the effects of the surround and the center of the receptive field are not exactly balanced.

Indeed, some recent studies relate rapid object classification to the M-pathway. Bar (2003) proposes that low spatial frequencies in an image are projected rapidly from early visual areas to the prefrontal cortex (PFC). This coarse information is used by the PFC to deduce an "initial guess" about the input image, triggering corresponding object representations in inferotemporal cortex (IT), which are then integrated with more detailed bottom-up information. In this study, the M-pathway is considered a candidate for the rapid bottom-up process, since it conveys information about low spatial frequencies early and rapidly. However, it is not discussed if the K-pathway could be a part of the rapid bottom-up process, in spite of its suitable functional and physiological properties, such as its low spatial resolution and direct connections from the LGN to higher brain areas.

Also Delorme et al. (1999) relate ultra-rapid categorization to the magnocellular pathway. In this work, too, a contribution by the K-pathway is not discussed. However, these studies do not consider homogeneity detection, either.

### 5.2.2 Accuracy of spike timing

In the model, the accuracy of spike timing in response to the stimulus is very high. LGN spike timing depends only on the intensity (or gray-value) of the stimulus pixel, with no stochastic component. For real neurons, this is not given, since there is always some stochastic aberration in spike timing, be it by fluctuations of the membrane potential, or by intrinsic imprecisions in the spiking mechanism.

The spike latency code used in the model can only be realized in real neurons, if they can produce their action potentials with a sufficient degree of temporal accuracy. This question has been actively debated over the years. Experimental results show that neurons can indeed produce spikes with an accuracy of about 1 ms (Tanaka, 1983; Mainen & Sejnowski, 1995; Reid & Alonso, 1995; Stevens & Zador, 1998; Reinagel & Reid, 2002). However, there is also evidence that, at least in some neurons, the incoming spike *rate* rather than the exact timing of those spikes determines the neuron's response (Shadlen & Newsome, 1994; Shadlen & Movshon, 1999). But, as Usrey (2002) points out, precise spike timing plays an important role in thalamo-cortical processing, so we have to assume that neurons in this brain

areas actually are capable of responding with a high degree of temporal accuracy, which is in favor of the model presented here.

Another caveat is the synchronization of LGN K cells to stimulus onset. That is, some mechanism is required to synchronize the reference point in time for spike latency of the K cells.

The experiments by Fabre-Thorpe et al. (2001) used *flashed* images. Here, synchronization is achieved by the stimulus, especially if a noise mask precedes stimulus presentation. If the subject faces a plain screen before stimulus onset, homogeneity-detectors respond to the whole screen. If now an image is flashed on the screen, K cell responses get desynchronized in non-homogeneous regions, and homogeneity-detectors with receptive fields in these regions stop firing.

The situation changes in freely-looking individuals, since the retinal image of the outside world seems to change in a continuous manner. However, the process of *looking* is intersected by saccadic eye movements, or *saccades*. During the quick movement of a saccade, vision is effectively inactivated (Rodieck, 1998). This is consistent with recent experimental data showing that, in behaving macaque monkeys, neuronal response in V1 is reduced during saccades (Slovin et al., 2002). The mechanism behind this effect is not clear; it could be due to some kind of trigger signal, which actively attenuates neuronal response, or simply to the very fast motion of the retinal image (approx. $800^\circ/s$) during a saccade, which typically takes some tens of milliseconds, depending on its magnitude. In any case, this effect could provide the "reset"-signal for the K cells.

In addition, Greschner et al. (2002) demonstrate how fixational eye movements (*microsaccades*) effectively synchronize ganglion cell responses. Here, synchronization of neuronal response is not linked to the onset of fixation, but throughout the entire intersaccadic interval.

A very intensely discussed mechanism for response synchronization are stimulus-evoked oscillations of neuronal activity. Rodemann (2003, chapter 8) shows a mechanism related to the one presented in this work, using an oscillatory signal for neuronal synchronization.

## 5.3 Benefits of homogeneity detection

It is widely assumed that mostly contour information counts in cortical processing of visual input in primates and most mammals in general. Homogeneous regions are thought not to play a big role. There are certainly reasons for this assumption, such as the fact that many of the cells found by physiologists in primate visual cortex actually do *not* respond to diffuse illumination of their receptive field, but require at least some kind of contrast. Moreover, for most cells the stimulus is required to *move* in order to elicit a response. However, in temporary disregard of these findings, I want to point out some benefits which can be drawn from information about homogeneous regions in the visual field.

I have already stated the first point: Homogeneity and edge detection are complementary mechanisms, which do not *break* each other in the first place. However, *not breaking* another mechanism is not a merit in itself. So, let's look at what can be achieved by the mechanism alone, i.e. *without* interaction with edge-detectors.

### Segmentation

The term *segmentation* in the context of visual processing describes the process of dividing an image into separate parts, such as objects. It is commonly assumed that segmentation is necessary to achieve robust object recognition (see Ullman, 1995). A low-pass filtered image with enhanced contrast, as resulting from large retinal receptive fields (see section 4.3), facilitates this process, because much high-frequency information is removed from the image, while coarser contrast information is enhanced.

Additionally, information from homogeneity-detecting cells can sort of "highlight" large continuous regions. These regions can be considered as a pre-segmentation of the image, and could then be analyzed in more detail by cells sensitive to other features.

Especially when considering the direct projections of LGN K cells to higher visual processing areas in the ventral pathway, this thought becomes interesting: cells in these areas can be "primed" for stimuli coming through the other pathways, e.g. by a slight subthreshold depolarisation making them more sensitive. Thus, the coarse hypothesis about the input, as conveyed by K-cells, can be refined by analyzing the input in more detail.

Moreover, as I already stated above, the model presented here does not exploit hints from color. As Gegenfurtner & Rieger (2000) state, color cues might be of great use for object segmentation. Thus, it may be valuable to include some kind of color representation in future studies.

## 5.4 Outlook

### 5.4.1 Psychophysics to test the hypothesis

Having postulated a model for homogeneity detection in primate visual cortex, it is desirable to test the hypothesis.

A psychophysical study based on the object classification experiments by Fabre-Thorpe et al. (2001) would be suitable to test whether homogeneous image regions do at all play a role in rapid object classification. One could remove low-pass information from stimulus images, for instance by replacing homogeneous regions with a texture or random pixel noise and see whether the performance in rapid object classification decreases.

Further, by using S-cone isolating as well as L/M-cone isolating stimuli (i.e. stimuli, to which only cells receiving S- resp. L/M-cone input respond differentially) would allow to test whether cells in the K-pathway, which are mainly S-cone driven, play an important role in rapid object classification. Maybe it is possible to construct stimuli, in which contrast edges are invisible to cells in the K-pathway, and test rapid object classification with these stimuli. The model postulates that in this case fast object classification would be impaired, since the K-pathway would signal a completely homogeneous image.

An even more tricky approach implies making homogenous regions invisible to the K-pathway, such that it can not detect homogeneous regions. Also here, rapid object classification would be impaired, if the model is correct.

### 5.4.2 Other uses for the processing concept

#### Coding and learning in cortical networks

The fast feed-forward, one-spike based processing concept used in this work could also be transfered to other tasks than the detection of homo-

geneous image areas, for instance to detect regular patterns (*textures*). A simple texture, consisting of parallel oriented lines, could be detected with orientation-selective cells projecting to a coincidence-detecting neuron. The preferred orientation of the orientation-selective cells should be similar, and their spike-timing would have to depend on contrast and orientation of the stimulus. When a regular texture of parallel identical lines is presented, the orientation-selective cells would produce a focused pulse packet, causing the coincidence-detecting neuron to produce an action potential.

Generally, by choosing appropriate receptive field characteristics, any kind of regular pattern can be detected with this processing concept. The only requirement is the correct "wiring", i.e. the respective feature-detecting cells must project correctly on an coincidence-detecting neuron, which represents the stimulus. Admittedly, the thought of having specialized cells for every kind of pattern is unrealistic regarding the finite number of cells in the brain and the infinite number of possible stimuli. But, given the highly divergent/convergent architecture of synfire chains, it is reasonable to assume that in such an architecture almost every feature can be represented. This is related to the concept of *liquid computing*: given a complex network, any feature in the input can be detected by a *read-out cell* connected to the "right" neurons (cf. Maas et al., 2002; Jaeger, 2001).

Also when considering *learning* in neural networks, in the sense of changing synaptic weights to represent a given feature, coincident spikes will provide a strong learning signal. A suitable learning rule would strengthen those synapses through which an action potential arrived in temporal coincidence with other action potentials. The amount of potentiation depends on the number of coincident spikes. Now imagine a two-layered network of spiking neurons receiving visual input: the first layer contains various types of feature detectors, i.e. orientation-selective cells, color-contrast cells, but also cells with homogeneous receptive fields. Spike-timing of these neurons depends on how good the stimulus matches the optimal stimulus, i.e. preferred orientation. The spike timing neurons weakly connect in a highly divergent manner to a layer of coincidence-detecting neurons. A given stimulus, presented to the network, will elicit a characteristic pattern of spikes, some of which will be coincident. Given that the projections are divergent enough, neurons in the coincidence-sensitive layer will develop selectivity for this stimulus. So, a spike-latency based network

for object recognition may emerge. It remains to be investigated how stable this learning process can be, and how good it performs in comparison with other established learning rules.

# Part 6: Appendix

## A Implementation of the rate code model

The rate code model is described in detail by Gewaltig et al. (2002). For reference, I reprint relevant details here.

### A.1 Retinal input

The receptive field (or *retinal input*) of LGN K cells (large, often no antagonistic surround) is modeled by a two-dimensional gaussian $h(x, y)$ with standard deviation $\sigma_1$:

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

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

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

and $B[x, y] = (A * h)(x, y)$ with threshold $\theta_1$ and slope $b_1$. The threshold is determined by the mean activity in the image. $A$ is the input image.

Please note, that in the spike-based model, there is an important difference in the assignment of the model stages to cells: In the rate code model, the application of a sigmoidal function accounts for LGN K-cell activation. However, in the spike-based model, the LGN K-cells are modeled by integrate and fire neurons. The sigmoid activation function there models the activation of *retinal ganglion cells*. The formula for preprocessing input images however stays the same.

## A.2 Rate code homogeneity-detectors

The K-cell layer $C$ projects to a layer of homogeneity-detectors $E$. The receptive field of a homogeneity-detector at position $(x, y)$ is modeled by a circular patch $\gamma(x, y)$ within $C$. Homogeneity-detectors effectively evaluate the variance of color or gray-level in their circular receptive field.

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

with

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

The activation of the homogeneity-detectors is given by $E[x,y] = \Theta_2\left(D[x,y]\right)$, where $\Theta_2(\cdot)$ is a sigmoid activation function equivalent to (2). The threshold $\theta_2$ is again chosen according to the mean activity of their input.

# B Simulation environment

## B.1 NEST - NEural Simulation Toolbox

All simulations were carried out using the NEST-simulator (Diesmann & Gewaltig, 2002). Available for many computer platforms, it is designed for parallel computer architectures, but also runs on single processor machines. The simulator can handle very large neuron populations. In this work, simulations have up to 1,8 million neurons, depending on the image size. Simulating 55 ms (550 steps with a resolution of 0.1 ms) with this network took about 14 minutes on an 8-processor SunFire V880.

## B.2 Neuron model implementation

The Lapicque model described above is implemented in NEST using the method for *exact digital simulation* presented by Rotter & Diesmann (1999). The implementation itself is documented in Mohns (2000, p. 9ff. ).

## B.3 Adjustable parameters

The implementation of the neuron model in NEST (therein called `iaf_neuron`) lets the user adjust most of the parameters present in the equations describing the neuron. One parameter which can only be modified indirectly is $R$, the membrane resistance. Since $R \cdot C = \tau$ (see eq. (2.3)), one can adjust either $\tau$ or $C$ in order to change $R$.

The parameters and their default values are shown in table 1.

| Parameter name | Default | Description |
|---|---|---|
| `U0` | -70 mV | resting membrane potential $V_0$ in mV |
| `Theta` | -55 mV | threshold $\theta$ in mV |
| `C` | 250 $\rho$F | membrane capacitance $C$ in $\rho$F |
| `Tau` | 10 ms | membrane time constant $\tau$ in ms |
| `TauSyn` | 2 ms | time constant of synaptic current $\tau_{syn}$ in ms |
| `TauR` | 2 ms | refractory period $t_R$ in ms |

Table 1: Adjustable parameters of the `iaf_neuron` model in NEST.

# C Parameters of the spike-based model

## C.1 Parameter search for coincidence detection

To achieve sufficient separation of the homogeneous and non-homogeneous case, some parameters of the detector neuron have to be adjusted. With its default parameters (see Appendix B.3), the detector responds even to very wide pulse packets.

As explained in section 2.2.3, I use $\tau_{syn}$, the time-constant of the post-synaptic current in response to a spike to adjust the "homogeneity threshold" of the detector neuron.

According to eq. (2.13), for smaller values of $\tau_{syn}$, less total charge is delivered to the neuron, which in turn leads to a smaller PSP amplitude. In order to yield the same PSP amplitude, the current amplitude must be adjusted. In the simulation, this was achieved by adjusting $C$, the capacity of the membrane. The NEST simulation program keeps $\tau$, which is defined as the product $R \cdot C$, at its assigned value, which can be adjusted separately. In consequence, when changing $C$, effectively $R$, the membrane resistance, is changed too.

Smaller values of $C$ result in greater values of $R$. In turn, higher membrane resistance results in a higher resulting membrane potential for a fixed current $I$, according $I = \frac{U}{R}$.

One could also have changed the weight of the synaptic connection, but this parameter is used for adjusting the maximum PSP amplitude when changing the receptive field size (i.e. the number of LGN K cells projecting to one homogeneity-detecting cell), which leads to a change in $n_{spikes}$, the number of incoming spikes.

## C.2 Neuron parameters in the simulation

Most neuron parameters were left at their default values (see table 1). Table 2 lists the adjusted parameters, and a description why they are changed.

| Parameter name | Default | Description |
|---|---|---|
| TauR | 200 ms | Larger than simulation time to allow only one spike per neuron. |
| TauSyn | 0.63 ms | Adjusting the coincidence threshold |
| C | 0.75 $\rho$F | Counteracting response diminuition by lower $\tau_{syn}$ |

**Table 2: Adjusted parameters.**

# D Used images

The images were either pulled off the internet or stem from the COIL-database (Nayar et al., 1996). All images were converted to grayscale with gray values between 0 and 255 using a graphics processor (The Gimp).

In all figures, all images were scaled for convenience. Figure 1 shows all images in correct relative sizes. Real sizes (in pixels) are listed in table D.

| Picture | size [px] |
|---|---|
| Elephant | 266 × 419 |
| Black forest mountain hotel | 409 × 297 |
| Robot head | 196 × 196 |
| Toy cat | 128 × 128 |

**Figure 1: Relative sizes of the pictures used in the simulations.**

# E Some words on modeling

### Making things simpler

In the IAF-neuron model, as in every scientific model, some intricate details of the modeled system have been discarded, others have been simplified. For example, the IAF-neuron is a *point-shaped* neuron model, i.e. it has no spatial extent. This implies that e.g. nonlinearities in dendritic integration are ignored. The capacity $C$ and resistance $R$ are assumed constant, as well as the threshold for producing an action potential and the subsequent refractory period. Also, the spikes produced by the IAF-neuron are unstructured, thus it cannot produce *bursts* of action potentials as some real neurons do.

These simplifications, on the one hand, make the neuron model behave less realistic. On the other hand, computational complexity is reduced significantly, allowing for larger simulation scenarios or simulating on less powerful, and therefore cheaper, hardware. Here, as always in modeling, a compromise has to be made between *accuracy* and *simplicity* of the model. Highest possible accuracy is not always desired, as it increases computation time. In addition, sometimes the more accurate model yields details which lie outside the study's scope, and do not contribute to a deeper understanding of the problem—rather is it possible that too many details make it harder to see the effects one wants to investigate. A more simplistic model may be easier to handle, and produce results of sufficient accuracy, if it has been

well designed and the relevant properties of the system under investigation have not been "simplified away".

### Other neuron models

Of course, the IAF-neuron is not the only neuron model out there. Simpler models for example do not use spikes but rather operate on a *rate code*. One of the simplest neurons, the *perceptron*, may not have much in common with a real neuron, and may not even have been designed as such in the the first place. Nevertheless, a lot of sensible computation can be done with it, and it is employed in neuroscience, for instance in McClelland & Rogers (2003). The advantage clearly is that a single neuronal unit does not require much computation power, and interactions such as learning are relatively simple to implement. Therefore it is relatively easy to build complex networks.

More complex model neurons include so-called *compartment models*. Here, the complete neuron including the dendrites is modeled by splitting it into compartments. Compartment models actually can deliver very accurate predictions for the behaviour of real neurons (see e.g. Traub et al., 1999). But since the complexity of one neuronal unit is very high, network simulations with compartment neurons can be very time-consuming. Moreover, compared to the IAF-model there are a lot more neuron parameters for which a realistic or functional correct setting has to be found, requiring even more time.

### Why the Lapicque neuron is good enough for the present work

Generally one may argue that the Lapicque model underestimates, rather than overestimates, the computational capabilities of its biological archetype. The simulation in this work exploits the coincidence detection properties of a neuron. Since this is already present in the IAF-model, there is no need for a more complicated model. However, further studies might require more realistic models, e.g. if dendritic structure shall be included in the network architecture, or if the network shall be susceptible to learning.

# References


Abeles M (1982a): *Local Cortical Circuits: An Electrophysiological Study*. Berlin, Heidelberg, New York: Springer-Verlag.

Abeles M (1982b): Role of cortical neuron: integrator or coincidence detector? *Israel Journal of Medical Sciences*, 18: 83–92.

Abeles M (1991): *Corticonics*. 1st edition. Cambridge: Cambridge University Press.

Abeles M, Bergman H, Margalit E & Vaadia E (1993): Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *Journal of Neurophysiology*, 70(4): 1629–1638.

Bar M (2003): A cortical mechanism for triggering top-down facilitation in visual object recognition. *Journal of Cognitive Neuroscience*, 15(4): 600–609.

Calkins D J (2001): Seeing with s-cones. *Progress in Retinal Eye Research*, 20: 255–287.

Delorme A, Richard G & Fabre-Thorpe M (1999): Rapid processing of complex natural scenes: A role for the magnocellular visual pathways? *Neurocomputing*, 26–27: 663–667.

Diesmann M (2002): *Conditions for Stable Propagation of Synchronous Spiking in Cortical Neural Networks: Single Neuron Dynamics and Network Properties*. Ph.D. thesis, Fakultät für Physik und Astronomie, Ruhr-Universität Bochum.

Diesmann M & Gewaltig M O (2002): NEST: An environment for neural systems simulations. In V Macho (editor), *Forschung und wissenschaftliches Rechnen, GWDG-Bericht*. Göttingen: Ges. für Wiss. Datenverarbeitung. In press.

Diesmann M, Gewaltig M O & Aertsen A (1996): Characterization of synfire activity by propagating "pulse packets". In J Bowers (editor), *Computational Neuroscience: Trends in Research*, pages 59–64. San Diego, CA: Academic Press.

Diesmann M, Gewaltig M O & Aertsen A (1999): Stable propagation of synchronous spiking in cortical neural networks. *Nature*, 402: 529–533.

Ding Y & Casagrande V A (1998): Synaptic and neurochemical characterization of parallel pathways to the cytochrome oxidase blobs of primate visual cortex. *Journal of Comparative Neurology*, 391: 429–443.

Fabre-Thorpe M, Delorme A, Marlot C & Thorpe S (2001): A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience*, 13(2): 171–180.

Friedman H S, Zhou H & von der Heydt R (2003): The coding of uniform colour figures in monkey visual cortex. *Journal of Physiology*, 548(Pt 2): 593–613.

Fukada Y, Sumimoto I, Sugitani M & Iwama K (1979): Receptive-field properties of cells in the dorsal part of the albino rat's lateral geniculate nucleus. *Japanese Journal of Physiology*, 29: 283–307.

Gegenfurtner K R & Rieger J (2000): Sensory and cognitive contributions of color to the perception of natural scenes. *Current Biology*, 10: 805–808.

Gewaltig M O (2000): *Evolution of synchronous spike volleys in cortical networks: Network simulations and continuous probabilistic models*. Berichte aus der Physik. Aachen: Shaker.

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

Greschner M, Bongard M, Ruhan P & Ammermüller J (2002): Retinal ganglion cell synchronization by fixational eye movements improves feature estimation. *Nature Neuroscience*, 5(4): 341–347.

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

Hubel D H & Wiesel T N (1962): Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160: 106–154.

Irvin G E, Norton T T, Sesma M A & Casagrande V A (1986): W-like response properties of intralaminar zone cells in the lateral geniculate nucleus of a primate (galago crassicaudatus). *Brain Research*, 362: 254–270.

Jack J J B, Noble D & Tsien R W (1985): *Electric Current Flow in Excitable Cells*. Oxford: Clarendon.

Jacobs G H, Neitz M & Neitz J (1996): Mutations in s-cone pigment genes and the absence of colour vision in two species of nocturnal primate. *Proceedings of the Royal Society London B Biological Sciences*, 263: 705–710.

Jaeger H (2001): The "echo state" approach to analysing and training recurrent neural networks. GMD technical report 148, German National Research Center for Information Technology.

Kandel E R, Schwartz J H & Jessel T M (2000): *Principles of Neural Science*. 4th edition. McGraw-Hill.

Komatsu H, Murakami I & Kinoshita M (1996): Surface representation in the visual system. *Cognitive Brain Research*, 5: 97–104.

Körner U (2001): The koniocellular pathway of the visual system. Internal Report 01/13, Honda R&D Europe (Deutschland), Future Technology Research, 63073 Offenbach/Main, Germany.

Lapicque L (1907): Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarisation. *Journal de Physiologie et de Pathologie générale*, 9: 620–635.

Li W & Gilbert C D (2002): Global contour saliency and local colinear interactions. *Journal of Neurophysiology*, 88(5): 2846–2856.

Maas W, Natschläger T & Markram H (2002): Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation*, 14(11): 2531–2560.

Mainen Z F & Sejnowski T J (1995): Reliability of spike timing in neocortical neurons. *Science*, 268: 1503–1506.

Martin P R, White A J, Goodchild A K, Wilder H D & Sefton A E (1997): Evidence that blue-on cells are part of the third geniculocortical pathway in primates. *European Journal of Neuroscience*, 9: 1536–1541.

McClelland J L & Rogers T T (2003): The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, 4: 310–322.

Mohns M (2000): *Interaction of Synchronous Spiking and Ongoing Activity in Cortical Networks*. MSc thesis, Albert-Ludwigs-Universität, Freiburg im Breisgau, Germany.

Morand S, Thut G, de Peralta R G, Clarke S, Khateb A, Landis T & Michel C M (2000): Electrophysiological evidence for fast visual processing through the human koniocellular pathway when stimuli move. *Cerebral Cortex*, 10: 817–825.

Nayar S K, Nene S A & Murase H (1996): Real-time 100 object recognition system. In *Proc. of ARPA Image Understanding Workshop*. Palm Springs.

Nicholls J G, Wallace B G, Martin A R & Fuchs P A (2001): *From Neuron to Brain*. 4th edition. Sinauer Associates.

Papoulis A (1977): *Signal Analysis*. New York, NY, USA: McGraw-Hill.

Reid R & Alonso J (1995): Specifity of monosynaptic connections from thalamus to visual cortex. *Nature*, 378: 281–284.

Reinagel P & Reid R C (2002): Precise firing events are conserved across neurons. *Journal of Neuroscience*, 22(16): 6837–6841.

Rodemann T (2003): *Information Processing with Spiking Neurons Under The Control of Gamma Oscillations*. Ph.D. thesis, Honda Research Institute Europe, Offenbach, Germany. To be published by Cuvillier Verlag, Göttingen.

Rodieck R W (1998): *The First Steps in Seeing*. Sinauer Associates.

Rotter S & Diesmann M (1999): Exact digital simulation of time-invariant digital systems with applications to neuronal modeling. *Biological Cybernetics*, 81: 381–402.

Schmuker M, Körner U, Körner E, Gewaltig M O & Wachtler T (2003a): A model of rapid surface detection in primate visual cortex. In H Bülthoff, K Gegenfurtner, H Mallot, R Ulrich & F Wichmann (editors), *Beiträge zur 6. Tübinger Wahrnehmungskonferenz*, page 161. Kirchentellinsfurt: Knirsch.

Schmuker M, Körner U, Körner E, Gewaltig M O & Wachtler T (2003b): A model of rapid surface detection in primate visual cortex. In N Elsner & H Zimmermann (editors), *The Neurosciences from Basic Research to Therapy: Proceedings of the 29th Göttingen Neurobiology Conference*, page 666. Stuttgart: Thieme.

Shadlen M & Movshon J (1999): Synchrony unbound: a critical evaluation of the temporal binding hypothesis. *Neuron*, 24: 67–77.

Shadlen M & Newsome W (1994): Noise, neural codes and cortical organization. *Current Opinion in Neurobiology*, 4: 569–579.

Shapley R, McLaughlin D & Shelley M (2003): Orientation selectivity. In M A Arbib (editor), *The Handbook of Brain Theory and Neural Networks*, 2nd edition, pages 831–835. Cambridge, MA: MIT Press.

Silveira L C, Lee B B, Yamada E S, Kremers J, Hunt D M, Martin P R & Gomes F L (1999): Ganglion cells of a short-wavelength-sensitive cone pathway in new world monkeys: morphology and physiology. *Visual Neuroscience*, 16: 333–343.

Slovin H, Arieli A, Hildesheim R & Grinvald A (2002): Long-term voltage-sensitive dye imaging reveals cortical dynamics in behaving monkeys. *Journal of Neurophysiology*, 88: 3421–3438.

Smith G D & Sherman S M (2002): Detectability of excitatory versus inhibitory drive in an integrate-and-fire-or-burst thalamocortical relay neuron model. *The Journal of Neuroscience*, 22(23): 10242–10250.

Stevens C F & Zador A M (1998): Input synchrony and the irregular firing of cortical neurons. *Nature Neuroscience*, 1(3): 210–217.

Tanaka K (1983): Cross-correlation analysis of geniculostriate neuronal relationships in cats. *Journal of Neurophysiology*, 49: 1303–1318.

Tani T, Yokoi I, Ito M, Tanaka S & Komatsu H (2003): Functional organization of the cat visual cortex in relation to the representation of a uniform surface. *Journal of Neurophysiology*, 89: 1112–1125.

Thorpe S, Delorme A & Van Rullen R (2001): Spike-based strategies for rapid processing. *Neural Networks*, 14: 715–725.

Thorpe S, Fize D & Marlot C (1996): Speed of processing in the human visual system. *Nature*, 381: 520–522.

Thorpe S J & Imbert M (1989): Biological constraints on connectionist models. In R Pfeifer & F Fogelman-Soulié (editors), *Connectionism in perspective*, pages 63–92. Amsterdam: Elsevier.

Traub R D, Jefferys J G R & Whittington M A (1999): *Fast oscillations in cortical circuits*. Cambridge, MA: MIT Press.

Tuckwell H C (1988): *Introduction to theoretical neurobiology, Vol. 1: Linear cable theory and dendritic structure*. Cambridge University Press.

Ullman S (1995): *High-Level Vision*, chapter 8. Cambridge MA: MIT Press.

Usrey W M (2002): The role of spike timing for thalamocortical processing. *Current Opinion in Neurobiology*, 12: 411–417.

Van Rullen R & Thorpe S J (2001): Rate coding versus temporal order coding: What the retinal ganglion cells tell the visual cortex. *Neural Computation*, 13: 1255–1283.