

# A mechanistic cortical microcircuit of attention for amplification, normalization and suppression

Frederik Beuth, Fred H Hamker\*

*Chemnitz University of Technology, Artificial Intelligence, Strasse der Nationen 62, D - 09111 Chemnitz, Germany*

---

## Abstract

Computational models of visual attention have replicated a large number of data from visual attention experiments. However, typically each computational model has been shown to account for only a few data sets. We developed a novel model of attention, particularly focused on explaining single cell recordings in multiple brain areas, to better understand the underlying computational circuits of attention involved in spatial- and feature-based biased competition, modulation of the contrast response function, modulation of the neuronal tuning curve, and modulation of surround suppression. In contrast to previous models, we use a two layer structure inspired by the layered cortical architecture which implements amplification, divisive normalization and suppression as well as spatial pooling.

*Keywords:* Attention models, biased competition, scaling vs. sharpening of tuning curves, contrast vs. response gain, surround suppression

---

## 1. Introduction

Attention is one of the fields in vision research that has been strongly influenced by models. While early models have been rather abstract, over the last 15 years, inspired by data from neural recordings, the neural correlates of attention have been increasingly addressed by neuro-computational models. Although attention operates across multiple brain areas and may help to bind stimulus properties processed in different parts of the brain, one important line of research in visual attention aims at understanding the local circuits of neural interactions. A very influential milestone has been the discovery of interactions between representations of stimuli that are placed within a common receptive field - coined as biased competition (Desimone & Duncan, 1995; Desimone, 1998).

According to this concept, attention should not be understood as a simple mechanism for gating or an enhancement of neural responses but rather as a competition for

---

\*Corresponding author: Fred H Hamker, phone: +49 (0)371 531-37875  
*Email addresses:* beuth@cs.tu-chemnitz.de (Frederik Beuth),  
fred.hamker@informatik.tu-chemnitz.de (Fred H Hamker)

neural representation which is typically strong if neurons in a single area have spatially overlapping receptive fields. Thus, if attention is directed away, the neural response to a single stimulus is typically larger than the response to two different stimuli, which suggests that the total stimulus energy is not additive. Attention towards a particular location in space or towards a stimulus feature implements a top-down signal which changes the properties of local interactive circuits almost as if the attended stimulus would be presented alone, i.e. the response of neurons tuned to the attended stimulus is amplified and the response of neurons tuned to the not attended one is reduced. Neural recordings in different brain areas have supported this basic concept (Desimone & Duncan, 1995; Chelazzi et al., 1998, 2001; Fallah et al., 2007; Motter, 1993; Lee & Maunsell, 2009, 2010a,b; Reynolds et al., 1999; Treue & Trujillo, 1999; Zhou & Desimone, 2011). Other observations of the neural effect of attention revealed multiplicative scaling (McAdams & Maunsell, 1999; David et al., 2008; Motter, 1993) or sharpening (Martínez-Trujillo & Treue, 2004; David et al., 2008) effects on the neural tuning curve. Moreover, several studies observed that attention typically leads to a shift in the contrast response function (Reynolds et al., 2000; Martínez-Trujillo & Treue, 2002; Li et al., 2008) which suggests that well visible stimuli are not further boosted. A large number of neuro-computational models have been developed and demonstrated to account for parts of these data (Reynolds & Heeger, 2009; Boynton, 2009; Compte & Wang, 2006; Hamker, 2004, 2005b; Lee & Maunsell, 2009; Ni et al., 2012; Spratling & Johnson, 2004; Spratling, 2008; Wagatsuma et al., 2013; Ardid et al., 2007; Buia & Tiesinga, 2008; Hugues & José, 2010). Basically, each of these models includes lateral or feedforward inhibition and some form of attentive gain increase. Further experimental studies shed more light on the nature of stimulus interactions, e.g. by probing the neural responses to two identical stimuli (Lee & Maunsell, 2009, 2010a) or led to conflicting observations of whether attention shifts or scales the contrast response function (Martínez-Trujillo & Treue, 2002), which cumulated into normalization models of attention (Lee & Maunsell, 2009; Reynolds & Heeger, 2009). Although these models are not completely different than their predecessors, they propose two essential components that determine the final population response, an attentive drive that scales the response multiplicatively and a suppressive drive that operates after the attentive drive within a local region and acts divisively on the neural response. The combination of these two components allowed to account for modeling competitive interactions of two identical stimuli (Lee & Maunsell, 2009, 2010a) and to solve the apparent conflict whether attention shifts or scales the contrast response function (Reynolds & Heeger, 2009).

Although Reynolds & Heeger (2009) demonstrated that their model can account for several data, existing models of attention have so far shown to account only for a very selective set of data. However, a particular strength of computational models is their ability to reveal a few potential underlying computational mechanisms that may account for a large variety of data. Thus, from the computational point of view it is important to find out how much data can be explained by a small set of computational mechanisms. The mechanisms of attention can be explained by different levels of implementation detail, e.g. detailed biophysical neuron models and their microcircuits, abstract spiking neurons, dynamic rate coded populations or more abstract mathematical descriptions. With respect to the available data we here developed a cor-

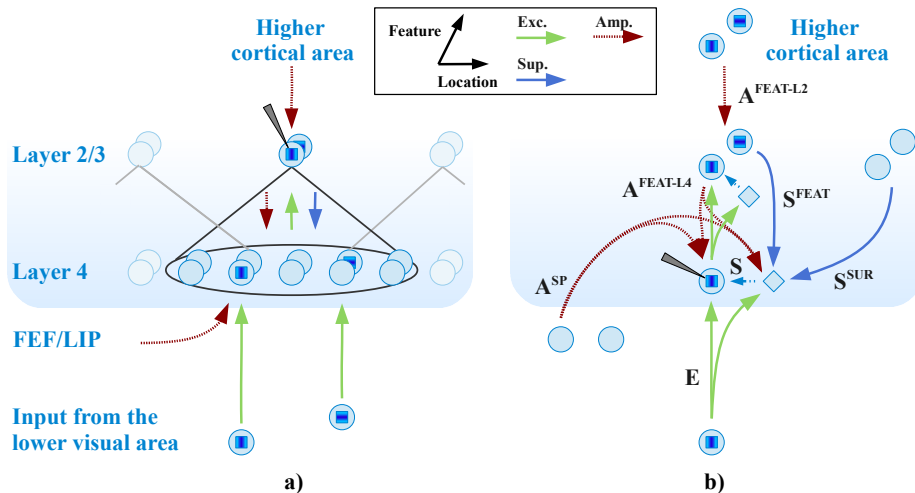


Figure 1: **a)** Proposed mechanistic cortical microcircuit of attention, consisting of a modulatory layer 4 and a spatial pooling layer 2/3. Feature-based attention can be deployed to layer 2/3 via an amplifying signal originating from a higher cortical area, e.g. PFC or IT. Spatial attention can be deployed to layer 4 via an amplifying signal originating from cortices encoding spatial information, e.g. FEF or LIP. **b)** Connectivity and influences centered on a single cell in layer 4, indicated by the electrode symbol. The cell receives feed-forward excitation ( $E$ ), feature-based amplification ( $A^{\text{FEAT-L4}}$ ) from layer 2/3, spatial amplification ( $A^{\text{SP}}$ ) from spatial cortices such as FEF or LIP, and suppression from an associated interneuron ( $S$ ). The interneuron receives several sources of suppression: the feedforward excitation of its associated neuron ( $E$ ), dissimilar features in layer 2/3 ( $S^{\text{FEAT}}$ ), and all similar features in the surround in layer 2/3 ( $S^{\text{SUR}}$ ).

tical microcircuit of attention using dynamic rate coded neural populations inspired by properties of the cortical microcircuit (Douglas & Martin, 2004). This leaves the exact biophysical implementation unanswered but allows to reveal the main computational mechanisms that may similarly operate in multiple brain areas regardless of their exact implementation. Although we mainly address data from neuronal recordings, the chosen level of implementation detail allows to discuss and better understand the influence of attention in psychophysical experiments, such as the effect of spatial and feature-based attention, target selection, and distractor inhibition.

## 2. Methods

The mechanistic cortical microcircuit of attention is an attempt to account for a very large dataset and thus to reveal the essential aspects of attention common in multiple brain areas. We identify three core mechanisms necessary for this approach: amplification, normalization and suppression. A set of Matlab routines implementing the model can be downloaded at: [www.tu-chemnitz.de/cs/KI/supplement/BeuthHamker2015/](http://www.tu-chemnitz.de/cs/KI/supplement/BeuthHamker2015/).

### 2.1. Model overview and structure

The model (Fig. 1a) is structured in a modulating (cortical layer 4) and in a pool (cortical layer 2/3) layer. A neuron in this model abstracts from a biophysical descrip-

tion of a cell and rather refers to a population of interconnected real cells. Similarly, gain amplifications are described by multiplicative operations. This level of description is similar to proposed concepts of normalization (Carandini & Heeger, 2012) and the normalization model of attention (Reynolds & Heeger, 2009), although we will suggest in more detail how abstract concepts of normalization and suppression may be implemented in a simple cortical microcircuit of attention. A neuron in each layer encodes a certain feature at a particular (one-dimensional) location, so that a layer of neurons is represented by a two dimensional matrix. The mechanistic cortical microcircuit of attention has three inputs: layer 4 receives input from lower visual areas, from areas that encode space such as the frontal eye field (FEF) and the lateral intraparietal cortex (LIP), and layer 2/3 receives a signal from a feature selective higher cortical area. The model mainly refers to physiological properties of area V4 and MT. However, as attention is a canonical property of the visual cortex (Saenz et al., 2002), the model is not restricted to these areas and may account also for other visual areas.

The tuning curve of a neuron has to be determined by the user to match the particular characteristics of a brain area as tuning curves largely differ across visual areas. Layer 4 is denoted as modulating layer as its neuronal responses are calculated by amplifying or suppressing the excitation given from the input. Neurons in layer 4 converge onto neurons in layer 2/3 which spatially pool over them in a broad area (a similar idea as in Hamker & Zirnsak (2006); Compte & Wang (2006)). Therefore layer 2/3 neurons have larger receptive fields than layer 4 neurons. The receptive fields in layer 2/3 are overlapping by a user defined amount (standard is 50% to each side), so a stimulus will fall within several receptive fields. Dependent on the amount of overlap and the receptive field sizes, layer 2/3 contains typically less cells in the spatial dimension than layer 4. In all experiments, we report the responses of cells in layer 2/3 (also indicated by the electrode symbol in each figure).

## 2.2. Divisive normalization

The neural responses in the model largely follow concepts of divisive normalization (Carandini & Heeger, 2012; Reynolds & Heeger, 2009). As a concretization of the original concept used by Reynolds & Heeger (2009), we propose an implementation via interneurons. In this implementation, each pyramidal neuron receives via an inhibitory connection a normalization signal from an associated interneuron (Fig. 1b,  $S$ ), which normalizes the response of the pyramidal neuron. The interneuron receives suppressive influences ( $S^{\text{FEAT}}$ ,  $S^{\text{SUR}}$ ) via excitatory connections plus the feedforward excitation of the associated pyramidal neuron. This is supported by the data from Mitchell et al. (2007), who found that attention modulates interneurons twice as strongly as pyramidal cells, and by the anatomical meta-analysis from Potjans & Diesmann (2012) reporting the required inhibitory, intra-layer connections from an interneuron to a pyramidal neuron, and the necessary excitatory connections from an pyramidal neuron to an interneuron or another pyramidal neuron.

## 2.3. Modeling of attention via amplification

Attention is thought to emerge from recurrent connectivity within the visual and prefrontal cortex (Hamker, 2005b). This computational framework proposes that neurons receive feedback from higher cortical areas which acts as an amplifying signal.

To distinguish between mechanism and result, we will use “attention” to describe the psychological effect and “amplification” to refer to a particular mechanism at the single neuron.

We assume that spatial attention implies a spatial amplification signal from spatial maps in the brain, such as the FEF (Hamker, 2005b) or the LIP (Steenrod et al., 2013), denoted in the model as  $A^{SP}$  (Fig. 1b). In the model, this signal amplifies the response of neurons in layer 4 (similar as in Hamker (2005b) and Spratling (2008)) as inspired by anatomical observations, e.g. layer 4 of area V4 receives input from the FEF (Barone et al., 2000) and LIP (Ninomiya et al., 2012). By this mechanism, spatial amplification can be directed to one or multiple spatial locations within the receptive field as needed in experiments with spatial attention.

We presume that feature-based attention implies a feature-specific amplification of neuronal responses in the visual cortex, typically invariant to spatial location. This amplification can occur via an amplifying signal originating from prefrontal cortex (PFC) which is then projected backwards through the visual cortex. This assumption is consistent with the attentional control of the visual cortex as reviewed by Kastner & Ungerleider (2000) for feature-based and by Miller & Buschman (2013) for spatial and feature-based attention. Therefore, we model a feature-based amplification signal originating from a higher visual area, e.g. PFC or IT, which is projected to V4 or MT. This signal is firstly projected to layer 2/3 via excitatory connections (Fig. 1b, signal  $A^{FEAT-L2}$ ) and then afterwards sent to layer 4 (Fig. 1b, signal  $A^{FEAT-L4}$ ). We assume the projection from layer 2/3 to layer 4 occurs indirectly via the cortical excitatory pathways of layer 2/3 to layer 5 and then to layer 4. Both projections were found in the physiological connectivity data of the microcircuit (Douglas & Martin, 2004) which appears to generalize across areas in the visual cortex. In layer 4, the effects of both spatial and feature-based amplification signals are summed up additively because this behavior was observed in several areas in the visual cortex (Treue & Trujillo (1999): area MT; Saenz et al. (2002): areas V1, V2, V3A, V4 and MT+).

#### 2.4. Modeling of suppression

Suppression from other neurons (Fig. 1b, signals  $S^{FEAT}$  and  $S^{SUR}$ ) is modeled via connections from layer 2/3 to layer 4. To our knowledge, it is not known in detail which intra-area connections are responsible for transporting suppressive signals, however layer 2/3 is beneficial as starting point because feedback connections from higher areas end in layer 2/3 and feedback seems to play an essential role in suppression (Angelucci et al., 2002; Bair et al., 2003; Carandini & Heeger, 2012; Gilbert, 1998; Hunt et al., 2011). Layer 4 is favorable as end point as its neurons need to receive suppression to balance out the amplification. To keep the model simple, suppression is implemented by two separate signals, each representing a particular experimental condition: feature-based suppression within the classical receptive field ( $S^{FEAT}$ ), and surround suppression from outside the receptive field ( $S^{SUR}$ ). This facilitates the usability of the model as the amount of suppression can be fitted to a particular experiment without influencing the other condition.

## 2.5. Mathematical description of the model

### 2.5.1. Notation

Firing rates of mean-rate neurons are denoted by  $r$ :  $r^{\text{In}}$  (Input),  $r^{\text{L4}}$  (layer 4),  $r^{\text{L2}}$  (layer 2/3),  $r^{\text{PFC}}$  (higher cortical area like PFC or IT) and  $r^{\text{FEF}}$  (spatial areas like FEF or LIP). The rates  $r^{\text{In}}$ ,  $r^{\text{PFC}}$  and  $r^{\text{FEF}}$  have to be set by the user to model a particular experimental condition. Big letters denote influences towards a neuron:  $E$  denotes the excitatory influence,  $A$  the amplifying one and  $S$  the suppressive one. Each neuron is addressed within a layer by the spatial index  $x$  and the feature index  $l$ . The layer dimensions are denoted by  $X$  and  $L$  respectively. The feature dimension is modeled circular, hence the features  $l = 1$  and  $l = L$  encode adjacent features. As the model is a recurrent system, we simulate it via ordinary differential equations over time, denoted by  $t$ . All firing rates were recorded in a steady state condition after the system has converged, which typically occurs after 100ms.

### 2.5.2. Influences towards a neuron and connectivity

A neuron is influenced via several connections which are either designed as one-to-one connections or as complex connectivity patterns. A one-to-one connection connects a single presynaptic cell with a single postsynaptic cell. Both cells encode the same location and feature. Influences via one-to-one connections:

*Excitation  $E^{\text{In-L4}}$  (input  $\rightarrow$  layer 4):* Excitatory connection from input layer to layer 4 with the connection strength  $v^{\text{In-L4}}$ .

*Amplification  $A^{\text{SP}}$  (FEF/LIP  $\rightarrow$  layer 4):* Amplifying connection from a spatial cortical area, e.g. FEF or LIP, to layer 4 with the strength  $v^{\text{FEF-L4}}$ .

*Amplification  $A^{\text{FEAT-L2}}$  (higher cortical area  $\rightarrow$  layer 2):* Amplifying connection from a higher cortical area, like PFC or IT, to layer 2/3 with the strength  $v^{\text{PFC-L2}}$ .

A complex connectivity pattern connects multiple presynaptic cells to a single postsynaptic cell, thus multiple presynaptic features or locations are integrated to a single postsynaptic feature and location. We modeled such an influence by an equation like  $E_{\text{post}} = v \cdot \sum_{\text{pre}} (w_{\text{post,pre}} \cdot r_{\text{pre}})$ , using a global scaling factor  $v$  and a connectivity matrix  $w_{\text{post,pre}} \in [0, 1]$ . This modeling via two separate variables was chosen as the experimental data can be well fitted by adapting only the scaling factor  $v$  and keeping the connectivity pattern  $w$  unchanged. Influences via complex connectivity patterns:

*Excitation  $E^{\text{L4-L2}}$  (layer 4  $\rightarrow$  layer 2/3):* The connectivity pattern  $w_{x,x'}^{\text{L4-L2}}$  for each postsynaptic cell  $(x, l)$  is a Gaussian over the spatial dimension of the receptive field ( $x'$ ) with constant feature ( $l$ ). The parameters of the Gaussian are chosen so that the Gaussian is centered and that it has a value of about 0.5 at the borders of the receptive field:  $\mu = 0$ ,  $\sigma = 0.4 \cdot s$ , whereby  $s = 5$  denotes the spatial extend of the field. Connections exist only inside the receptive field, except the experiment of Cavanaugh et al. (2002a) which needs a broader connectivity pattern ( $2s$ ). The scaling factor of the excitation is fixed to 1.

*Amplification  $A^{\text{FEAT-L4}}$  (layer 2/3  $\rightarrow$  layer 4):* The connectivity pattern  $w_{x,x'}^{\text{L2-L4}}$  for each postsynaptic cell  $(x, l)$  is a Gaussian over the spatial dimension of the receptive field ( $x'$ ) with constant feature ( $l$ ). It is reciprocal modeled to  $w^{\text{L4-L2}}$ . Scaling factor:  $v^{\text{L2-L4}}$ .

*Feature-based suppression  $S^{\text{FEAT}}$  (layer 2/3  $\rightarrow$  layer 4):* The connectivity pattern  $w_{dl}^{\text{FEAT}}$  for each postsynaptic cell  $(x, l)$  is a linear function (Eq. 1) over the feature dimension of the presynaptic layer ( $l'$ ) at location ( $x$ ). It is 0 if pre- and postsynaptic features are equal ( $dl \leq L_0$ ) and 1 if they are at most dissimilar. As the feature dimension is modeled circular, the features are most dissimilar at medial distance ( $dl = L/2$ ) in the feature space  $[1, L]$ . Scaling

factor:  $v^{\text{FEAT}}$ .

$$w_{dl}^{\text{FEAT}} = \begin{cases} 0 & dl \leq L_0 \\ \frac{dl-L_0}{0.5L-L_0} & L_0 < dl \leq 0.5L \\ 1 - \frac{dl-0.5L}{0.5L-L_0} & 0.5L < dl \leq L-L_0 \\ 0 & \text{else} \end{cases} \quad (1)$$

whereby  $dl = |l - l'|$  and  $L_0 = L/8$ .

*Surround suppression*  $S^{\text{SUR}}$  (layer 2/3  $\rightarrow$  layer 4): The connectivity pattern  $w_{x,x'}^{\text{SUR}}$  for each post-synaptic cell  $(x, l)$  is a function (Eq. 2) over the spatial dimension of the presynaptic layer  $(x')$  with constant feature  $(l)$ . It is 0 for the same location, maximal (1.0) for close surround locations and moderate (0.4) for distant ones. Scaling factor:  $v^{\text{SUR}}$ .

$$w_{x,x'}^{\text{SUR}} = \begin{cases} \frac{|x-x'|-1}{\psi-1} & |x-x'| \leq \psi \\ 1 - 0.6 \frac{|x-x'|-\psi-1}{\psi-1} & \psi < |x-x'| \leq 2\psi \\ 0.4 & \text{else} \end{cases} \quad (2)$$

whereby  $\psi = (s - 1)/2$  denotes the radius of the receptive field  $s$ .

Fig. no.	E			$S^{\text{FEAT}}$		$S^{\text{SUR}}$		A		$\beta$	Tuning function
	$v^{\text{In-L4}}$	$pE$	$p^{\text{Pool}}$	$v^{\text{FEAT}}$	$p^{\text{FEAT}}$	$v^{\text{SUR}}$	$p^{\text{SUR}}$	$v^{\text{FEF-L4}}$	$v^{\text{PFC-L2}}$		
Std.	3	2	4	3	2	0.5	1	3	0.5	0	$c = \frac{8}{L}, a = 0$
2				2.5				2		0.15	$c = \frac{6}{L}$
3a				2.75				1			$c = \frac{6}{L}$
3b				3.5							$c = \frac{8}{L}, a = 0.15$
3c-e				3.5	3			1.25			$c = \frac{8}{L}, a = 0.1$
3f	3.5			2.5						0.05	$c = \frac{8}{L}, a = 0.15$
4					3				0.17	0.15	$c = \frac{8}{L}, a = 0.1$
5	2	1.75									$c = \frac{8}{L}$
6						1					$c = \frac{8}{L}$
7	2	1.5				0.75				0.1	$c = \frac{8}{L}, a = 0.1$
8a				2	3						$c = \frac{8}{L}, a = 0.3$ , Amplification: $c = \frac{4}{L}$
8b				2.4	3.5				0.25		$c = \frac{6}{L}$
9a	0.5	0.75	2			0.8	2				$c = \frac{8}{L}$
9b	1.6	1.25				0.55	2.5				$c = \frac{8}{L}$
10											$c = \frac{8}{L}$

Table 1: Model parameters which were varied between experiments are listed in the table. The second row “std” denotes the standard values which were used in all experiments if not otherwise denoted. The input is modeled via a symmetrical exponential function in the feature space with width  $c$  and baseline  $a$ :  $f(x) = a + (1 - a) \cdot e^{(-c \cdot x)}$ . If not otherwise noted, the amplification is modeled via the same parametrized function, but without the baseline ( $a = 0$ ). The model contains also an unmodulated baseline activity  $\beta$  to simulate conditions in which attention does not modulate the response.

### 2.5.3. Neuronal layers

The neuronal responses are simulated by the following equations. The values of parameters varied across experiments are listed in Tab. 1, and the ones kept constant are the following:  $\sigma^{\text{L4}} = 0.3$ ,  $\sigma^{\text{L2}} = 2$ ,  $v^{\text{L2-L4}} = 1$ ,  $\tau = 10ms$ .

#### Layer 4

$$\tau \frac{\partial r_{x,l}^{L4}}{\partial t} = g \cdot \frac{E_{x,l} \cdot A_{x,l}}{\sigma^{L4} + S_{x,l}} - r_{x,l}^{L4} \quad (3)$$

$$S_{x,l} = E_{x,l} \cdot A_{x,l} + b(E_{x,l}) \cdot S_{x,l}^{\text{SUR}} + b(E_{x,l}) \cdot S_{x,l}^{\text{FEAT}} \quad (4)$$

$$A_{x,l} = 1 + A_{x,l}^{\text{SP}} + A_{x,l}^{\text{FEAT-L4}} \quad (5)$$

$$E_{x,l} = \left( v^{\text{In-L4}} \cdot r_{x,l}^{\text{In}} \right)^{p^E} \quad (6)$$

$$\text{With: } A_{x,l}^{\text{SP}} = v^{\text{FEF-L4}} \cdot r_x^{\text{FEF}} \quad (7)$$

$$\text{With: } A_{x,l}^{\text{FEAT-L4}} = v^{\text{L2-L4}} \cdot f_1 \left( \sum_{x'} f_2 \left( w_{x,x'}^{\text{L2-L4}} \cdot r_{x',l}^{\text{L2}} \right) \right) \quad (8)$$

$$\text{With: } S_{x,l}^{\text{SUR}} = \sum_{x'} w_{x,x'}^{\text{SUR}} \cdot \left[ v^{\text{SUR}} \cdot r_{x',l}^{\text{L2}} \right]^{p^{\text{SUR}}} \quad (9)$$

$$\text{With: } S_{x,l}^{\text{FEAT}} = \sum_{l'} w_{l,l'}^{\text{FEAT}} \cdot \left[ v^{\text{FEAT}} \cdot f_1 \left( \sum_{x'} f_2 \left( w_{x,x'}^{\text{L2-L4}} \cdot r_{x',l'}^{\text{L2}} \right) \right) \right]^{p^{\text{FEAT}}} \quad (10)$$

$$\text{With: } b(E_{x,l}) = E_{x,l} \cdot \left( \frac{\sigma^{L4} + E_{x,l}}{g \cdot E_{x,l}} \right)^2 \quad (11)$$

Whereby  $b(E_{x,l})$  represents a correction function based on the excitation ( $E_{x,l}$ ) and the inverted divisive normalization function without modulation, i.e without amplification and suppression (terms in brackets). The function ensures that suppression causes the same proportional decrease of the response for all stimulus contrasts. The parameters  $\sigma^{L4}$ ,  $v^{\text{In-L4}}$  and  $p^E$  control the shape of the contrast response function similar as in Carandini & Heeger (2012). The parameters  $p^{\text{FEAT}}$  and  $p^{\text{SUR}}$  control the non-linearity behavior of feature-based and surround suppression.

The factor  $g$  normalizes the firing rates to a maximum of 1 as in other divisive normalization approaches (Albrecht & Hamilton, 1982; Carandini & Heeger, 2012). It should be chosen as:  $g = 1 + \sigma / (A \cdot E')$  whereby  $A \cdot E'$  denotes the maximum possible excitation (here 1).

#### Layer 2/3

$$\tau \frac{\partial r_{x,l}^{\text{L2}}}{\partial t} = g \cdot \frac{E_{x,l} \cdot \left( 1 + A_{x,l}^{\text{FEAT-L2}} \right)}{\sigma^{\text{L2}} + S_{x,l}} - r_{x,l}^{\text{L2}} \quad (12)$$

$$S_{x,l} = E_{x,l} \cdot \left( 1 + A_{x,l}^{\text{FEAT-L2}} \right) \quad (13)$$

$$\text{With: } E_{x,l} = f_1 \left( \sum_{x'} f_2 \left( w_{x,x'}^{\text{L4-L2}} \cdot r_{x',l}^{\text{L4}} \right) \right) \quad (14)$$

$$\text{With: } A_{x,l}^{\text{FEAT-L2}} = v^{\text{PFC-L2}} \cdot r_l^{\text{PFC}} \quad (15)$$

$$\text{With: } f_1(x) = \frac{p^{\text{Pool}}}{4} \cdot x^{\frac{1}{p^{\text{Pool}}}} \text{ and } f_2(x) = x^{p^{\text{Pool}}} \quad (16)$$

Whereby  $f_1(x)$  and  $f_2(x)$  represent non-linearities allowing the non-linear summation necessary for a spatial pooling operation. The pooling operation is modeled after the pooling and surround suppression data of Bonin et al. (2005); Heuer & Britten (2002); Hunter & Born (2011); Jones et al. (2002); Pack et al. (2005). The non-linearity parameter is chosen as  $p^{\text{Pool}} = 4$ .

Baseline activity can be modulated by attention like in the sharpening of neuronal tuning curves (no. 8a in Tab. 1), but is also often unaffected by attention. To model the former case, we add a small value to the input (parameter  $a$ ). For the latter case, we add a small value (parameter  $\beta$ ) to the firing rates after the differential equation system has been converged (Eq. 17, 18). Our



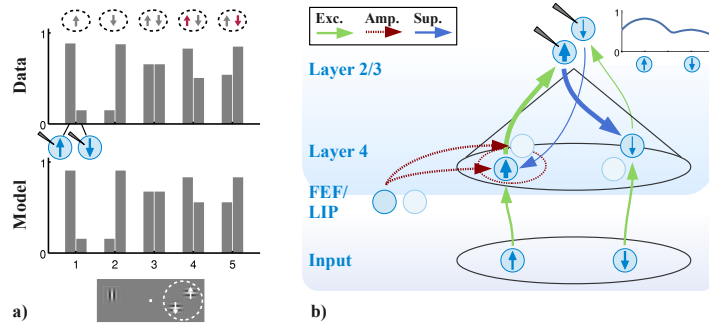


Figure 2: Biased competition experiment with spatial attention. **a)** Neurophysiological data (top) in relation to our simulation results (bottom). Data is adapted from Lee & Maunsell (2010a, Fig5). In each column, the left bar denotes the response of a neuron preferring the upward moving stimulus and the right bar the downward direction. Spatial attention is either directed away (columns 1 - 3) or to the location of the stimulus marked in red (columns 4 and 5). **b)** Model mechanism of biased competition, within the condition of spatially attending the left location (condition 4 in a). In each layer, exemplary neurons reacting for one of the two stimuli are shown. In addition, the full population response of layer 2/3 at the recorded location is illustrated at the top right. The connection type (excitation, amplification, suppression; see Section 2) describes the influence on its postsynaptic neuron. Attentional modulation of neurons and connections are denoted by the increase or decrease of the symbols thickness. For clarity, interneurons and unrelated connections are not illustrated in the figure.

approach is similar as in the normalization model of attention (Reynolds & Heeger, 2009) to facilitate a comparison of the data fits.

$$r^{L4} = \beta + (1 - \beta) \cdot r^{L4} \quad (17)$$

$$r^{L2} = \beta + (1 - \beta) \cdot r^{L2} \quad (18)$$

### 3. Results

In the following, we will identify essential mechanisms of our model for each attention experiment. We see these mechanisms as a proposal of how attention might work in the cortical microcircuit of the visual cortex. In each experiment, we first verify the model by comparing its simulated neuronal responses to physiological data and then link the simulated data to the mechanisms in the model. Further, we will predict attention mechanisms in novel paradigms, e.g. in biased competition with surround suppression.

#### 3.1. Biased competition

In biased competition paradigms, suppression is observed between two different stimuli presented inside a receptive field of a neuron and attention has been observed to modulate this suppression. Biased competition with spatial attention has been reported in several areas (MT: Lee & Maunsell (2010a), also Lee & Maunsell (2009, 2010b); V2 and V4: Reynolds et al. (1999); V1, V2, V4: Motter (1993)). In Lee & Maunsell (2010a), two inversely moving gratings were placed in the receptive field of a neuron

in area MT (Fig. 2a). Attention was either spatially directed to a grating placed in the opposite hemifield (attend away) or to one of the two stimuli. They observed that an unattended stimulus alone elicits a much stronger neuronal response in its preferred neuron (column 1 left) as the unattended stimulus pair (column 3). However, if the preferred stimulus of the pair (column 4) is attended, the response of the preferred cell will be as strong as to the stimulus alone (column 4, left) and the response of the anti-preferred cell will be decreased (column 4, right). The reverse effect is observable for attending the anti-preferred stimulus (column 5). The data is adapted from Lee & Maunsell (2010a, Fig. 5) and shows the average firing rate over the full population and over a time window between 50ms and 200ms after stimulus onset.

In the model (Fig. 2b), spatial attention is modeled by an amplification signal originating from FEF/LIP to layer 4, amplifying the response of all cells at the attended location. This location comprises only the attended stimulus, here the upward moving one, so only the response of this feature is increased. This drives layer 2/3 cells, resulting also there in an amplified response of this feature as observed in the data (column 4, left in Fig. 2a). These neurons in layer 2/3 suppress cells in layer 4 encoding the anti-preferred stimulus (downward motion), denoted as feature-based suppression. This in turn decreases also the response of cells in layer 2/3 encoding the anti-preferred stimulus as observed in model and data (column 4, right in Fig. 2a).

For simulating biased competition with spatial attention, the spatial pooling operation in the second layer is necessary. This operation is one of the main model improvements as previous models contain only a single layer (Boynton, 2009; Lee & Maunsell, 2009; Ni et al., 2012; Reynolds & Heeger, 2009) as discussed in more detail later. Pooling mediates the interaction of the attended location with the recorded location. The recorded location is here defined as the center of the receptive field, because the model uses symmetric receptive fields. The attended location is classically in the periphery of the receptive field. Therefore, spatial pooling is required to transfer the attentional modulated response from the attended location to the center. In addition, spatial pooling is accompanied by a distribution of the suppression. This is implemented via the suppressive feedback connections from layer 2/3 to layer 4 originating from the center to all locations within the receptive field. Both mechanisms together allow that a spatial amplification at one location induces a suppression at the other location.

Thus in summary, necessary mechanisms for biased competition are the amplification in layer 4 and in layer 2/3, the suppression of the anti-preferred feature, and the spatial pooling.

Biased competition has been traditionally tested with different stimuli raising the question what effect occurs if identical stimuli are placed within the receptive field. Recent data suggests that a “feature pooling” effect occurs instead of biased competition. The effect without attention is shown by MacEvoy et al. (2009) for area V1: no suppression was observed between equal stimuli, i.e. the response to a single stimulus was the same as presenting both stimuli (Fig. 3b, right). In the study, they varied systematically the similarity between the stimuli and showed that the amount of suppression decreases with increasing similarity (Fig. 3b). Yet it has to be noted that they used overlapping stimuli whereby the stimuli in biased competition are normally spatially distinct. Nevertheless, the pooling effect also occurs with deploying attention to one of two spatially distinct stimuli in area MT: Lee & Maunsell (2010a) reported for two

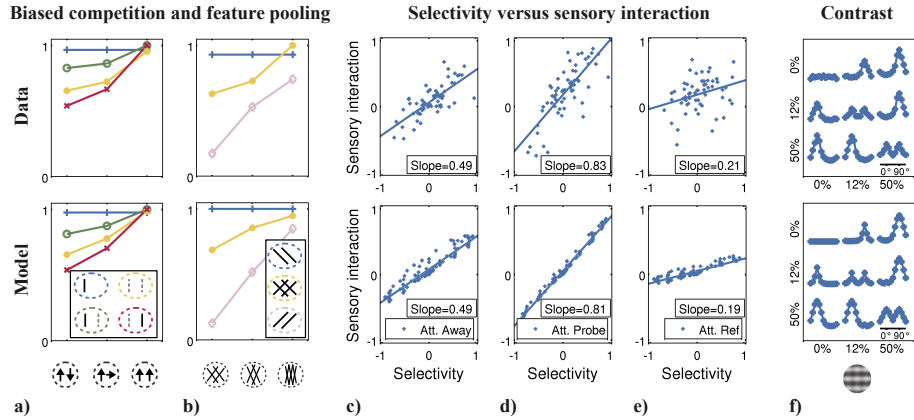


Figure 3: Characteristics of biased competition and related paradigms. Neurophysiological data (top row) is shown in relation to simulation results (bottom row). **a)** Neuronal response depends on the similarity of the two spatially disjunctive stimuli. Biased competition (two different stimuli, left) and feature pooling (two equal stimuli, right) mark the extreme cases. Experimental conditions are: preferred stimulus alone attended (blue), pair attend preferred (green), pair attend away (yellow) and pair attend anti-preferred (red). The data shows the average population response in area MT of cells preferring the left stimulus and is adapted from Lee & Maunsell (2010a, Fig. 5). **b)** The same similarity dependency as in a) can be observed also for two overlapping gratings in area V1. In this study (MacEvoy et al., 2009, Fig. 2a-c), conditions are all without attention: preferred stimulus alone (blue), pair (yellow) and anti-preferred stimulus alone (pink). **c-e)** Effect of attention on the relationship between selectivity (probe - reference) and sensory-interaction (pair - reference). See text for details. The data represents a population response in area V4 in a biased competition paradigm with spatial attention (Reynolds et al., 1999, Fig. 11a,b,d). **f)** Contrast dependency of responses in area V1 with overlapping competing stimuli (Busse et al., 2009, Fig. 4a). The contrast of two stimuli ( $0^\circ$  and  $90^\circ$  oriented gratings) is systematically varied (0%, 12%, 50%) and the population response is plotted. Contrasts of horizontal gratings ( $90^\circ$ ) are denoted at the x-axis and contrasts of vertical gratings ( $0^\circ$ ) at the y-axis.

equal stimuli a similar attentional modulation as for a single stimulus (Fig. 3a right). The effect could not be induced by the experimental setup as the stimulus similarity has been systematically varied within the same setup, and biased competition has been found for different stimuli and pooling for identical stimuli. Our model reflects this effect as the strength of the suppressive connections from layer 2/3 to layer 4 depends on the feature similarity (Fig. 1), so the strength is maximal for different ones and zero for similar features.

A possibility to quantitatively measure biased competition is by relating the response of a cell to a single stimulus (selectivity) to its response to a stimulus pair (sensory interaction, Reynolds et al. (1999)). The selectivity of a neuron is defined as the response to a variable probe stimulus minus the response of an arbitrarily chosen reference stimulus. A positive selectivity refers to cells preferring the probe whereby a negative value refers to cells preferring the reference stimulus. The sensory interaction is defined as the response to the stimulus pair (probe and reference presented simultaneously) minus the reference alone. A positive value indicates that the cell's response to the reference is increased by adding a probe whereby a negative value denotes suppression by adding the probe. The selectivity and sensory interaction indices for several

randomly chosen probe and reference stimuli have been plotted cell-wise against each other and fitted with a linear regression curve. In area V4, the slope in the unattended condition is 0.49 (Fig. 3c). A slope of 1.0 would indicate no suppression between the stimuli, hence the slope of 0.49 shows for a cell preferring the probe (positive selectivity) a suppression from the reference of about 50%. Likewise for a cell preferring the reference (negative selectivity), it shows a suppression from the probe of about 50%. Attending the probe results in a slope of 0.83 (Fig. 3d) which demonstrates an increase to the pair response for cells preferring the probe (positive selectivity) and a decrease to the pair response for cells preferring the reference (negative selectivity). Contrary, attending the reference (Fig. 3e) will evoke the opposite effect and thus resulting in a slope of 0.21. Additionally to area V4, Reynolds et al. (1999) investigated also area V2 and found a qualitatively similar effect. The model can replicate the data well (Fig. 3c-e) with the previously outlined mechanisms involved in biased competition.

All the above experiments were carried out with stimuli of equal contrast. However, the amount of suppression depends on the contrast of each stimulus. Stimulus contrast is defined in this context as the contrast between the stimulus and the background, hence it describes the strength of a stimulus. Busse et al. (2009) used overlapping gratings of 0° and 90° orientation, varied systematically the contrast of both orientations and recorded the neuronal population via electrode arrays (Fig. 3f). If a single stimulus is presented alone (shown as 0% contrast of the other stimulus in Fig. 3d), the response is characterized by the contrast response function (CRF, Albrecht & Hamilton (1982)). The CRF is implemented in the model by divisive normalization in layer 4:  $r^{L4} = (1 + \frac{1}{\sigma}) \cdot \frac{E}{\sigma + E}$ , whereby the excitation  $E$  represents the stimulus contrast. This divisive normalization explains the strong non-linear influence of the contrast on the response. If a pair of differently contrasted stimuli is presented (12% and 50% contrast in Fig. 3f), the population response is dominated by the higher contrasted stimulus. This implies a non-linear suppression towards cells preferring the lower contrasted stimulus, realized in the model by the same feature-based suppression mechanism as in biased competition: suppression occurs from cells in layer 2/3 preferring the 50% contrasted stimulus towards cells in layer 4 preferring the 12% contrasted stimuli. Afterwards, the decreased response in layer 4 is projected to the recorded neuron in layer 2/3. The observed non-linear effect occurs because the suppression originates in a non-linear manner from the neuronal response in layer 2/3 ( $r^{L2}$ ):  $S = (r^{L2})^{p_{FEAT}}$ ; and because this suppression operates divisively on the firing rate in layer 4:  $r^{L4} = (1 + \frac{1}{\sigma}) \cdot \frac{E}{\sigma + E + b(E) \cdot S}$ . These mechanisms are present in the model in all conditions, but the non-linear effect is mostly visible in cases with different contrasted stimuli. With two equally contrasted stimuli, the data shows another effect, an equally strong suppression towards both stimuli and under all contrasts. This expected result proves that suppression occurs at all contrasts which implies that also biased competition should occur at all contrasts.

Feature-based attention (review: Maunsell & Treue (2006)) can also modulate the suppression between stimuli in a biased competition paradigm, similar to the effect described for spatial attention but with a feature-selective bias. Studies reported the effect for area V4 (Chelazzi et al., 2001; Fallah et al., 2007; Zhou & Desimone, 2011), MT (Treue & Trujillo, 1999), and IT (Chelazzi et al., 1998). In the study of Chelazzi et al. (1998), the task of three rhesus monkeys was to execute a saccade to a previously shown target object that was presented together with distracting objects (Fig. 4a). The

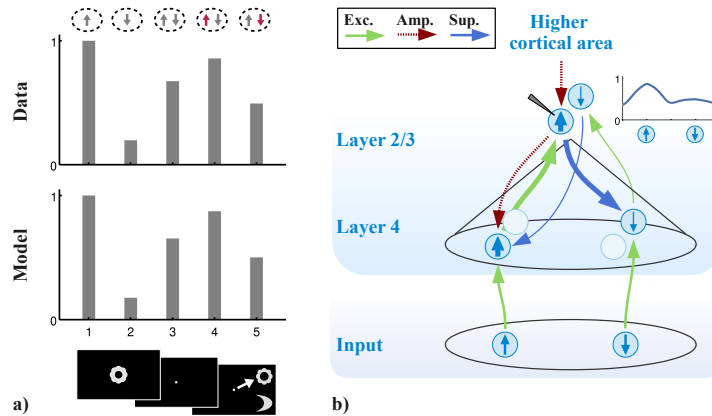


Figure 4: Biased competition experiment with feature-based attention. **a)** Neurophysiological data (top) in relation to our simulation results (bottom, recorded cell is indicated in (b)). The average neuronal response of area V4 for the preferred stimulus (1), anti-preferred stimulus (2) and pair (3-5). In the latter, feature-based attention (denoted by red) is either directed away (3), to the preferred (4) or anti-preferred (5). Data adapted from Chelazzi et al. (1998, Fig. 10a and 10b). **b)** Mechanics of biased competition with feature-based attention. The notation of the figure is similar to Fig. 2. Attentional modulation of neurons and connections are denoted by the increase or decrease of the symbols thickness. The mechanics are based on the experiment illustrated in (a), condition 4.

target object was previously shown at a different location to ensure that the setup causes only feature-based attention. The presented data (Fig. 4a, top) illustrates the population response of neurons preferring the target in area IT. The response was obtained by averaging over 58 neurons and over a time span from 150ms to 300ms after stimulus onset. The time window was chosen so that it starts at the onset of the full response and ends before the neuronal response changes due to onset of the saccade, similar as used by Lee & Maunsell (2010a) for their setup of biased competition. The data demonstrates the typical biased competition effects: the unattended, preferred stimulus alone (Fig. 4a, condition 1) elicits a much stronger neuronal response than the unattended stimulus pair (condition 3). However, if the preferred stimulus of the pair is attended, the response is amplified (condition 4) whereby a decrease of the response is observed if the anti-preferred stimulus is attended (condition 5). Yet, it is still higher as the response to the unattended, anti-preferred stimulus alone (condition 2). Although this task initially deploys feature-based attention as the location of the target object is varied across trials, it has been proposed and modeled that saccade preparation induces a delayed spatial reentrant signal which affects the late response (Hamker, 2005b). However, as we here focus on the cortical microcircuit of attention, we simplify the modeling and only apply a feature-specific bias.

The mechanisms of biased competition with feature-based attention are illustrated in Fig. 4b. The search array composed of target and distractor excites respective cells in layer 4 and layer 2/3. A feature-based amplification originates from a higher cortical area and is projected to layer 2/3 amplifying the neuronal response of the attended feature. This in turn enhances via amplifying connections the neuronal response of this feature in layer 4 and decreases via suppressive connections the response of the

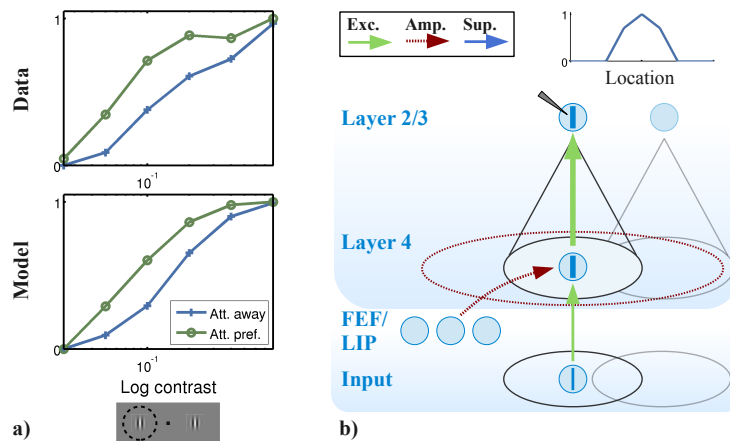


Figure 5: Contrast gain resulting from the attentional modulation of the contrast function. **a)** Neurophysiological data (top) in relation to simulation results (bottom, recorded cell is indicated in (b)). The average neuronal response of area V4 as a function of contrast (data adapted from Reynolds et al. (2000, Fig. 5a)). **b)** Mechanics of contrast gain. Neuronal activity in layer 2/3 occurs only at a small region as illustrated by the spatial response profile at the top right, showing across space the response of neurons preferring the stimulus. The notation of the figure is similar to Fig. 2. Attentional modulation of neurons and connections are denoted by the increase or decrease of the symbols thickness.

distracting feature in layer 4, which also reduces the excitatory signal to layer 2/3 neurons selective for the distracting feature.

### 3.2. Attentional modulation of the contrast response function: contrast gain or response gain

The amount of attentional modulation depends on the stimulus contrast. Contrast gain (definition: Reynolds et al. (2000)) describes that attention amplifies the neuronal response mainly at lower stimulus contrasts whereby at high contrast, the neuronal response saturates to the same value without an explicit threshold on the response. This results in a leftward shift of the contrast response function (CRF, Albrecht & Hamilton (1982)). The response gain denotes the appealing simple idea that attention amplifies the neuronal response by a fixed factor, as already suggested very early by McAdams & Maunsell (1999) or Treue & Trujillo (1999). There has been experimental evidence for both, contrast gain and response gain, as explained in the following.

Reynolds et al. (2000) observed contrast gain in single-cell recordings of area V4 involving spatial attention. Their data (Fig. 5a) illustrates the contrast response function of a population of V4 neurons to a grating. In this study, the task of two rhesus monkeys was to spatially attend inside (attend preferred, Fig. 5a left) or outside the receptive field of the recorded cell (attend away, Fig. 5a right).

Reynolds & Heeger (2009) already suggested by means of model simulations that whether attention leads to contrast gain or response gain depends on the stimulus size in relation to the attended region. Reynolds et al. (2000) used a rectangle as a cue that was much larger than the stimulus, presumably resulting in a broader field of attention compared to the size of the stimulus. Besides area V4, Li et al. (2008) reported contrast

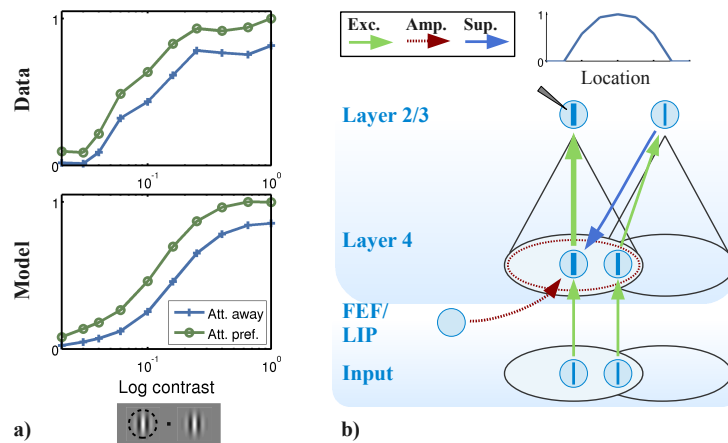


Figure 6: A mix of response and contrast gain resulting from the attentional modulation of the contrast response function. **a)** Neurophysiological data (top) in relation to simulation results (bottom, recorded cell is indicated in b). The average neuronal response of area V4 as a function of contrast (CRF, data adapted from Williford & Maunsell (2006, Fig. 6g)). **b)** Mechanisms responsible for a mix of response and contrast gain. Neuronal activity in layer 2/3 occurs at multiple locations as illustrated by the spatial response profile at the top right, showing across space the response of neurons preferring the stimulus. The notation of the figure is similar to Fig. 2. Attentional modulation of neurons and connections are denoted by the increase or decrease of the symbols thickness.

gain in area V1 with fMRI using spatial attention. Furthermore, Martínez-Trujillo & Treue (2002) reported the effect in area MT for exemplary neurons, but they also found neurons with a response gain.

In our model, contrast gain occurs if the attended region is much larger than the stimulus (Fig. 5b). Thus, spatial attention is deployed very broadly to layer 4, but amplifies only a few neurons as only a few features at a single location will be excited due to the small stimulus size. This response is projected to layer 2/3, activating only neurons within a very narrow region. Thus no suppression from neurons at spatially adjacent locations occurs, allowing the response to saturate. At high contrast, the response saturates to its maximal possible value, so attentional amplification occurs only at low contrasts. Both effects together, the saturation at high and the attentional amplification at low contrasts, induces the typical leftwards shift of the contrast response function (Fig. 5a).

A mixed effect between contrast and response gain was observed by Williford & Maunsell (2006), in area V4. Furthermore, evidence for response gain has been found in fMRI data by Boynton (2009) in areas V1, V2; by Li et al. (2008) in areas V2, V3, V3A, V4s; and by Murray (2008) in areas V1, V2, V3. Williford & Maunsell (2006) recorded the contrast response function (CRF, Albrecht & Hamilton (1982)) for a population of V4 neurons and obtained an average CRF (Fig. 6a) over all neurons with a significant attentional modulation and a significant contrast enhancement of the CRF. The task of two macaque monkeys was to spatially attend a grating either inside the receptive field of the recorded cell (attend preferred, Fig. 6a left stimulus) or outside of it (attend away, Fig. 6a right). The cueing stimulus was a grating of the same size as

the stimuli, presumably resulting in an attended area that has roughly the same size as the stimulus.

In our model, spatial attention leads to a pure response gain if the stimulus is much broader than the attended area (similar as in Reynolds & Heeger (2009)) and in a mix of response and contrast gain if the stimulus and the attended area are similar in size (Fig. 6b). In both cases, the stimulus falls within the receptive fields of multiple, spatially-adjacent neurons in layer 2/3. In the unattended case, the broad responses in layer 2/3 result in a suppression to layer 4 which in turn results in a moderate response in layer 4 (left cell) and also in layer 2/3 (left cell). This occurs similarly at all contrast levels. The suppression occurs between different locations similar as surround suppression, thus we model it via the same signal  $A^{\text{SUR}}$ . In the attended case, attention is deployed to only a small area amplifying only there the neuronal response. Thus, the previously moderate response is amplified. As this mechanism is contrast independent, the response gain is observed under all contrasts.

A response gain is observed in many of the other simulated experiments (biased competition: Fig. 2, 4; attentional modulation of neuronal tuning curve: Fig. 7, 8; attentional modulation of surround suppression: Fig. 9b). In some experiments, it is shown directly by the modulation of the contrast response function (Fig. 9b). In the other experiments, only high contrasted stimuli were presented. In this case, an increase of the neuronal response through attention indicates the response gain effect (Fig. 2, 4, 7, 8). According to our model, a response gain is observed if suppression is similar strong among different contrasts, as observed for example in the study of Lee & Maunsell (2010b) (same data as Lee & Maunsell (2010a) shown in Fig. 2). This setup involves spatial attention and the neuronal responses in the conditions attend preferred and attend anti-preferred were recorded. A response gain is observed because this setup induces similar effects across all contrasts: In the condition attend anti-preferred, a strong suppression from layer 2/3 to layer 4 results in a moderate response (Fig. 2); and in the attend preferred condition, the neuronal responses layer 4 are amplified. As amplification and suppression mechanisms modulate the responses similarly strong for all contrasts, a response gain effect is observed.

### 3.3. Attention modulation of neuronal tuning curves

#### 3.3.1. Scaling of the neuronal tuning curve

Attention also affects the neuronal tuning curve in the feature space, resulting in either a scaling (Fig. 7a) or sharpening (Fig. 8a) of the curve (Ling et al., 2009).

Typically, the deployment of spatial attention results in a multiplicative scaling of the tuning curve, as observed by single-unit recordings in V1 (Motter, 1993), V2 (Motter, 1993) and V4 (David et al., 2008; McAdams & Maunsell, 1999; Motter, 1993). We simulated the study of McAdams & Maunsell (1999) who have recorded tuning curves in area V4 (Fig. 7a). In their study, two rhesus monkeys have to either report the color equality/difference of successive blob-stimuli presented in the left hemifield (condition: attend away) or the equality/difference of successive presented gratings presented in the right hemifield (condition: attend preferred). The stimulus in the right hemifield was placed within the receptive field of the recorded neuron. Furthermore, its orientation was varied to obtain the neuronal tuning curve. The data shows the average population



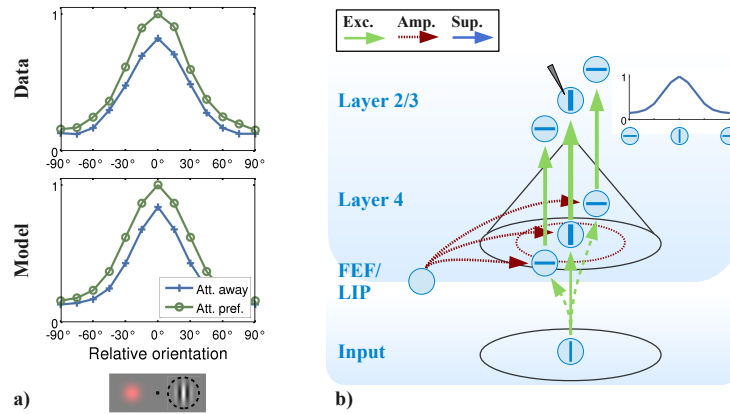


Figure 7: Scaling of the neuronal tuning curve due to spatial attention. **a)** Neurophysiological data (top) in relation to simulation results (bottom, recorded cell is indicated in **b**). Color denotes conditions of attend away (blue) and attend preferred (green). Data adapted from McAdams & Maunsell (1999, Fig. 4). **b)** Mechanics: spatial amplification enhances all features at the attended location with a similar factor, resulting in a multiplicative scaling of the tuning curve. The notation of the figure is similar to Fig. 2. The variable stimulus excites each neuron differently strong according to its feature preference (indicated by dotted arrows). Attentional modulation of neurons and connections are denoted by the increase or decrease of the symbols thickness.

response which was acquired by shifting the preferred tuning of each single neuron to the center of the plot and then afterwards averaging the responses.

According to our model, the scaling effect occurs because spatial attention amplifies the response of all features within a population in layer 4 with a similar factor (Fig. 7b). The increased activity is projected to layer 2/3 neurons. Feature-based suppression is not important in this paradigm as the amount of suppression is relatively similar in both experimental conditions attend away and attend preferred. The amount of suppression depends on the differences between the neuronal responses within layer 2/3, because a cell receives feature-based amplification from its own feature and suppression from opposing ones. As these response differences do not change with spatial attention, the suppression remains constant between both conditions. In addition, suppression from adjacent neurons occurs as the stimulus is large enough to fall within several receptive fields, similar as in the response gain experiment (Fig. 6). In summary, suppression occurs, but the effect is primarily based on the very crucial property of attention, the amplification mechanism.

### 3.3.2. Sharpening of the neuronal tuning curve

Feature-based attention leads to an amplification of the attended feature and to a suppression of unattended ones, known as sharpening (Fig. 8a), as suggested by single-unit recordings in MT (Martinez-Trujillo & Treue, 2004) and V4 (David et al., 2008). Our simulations suggest that the effect is a general property of attention in the cortical microcircuit as it is based on two very basic properties: the amplification mechanism and suppression.

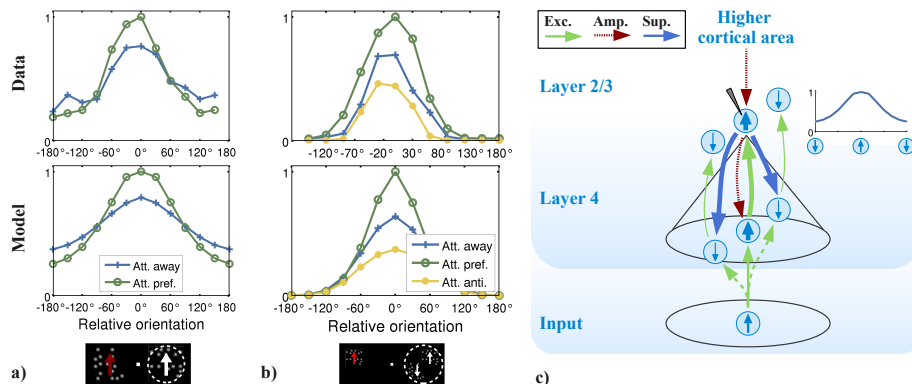


Figure 8: Modulation of neuronal tuning curves due to feature-based attention. **a-b)** Neurophysiological data (top) in relation to simulation results (bottom). **a)** Sharpening of the neuronal tuning curve due to feature-based attention in a single-stimulus setup. The recorded cell is indicated in c). The data illustrates the response of a representative MT neuron for different orientated stimuli (adapted from Martinez-Trujillo & Treue (2004, Fig. 4a). **b)** In biased competition, feature-based attention leads to a scaling of the neuronal tuning curve. Data is adapted from Treue & Trujillo (1999, Fig. 3b). **c)** Mechanisms: in the single-stimulus setup of (a), feature-based attention results in a sharpening as it amplifies the response of only the attended feature and increases suppression to opposing ones. The notation of the figure is similar to Fig. 2. The variable stimulus excites each neuron differently strong according to its feature preference (indicated by dotted arrows). Attentional modulation of neurons and connections are denoted by the increase or decrease of the symbols thickness.

Martinez-Trujillo & Treue (2004) illustrate this effect for a representative neuron in area MT (Fig. 8a). Their experimental setup contains two identically moving random dot patterns (RDPs) at two locations (right location was recorded). In the task, two macaque monkeys have either to spatially attend to the fixation point (condition: attend away; Fig. 8a blue) or to the moving RDP at location 'left' (condition: attend preferred; Fig. 8a green). This setup ensures feature-based attention at the 'right' position: spatially attending the left RDP induces spatial and feature-based attention to it, presumably resulting in a deployment of feature-based attention to all identical RDP in the scene, hence also to the right RDP. The tuning curve is sharpened by attention as the neuronal response to the preferred motion direction is increased and the response to the anti-preferred direction is decreased.

In our model (Fig. 8c), sharpening occurs if feature-based attention is narrowly tuned and thus amplifies only the response of cells in layer 2/3 and layer 4 preferring the attended feature. Amplifying this response leads to a stronger suppression from layer 2/3 to the other (and non-attended) features in layer 4 resulting in a decreased neuronal response. This in turn decreases also the response of opposing features in layer 2/3 (opposing motion direction are at  $180^\circ$  and  $-180^\circ$ ). These suppression effects induce the sharpening of the tuning curve (Fig. 8a).

### 3.3.3. Neuronal tuning curve in biased competition

Treue & Trujillo (1999) investigated the modulation of the tuning curve within an additional distractor by using a biased competition paradigm with feature-based

attention. Unexpectedly, this setup results in a scaling of the tuning curve instead of a sharpening (Fig. 8b). They reported neuronal tuning curves of a representative neuron in area MT and placed two stimuli in the receptive field of a cell: a target random dot pattern (RDP) with variable motion direction to record a tuning curve and a distractor RDP with fixed anti-preferred motion direction of the recorded cell. Attention was deployed either to the target or to the distractor by cueing a third RDP placed in the contralateral hemifield which had the same feature as the attended object. As cue and attended object have different locations, this setup ensures pure feature-based attention. The task of the two macaques was to detect a motion change in the target stimulus. The data shows a multiplicative scaling of the tuning curve when attending the target (up-scaling) or distractor (down-scaling).

A pure multiplicative scaling is surprising as other data shows a sharpening of the tuning curve (Fig. 8a) and other models (Reynolds & Heeger, 2009) could not provide a solid explanation for this discrepancy. However, our modeling of the experiment clarifies that the sharpening effect does not occur due to a strong suppression from the additional distractor. The presence of the distractor suppresses the flanks of the tuning curve as the neuronal responses are zero for anti-preferred targets in the attended-away condition (Fig. 8b, blue curve at motion directions in the range:  $-180^\circ$  to  $-120^\circ$  and  $120^\circ$  to  $180^\circ$ ). Contrary, these responses are non-zero in the data showing the sharpening (Fig. 8a). In the model, suppression is most powerful at the flanks of the tuning curve. It is weaker at the center due to competition between neurons preferring the target and the distractor. Due to this suppression of the flanks, a further sharpening does not occur.

#### 3.4. Attentional modulation of surround suppression

Attention can also modulate the amount of received suppression in center surround experiments. Within a classical receptive field (cRF), attentional modulated suppression is well explained by the biased competition paradigm (Section 3.1). Thus, the surround paradigm is especially interesting as it addresses the question of attentionally modulated suppression beyond the cRF. In such center surround experiments (Gilbert, 1998), a stimulus in the surround of the cRF suppresses the cell’s response as demonstrated by Cavanaugh et al. (2002a) for area V1. They systematically increased the size of a grating beyond the size of the cRF and observed that the average neuronal response starts to decrease after the stimulus size exceeds the cRF showing the suppression effect (Fig. 9a). Furthermore, they found a larger cRF at lower contrasts. The cRF size was defined in their study as the stimulus size invoking the maximum response (grating summation field). Our attention model can account for this effect if we presume a weaker surround suppression at lower contrasts by setting  $p^{\text{SUR}} = 2.0$ . At higher contrasts, the strong suppression removes layer 4 activity at the flanks of the stimulus area, leading to a narrow spatial response curve. Contrary at lower contrasts, the response remains broad in the spatial dimension. Layer 2/3 neurons integrate so over a broader activity at lower contrast, resulting in higher responses. This approach is similar to the “gain” model of Cavanaugh et al. (2002a) which fitted at best the data.

We have simulated the experiment as follow. The input was simulated via a Gaussian in the spatial dimension. Its standard deviation represents the area of the stimulus:  $\sigma \approx \frac{\pi d^2}{4}$ , with stimulus diameter  $d$ . Layer 2/3 neurons need to integrate information

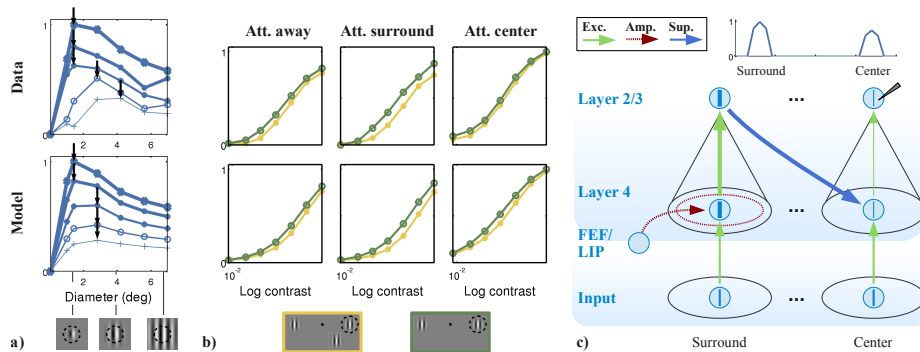


Figure 9: Attentional modulation of surround suppression. **a - b)** Neurophysiological data (top) in relation to simulation results (bottom, recorded cells are indicated in c). **a)** The average neuronal response in area V1 for gratings of increasing diameter size (data adapted from Cavanaugh et al. (2002a, Fig. 8a)). The stimulus contrast was also varied whereby the highest is denoted by the darkest points. The peak in a response curve marks the border of the receptive field, denoted by black arrows. **b)** The average response of neurons in area V4 as a function of contrast if one stimulus is presented in the center and another one in the surround of the classical receptive field (cRF). Yellow denotes this case (condition: surround), green denotes a control setup without the surround stimulus (condition: center). Data was adapted from Sundberg et al. (2009, Fig. 6f). **c)** Mechanisms of attentionally modulated surround suppression. The notation of the figure is similar to Fig. 2. Attentional modulation of neurons and connections are denoted by the increase or decrease of the symbols thickness.

beyond their cRF borders, so layer 4 and layer 2/3 are weakly connected outside the cRF too. In all other experiments, the layers have for simplicity only connections inside a cRF. In both cases, the width of the cRF is the same as the connectivity weights are modeled via the same Gaussian.

Surround suppression is widely observed in the visual cortex: in LGN (Bonin et al., 2005); V1 (Cavanaugh et al., 2002a,b; Bair et al., 2003); V2 (Ito & Gilbert, 1999; Willmore et al., 2010); V4 (Sundberg et al., 2009); MT (Allman et al., 1985; Pack et al., 2005; Hunter & Born, 2011); MST (Orban, 2008); and LIP (Falkner et al., 2010). In IT, no study reported this effect, probably because receptive fields are very large thus it is experimentally very difficult to place a stimulus outside the receptive field. However, Miller et al. (1993) found suppression between pair of stimuli within a receptive field.

The surround effect can also be modulated by attention as attending the center of the receptive field or the surround stimulus amplifies or respectively suppresses the neuronal response. Sundberg et al. (2009) observed this effect by recording neuronal responses in area V4 (Fig. 9b). Their setup (yellow, Fig. 9b) consists of three gratings with variable contrast, one in the opposite hemifield (left stimulus in Fig. 9b), one in the classical receptive field (cRF) of the recorded cell (right stimulus marked with a ring) and one in the remote surround of the cRF (lower right stimulus). Stimulus contrast was varied for the stimulus inside the cRF and was fixed to the maximum for the stimulus in the surround. As a control (green), they replicate the known attentional modulation of the cRF for a single stimulus (Fig. 5 in Section 3.2). In both setups, they investigated the conditions of attending the grating in the opposite hemifield (attend away), in the center of the cRF (attend center, right stimulus), and in the

surround (attend surround). They found that in the attend away condition, the surround stimulus slightly suppresses the neuronal responses (yellow versus green curve). Compared to the data of Cavanaugh et al. (2002a) shown in Fig. 9a, the weak suppression may be explained by the fact that Cavanaugh et al. (2002a) covers the whole surround with a large surround stimulus whereby Sundberg et al. (2009) uses a single and small surround stimulus. Attending the surround slightly increases the amount of surround suppression. In contrast, attending to the center stimulus diminishes the surround suppression (overlapping of the yellow and green curve). In addition, attending to the center stimulus increases the absolute response, indicating a response gain effect in addition to the surround effect.

The mechanics in the model will be explained by the example of the study of Sundberg et al. (2009) (Fig. 9c). The center and the surround stimulus excite neurons in layer 4 and in turn in layer 2/3 at both the surround (left in Fig. 9c) and the center location (right). The layer 2/3 response from the surround is projected back via suppressive connections to layer 4 in the center, decreasing the neuronal response which in turn also decreases the layer 2/3 activations in the center. As this suppression does not occur if no surround stimulus is present, its influence is visible by comparing the responses with and without surround stimulus (yellow versus green in Fig. 9b). An opposite suppression effect occurs also from the center to the surround location (not shown). If the surround is spatially attended, the responses in both layer 4 and 2/3 are amplified which increases the suppression to layer 4 in the center. Attending the center invokes the opposite mechanism, resulting in a decreased response at the surround.

Besides the study of Sundberg et al. (2009), a few other studies investigated the effect of attention in center surround experiments and found similar results: Chen & Martinez-Conde (2008) found the effect in area V1 by spatially attending the center or the surround with a variable task difficulty. On the highest difficulty, they found an amplification of about 140% when attending the center. Ito & Gilbert (1999) shows the effect also for area V1 in the conditions of attending away, the center, or four surround locations simultaneously.

Until now, we have considered only the surround suppression between similar stimuli in cases where the suppression is independent of the spatial arrangement of the surround. This is the most typical case (Zanos et al., 2011) and so implemented as standard in the model. However in general, the effect is more complex as reviewed by Gilbert (1998) or Spratling (2010) (the latter study particularly focused on V1). Bonin et al. (2005) found the condition that suppression occurs from all features in the surround, independently of their similarity. This was also observed for random dot patterns which create a strong suppression independent from spatial arrangement. However in the paradigm denoted as contour linkage (or flanker effect, Gilbert (1998)), adjacent stimuli can even enhance each other if they form together a large contour. The model might also account for these effects if the amplifying and suppressive connections from layer 2/3 to layer 4 ( $A^{\text{FEAT}}$  and  $S^{\text{SUR}}$ ) are adjusted appropriately.

### 3.5. Predictions

The model predicts interactions between many of the previously outlined paradigms as the underlying mechanisms will normally interact with each other. We will exemplarily investigate the interactions of biased competition and surround suppression.

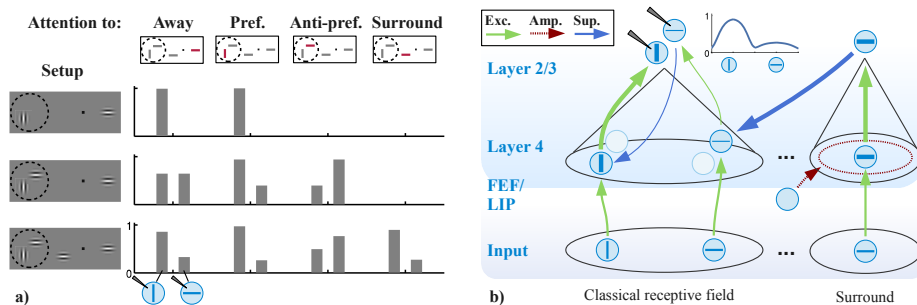


Figure 10: Prediction of biased competition with surround suppression. **a)** The setup consists of a preferred (vertical grating) and anti-preferred stimulus (horizontal grating) inside the classical receptive field (cRF), plus an anti-preferred stimulus in the surround. Three setup variations were simulated: the preferred stimulus alone (first row), biased competition alone containing the preferred and anti-preferred stimulus in the cRF (second row), and biased competition with the additional anti-preferred stimulus in the surround (third row). Model responses are illustrated in the condition of attending away or of attending one of the three stimuli. The attended stimulus is marked in red. Responses are not displayed if the attended stimulus is absent in a setup. **b)** Model mechanisms, illustrated in the condition of spatially attending the surround stimulus (last condition in a). The notation of the figure is similar to Fig. 2. Attentional modulation of neurons and connections are denoted by the increase or decrease of the symbols thickness.

The competition between a stimulus pair can be favored for one of them by attending it (biased competition). However, a surround stimulus can also favor one of them by selectively suppressing the other one. An exemplary setup is depicted in Fig. 10a, third row. It consists of a preferred (vertical grating) plus an anti-preferred stimulus (horizontal grating) inside the classical receptive field (cRF), and an additional anti-preferred stimulus in the surround (horizontal grating). Stimulus contrast is high and equal for all stimuli. We propose to record from a neuron in area V4 as biased competition and surround suppression effects were reported for this area (Reynolds et al., 1999; Sundberg et al., 2009). As surround suppression strength depends on the stimulus size (Cavanaugh et al., 2002a) and as the suppression is weak in V4 for small stimuli (Sundberg et al., 2009), we use a large surround stimulus to obtain a strong effect (3x of the cRF size). Compared to this, the stimuli inside the cRF are much smaller (1/5 of the cRF size). The interaction between surround and stimuli inside the cRF can be very complex (Section 3.4). We assume here the most typical case in which only similar stimuli suppress each other.

The model predicts that an anti-preferred surround stimulus decreases the response to an anti-preferred stimulus in the cRF, leading to an increased response of a preferred stimulus in the cRF. We simulated three setups (Fig. 10a): the preferred stimulus alone (first row), biased competition (second row), and biased competition with a surround stimulus (third row). For all setups, we simulated the attentional conditions of 1) attend away, 2) attend preferred, 3) attend anti-preferred, and 4) attend surround stimulus. The model uses in all simulations its standard parameters (Tab. 1 in Section 2). The first setup contains the preferred stimulus only and so illustrates the response without suppression from other stimuli. Unattended, the response is already strong due to the high stimulus contrast, so attending the stimulus increases the response only marginally. The second setup shows the classical biased competition effect, namely the average re-

sponse to the pair of stimuli if attention is directed away. Furthermore, if one of the two stimuli is attended, its response is amplified and the other one suppressed. The third setup combines biased competition with surround suppression. In the attended away condition, the response of neurons encoding the preferred stimulus is halfway between the response to the stimulus alone and to the pair in the biased competition setup. Complementary, the response of neurons encoding the anti-preferred stimulus is lower. Both responses indicate that the surround stimulus suppresses the anti-preferred stimulus in the cRF. Attending the surround stimulus increases this effect. Attending one of the two stimuli inside the cRF results in the previously shown biased competition effects with a minor influence of surround suppression, illustrating that the amplification has a stronger influence than the surround suppression.

Attending the surround stimulus has the most prominent effect (Fig. 10b). In the model, it amplifies the response of neurons encoding this stimulus, which strongly suppress neurons encoding the anti-preferred stimulus inside the cRF. Contrary, neurons encoding the preferred stimulus are not suppressed as surround suppression occurs only between neurons encoding similar features, here the features of the anti-preferred stimuli. In the cRF, the reduced response of those neurons diminishes the feature-based suppression to the neurons encoding the preferred stimulus. The response of the neurons selective for the preferred stimulus is so almost similar as in the condition of presenting the preferred stimulus alone. If the surround stimulus is not attended (attend away condition), the mechanisms operate identical but are weaker. If a stimulus inside the cRF is attended, the mechanisms of biased competition have more influence than surround suppression. Attending the preferred stimulus removes the influence of the weak suppression from the distractor in the cRF, so the response increases slightly. Attending the anti-preferred stimulus amplifies the response of neurons encoding it, which diminishes the influence of suppression. This suppression results from the neurons in the surround as well as from the neurons encoding the preferred stimulus (feature-based suppression). Thus, the accumulated suppression has still a notable influence and the response of those neurons is less increased as in biased competition. This lower response results in less feature-based suppression back to the neurons encoding the preferred stimulus, so their rate is less decreased as in biased competition.

#### **4. Discussion**

Experimental investigations that focused on the attentional modulation of the neuronal response in visual areas have led to a large set of data. In order to understand the underlying mechanisms of attention, computational models can help if they allow to explain the observations by only a few core mechanisms. In the proposed model, these are amplification, normalization, spatial pooling and suppression. In the following, we will review the abilities of existing models and the mechanisms they incorporate. Besides, we will discuss limitations of the proposed model. Furthermore, we will discuss the physiological foundation of the approach, especially concerning the implementation of the suppression.

Fig. no.	Experiment	Reynolds & Heeger (2009)	Boynton (2009)	Lee & Maunsell (2009)	Ni et al., (2012)	Spratling (2008)	Wagatsuma et al., (2013)
2, 3a	Biased competition with spatial att.			+	+	+	
3a, b	Feature pooling			+	+		
3c-e	Selectivity vs. sensory interaction						
3f	Contrast dependency of suppression						
4	Biased comp. with feature-based att.	+				+	
5	Contrast gain	+	+				
6	Response gain	+	+				
7	Scaling	+	+			+	+
8a	Sharpening	+	+			+	+
8b	Tuning curve in biased competition	imprecise fit					
9a, b	Surround suppression						

Table 2: Capabilities of existing single-area models to account for a particular attention experiment. The symbol ‘+’ denotes cases where the publication shows simulation results for a certain experiment.

#### 4.1. Relation to existing single-area models

The motivation for a unified model of visual attention arises mainly from the facts that each of the more recent existing single-area models can account for only a subset of the data (Tab. 2) and that most models do not explain the involvement of underlying mechanism within the cortical microcircuit.

The model of Reynolds & Heeger (2009) is probably the most popular one as it can replicate data from several experiments: modulation of the contrast response function, modulation of the neuronal tuning curve and biased competition with feature-based attention. Essential ingredients for its abilities are divisive normalization (Carandini & Heeger, 2012) accounting for all contrast related properties, and a generic attention plus a suppression field to simulate amplification and suppression respectively. Reynolds & Heeger (2009) state that they made no assumption how and over which connectivity these interactions are carried out in the cortical microcircuit. This improves the flexibility of the model, however it is a drawback in terms of specific predictions about the putative implementation within the microcircuit. As compared to our model, their model has no clear definition of a receptive field as it lacks the spatial pooling operation in layer 2/3. So it cannot replicate experiments where spatial attention is directed to only a part of the receptive field, e.g. biased competition with spatial attention and selectivity versus sensory interaction (Tab. 2, rows 1 and 3).

To recapitulate, the biased competition experiments with spatial attention (Reynolds et al., 1999) use two stimuli which are placed within a receptive field at different positions. In our model, a spatial pooling operation is realized by excitatory connections from layer 4 to layer 2/3. By these converging connections, the neuronal responses of



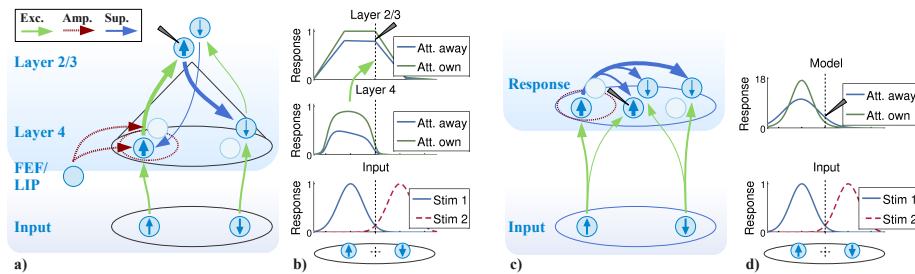


Figure 11: Comparison of our model with spatial pooling (a, b) with the model of Reynolds & Heeger (2009) lacking pooling (c, d), in a spatial biased competition experiment. **a)** The proposed mechanistic cortical microcircuit of attention. Notation is similar to Fig. 2. Attentional modulation of neurons and connections are denoted by the increase or decrease of the symbols thickness. **b)** For our model with spatial pooling, the responses of the input layer, layer 4 and layer 2/3 are illustrated over all location within the receptive field. The responses of the upward moving stimulus are denoted by blue (unattended) and green colors (attended own) whereby the downward one is shown in red, but for simplicity only in the input. The model assumes symmetrical receptive fields, thus the recorded cell is located in the center, denoted by a dotted line. **c)** Model without spatial pooling proposed by Reynolds & Heeger (2009). Notation is identical to a). To facilitate a comparison with a), the mathematical formulation of the model is illustrated as a neuronal circuit. **d)** For the model without spatial pooling, responses of the input and the recorded layer. Notation is identical to b).

a neuron in layer 4 stimulate the recorded neuron (Fig. 11b, green arrow) which defines the receptive field by its bottom up connectivity. Thus, any amplification caused by spatial attention to a part of the receptive field will amplify also the recorded neuron (blue versus green curve in layer 2/3). Using this mechanism, the proposed model replicates the physiological data (Section 3.1), i.e. that the attended response is higher as the unattended one.

Contrary to our model, Reynolds & Heeger (2009) (Fig. 11c, d) use a spatial broadening of the stimulus instead of a pooling operation. Indeed, in the unattended condition, this approach is effective as the neuron at the center location is activated (Fig. 11d, blue). Attending the stimulus location amplifies also the neuronal response at the location of the stimulus (Fig. 11d, green). However, this amplified response induces a stronger suppression at the other locations, thus the response will decrease at the recorded neuron located at the center. This decrease is obviously in contrast to the observed amplification in biased competition (Lee & Maunsell, 2010a) as shown in Fig. 2 (Section 3.1). However, if the model of Reynolds & Heeger (2009) would be extended by a spatial pooling layer, it could theoretically simulate this biased competition paradigm. In summary, any attention experiment which requires to spatially attend only a part of the receptive field and to record at another location, cannot be simulated by a single layer neuronal model.

Others (Boynton, 2009; Lee & Maunsell, 2009, 2010a,b) published simpler computational models based on divisive normalization. The model of Boynton (2009) is primary intended to simulate the modulation of the contrast response function. Thus, their model has a contrast and feature dimension, but no spatial dimension and no attention or suppression field. Attention is modeled in a very abstract parametric fashion and not by an intrinsic mechanism, rather it is applied via three independent parameters: the multiplicative scaling of a tuning curve, a baseline gain and a shift of the

tuning curve to the left. These parameters allow to fit the model to data with respect to the attentional modulation of the neuronal tuning curve and contrast response function (Tab. 2), but make it unlikely that the model can account for other data. In particular it lacks any suppression mechanism.

Another simple model based on divisive normalization is from Lee & Maunsell (2009) (also used and described in Lee & Maunsell (2010a,b)). Their model simulates a response based on a driving input and a normalization which may involve one or two stimuli. The normalization term in the denominator of the neuronal response function operates as suppression. It is simulated via an exponential function based on the contrast and a non-linearity, a somewhat different type of normalization compared to Carandini & Heeger (2012). A stimulus is only encoded by a single rate variable instead of a neuronal population, so the model lacks tuning curve characteristics. To overcome this limitation, they model only the interactions between two stimuli. More precisely, they firstly calculated the tuning curve for a single stimulus from the physiological data without using their model. Secondly, they simulate the interactions by passing the previously calculated tuning curves of the two stimuli to their model which then modulates and normalizes the responses (Lee & Maunsell, 2010a). Due to this procedure, their model can explain the change in the neuronal response when a second stimulus is added in the receptive field, but not the response towards the single stimulus itself. In summary, they simulate biased competition and feature pooling, both with spatial attention (Tab. 2, rows 1 and 2). The research group of Maunsell uses in Ni et al. (2012) a new model to explain their findings that neurons are modulated differently strong by attention. Their approach is called tuned normalization and it is more in line with the approach of divisive normalization. Otherwise, this model incorporates the same ideas as the previous one (Lee & Maunsell, 2009, 2010a,b). In summary, their model contains all necessary mechanisms (amplification, normalization and suppression) except spatial pooling to simulate all considered attention experiments (Tab. 2), so it could theoretically simulate more experiments as the published biased competition experiments.

Spratling (2008) proposed an attention model which has a higher biological plausibility as the previously reviewed models because it explicitly includes neural connections. The model contains two stages, possible representing two different brain areas (Spratling, 2012), whereby each stage contains two sublayers: a prediction and an error layer. The prediction layer sends an expectation signal to the error layer, so the error layer can minimize the reconstruction error. Attention signals are deployed to the prediction layer, which are then transported to the error layer. Spatial attention is applied to stage 1 and feature-based attention to stage 2. Thus their and our model have a similar structure because both use two layers to simulate a single cortical area, but their model is fundamentally different in all other aspects. Spratling (2008) can replicate a few attention paradigms with his model: modulation of the contrast response function, modulation of the neuronal tuning curve, and the decrease of reaction times (Tab. 2). A different version of his model (Spratling, 2010) can replicate additionally the data of MacEvoy et al. (2009) (competition and feature pooling without attention) and of Cavanaugh et al. (2002a) (surround suppression without attention). However, this model variant does not contain attention signals and so we do not include it into our list of attention models (Tab. 2).

An attention model which has the advantage to precisely simulate the cortical microcircuit is Wagatsuma et al. (2013). Their model is a revised version of an existing model of the V1 microcircuit (Potjans & Diesmann, 2012) consisting of layers 2/3, 4, 5, and 6. Their model uses spiking neurons to allow a precise modeling of the temporal effects of attention. Thus the model is focused on replicating such effects like oscillations, but it can also replicate two experiments considered here (Tab. 2): the scaling and the sharpening of neuronal tuning curves. However, due to the high complexity of the neuronal circuit, it cannot easily explain which mechanisms are necessary for a particular attention effects. Moreover, the range of simulated experiments is presently too small to draw general conclusions about the role of connections in the simulated microcircuit. There exist also a few other models based on spiking neurons (Ardid et al., 2007; Buia & Tiesinga, 2008; Hugues & José, 2010), but with much less realistic modeling of the microcircuit and no further advantages.

#### *4.2. Model limitations*

The proposed model of visual attention focuses on the explanation of single unit recordings by means of a mechanistic cortical microcircuit of attention. Thus, the model covers the information processing necessary for replicating the neural firing rates emerging in these experiments, but it simplifies all other aspects as much as possible to keep the model simple and focused. For example, the model contains only layer 4 and layer 2/3 of the cortical microcircuit (Douglas & Martin, 2004) and only the necessary part of its connectivity.

Furthermore, to allow an easy visualization of neuronal responses, we model only a single feature dimension, and a single spatial dimension instead of two spatial retinotopic dimensions. We also do not simulate any feature integration or learning in the visual cortex, hence all layers use the same features.

As the model focuses on the simulation of a cortical microcircuit of attention, experiments involving multiple areas cannot be simulated. In previous studies, we focused on the interaction of the visual cortex with the frontal eye field (Hamker, 2005b) to address the temporal dynamics of attention and the interplay between feature-based and spatial attention. Moreover, splits of spatial attention have been addressed (Zirnsak et al., 2011). Furthermore, attention effects operating simultaneously on multiple levels cannot be simulated as this case might also involve multiple areas. One exemplary experiment is object based attention in the form where a subpart of an object is attended which induces attention to the whole object and to other subparts (Mitchell et al., 2004; Reynolds & Heeger, 2009).

Furthermore, temporal effects of attention are beyond the grasp of this model. For example, attention modifies the pattern of neural oscillations (Jensen et al., 2007; Grossberg & Versace, 2008), decreases the transmission delays (Sundberg et al., 2012), decreases the variability of the neuronal response (Mitchell et al., 2007) or enhances the temporal coupling of neurons (Gregoriou et al., 2009). These findings may be crucial to understand the full neural implementation of the proposed cortical microcircuit of attention and would require a different level of implementation detail.

### 4.3. *Physiological origin of suppression*

At present, it is not accurately known (Reynolds & Heeger, 2009) how the basic mechanisms of attentional processing, i.e. amplification, normalization and suppression, are implemented in the cortical microcircuit of a visual area (Douglas & Martin, 2004). As the mechanics of amplification and normalization in our model are inspired by (Reynolds & Heeger, 2009) and (Carandini & Heeger, 2012), we would like to refer to this literature about the physiological foundations of the first two mechanisms. Very recent research shows also additional possibilities, for example Brosch & Neumann (2014) have proposed an implementation of amplification via interneurons. Concerning the suppression mechanism, our model proposes a new and more concrete implementation, so we will substantiate this point by relating our model to the suppression literature.

In our model, suppression is implemented for simplicity only via connections from layer 2/3 to layer 4. In the cortical microcircuit (Potjans & Diesmann, 2012), such connectivity could be realized by a chain of anatomical connections: excitatory connections from layer 2/3 to layer 5/6, excitatory connections from layer 5/6 to layer 4, and inhibitory connections within layer 4. However, we are aware that there exist also other possibilities to implement the suppression in the cortical microcircuit (Potjans & Diesmann, 2012). We have chosen layer 2/3 as starting point for the suppression as layer 2/3 is the entry point for feedback connections from higher areas (Douglas & Martin, 2004). Feedback seems to transport suppressive signals from higher to lower areas.

In principle, suppressive signals can be transmitted via feedforward, lateral or feedback connectivity (Gilbert, 1998). Feedback connections seem to be involved in surround experiments. The reviews of Carandini & Heeger (2012) and Gilbert (1998) rule out feedforward connections, Bair et al. (2003) concludes that suppression of neuronal response originates from feedback signals, and Angelucci et al. (2002) found a correspondence between the spatial extent of feedback connections and of surround suppression. Lateral connections seem mostly to be involved in low contrast experiments. Angelucci et al. (2002) found that the receptive field size of low contrast stimuli corresponded to the extent of lateral connections. Hunt et al. (2011) found that excitatory lateral connections typically resemble large structures available in natural scenes, like edges or circles, denoted as co-linearity or co-circularity. A study by Cass & Spehar (2005) links these two facts together: they found that the range of collinear contrast facilitation is consistent with long-range horizontal transmission. In conclusion, suppression mediated by feedback from higher areas can be easily incorporated into our model.

### 4.4. *Relevance of the model for psychophysical experiments*

The proposed attention model of the cortical microcircuit can explain the data of several physiological attention experiments, but the model mechanisms are also very relevant for psychophysical experiments.

One psychophysical aspect of attention is its selective processing, because it selects task relevant stimuli under many distractors (Carrasco, 2011). The selection might be necessary to not overload the visual systems limited capacity (Todd et al., 2005), or

to save energy as the brain’s metabolic cost depends mostly on the neuronal activity (Carrasco, 2011). In the model, such a selection can be achieved by amplification of the target and by suppression of distractors. Visual search (Lee & McPeck, 2013) is a typical experiment involving such a selection. Suppression is also observed in crowding (Whitney & Levi, 2011) where the visibility of a stimulus is reduced if it is shown with nearby stimuli. One influential factor for the reduced visibility may be our proposed top-down suppression mechanism.

Besides the selection of stimuli, attention increases the behavioral acuity of subjects by increasing contrast sensitivity, by decreasing reactions times, or by reducing external noise (Carrasco, 2011). The increased contrast sensitivity is also visible in the simulated physiological data as attention causes an appropriate modulation of the contrast response function (Section 3.2). Noise reduction can be realized by amplifying the neuronal response of the target, and by suppressing all other neuronal responses. Therefore, many psychophysical aspects of attention will rely on the mechanisms proposed in our model.

#### 4.5. Influence of model parameters

Our model contains several free parameters that have been tuned to fit different data sets. However, the model should not be dominated by the particular setting of its parameters. Hence, the model should be able to qualitatively fit the data with a fixed parameter set, and parameter variations should only be necessary for a quantitative fit. The model includes parameters (Tab. 1) to modify the contrast response function (parameters  $v^E$ ,  $p^E$ ), to scale the influence of amplification ( $v^{\text{FEF-L4}}$ ,  $v^{\text{PFC-L2}}$ ) or suppression ( $v^{\text{FEAT}}$ ,  $p^{\text{FEAT}}$ ,  $v^{\text{SUR}}$ ,  $p^{\text{SUR}}$ ), to add an un-modulated baseline ( $\beta$ ), and to change the neuronal tuning in the feature space (tuning curve). We evaluated if the model can qualitatively fit the data by fixing all parameters to a reasonable value, denoted as standard parameter set (Tab. 1, row ‘std’). Only the neuronal tuning curve was allowed to be freely chosen because of the heterogeneous response characteristics of neurons in different cortical areas.

The results, illustrated in the supplementary materials, show a satisfying qualitative fit for nine of the twelve experiments. In two experiments (Sundberg et al. (2009), Fig. 9b; Williford & Maunsell (2006), Fig. 6), our results reproduce all significant effects of the original publication, but differ notable from the data. In one experiment (Cavanaugh et al. (2002a), Fig. 9a), a significant effect cannot be reproduced. The deviations result in all three data sets from an inappropriate strength of the surround suppression.

We found that the surround suppression strength ( $v^{\text{SUR}}$ ) and non-linearity ( $p^{\text{SUR}}$ ) vary strongly among data sets (Tab. 1). In the V4 data of Sundberg et al. (2009) (Fig. 9b), the surround suppression is weak which we quantitatively fit by setting  $v^{\text{SUR}} = 0.55$ . Contrary, the V1 data of Cavanaugh et al. (2002a) (Fig. 9a) was best fitted with a strong surround suppression of  $v^{\text{SUR}} = 0.8$ . Besides these two data sets, the surround suppression mechanism is also used for suppression from close, adjacent neurons in four data sets (Fig. 3b, 6, 7, 8a). These were fitted optimally by  $v^{\text{SUR}} = 0.5, 1.0, 0.75, 0.5$ . Thus the standard value for  $v^{\text{SUR}}$  is chosen as a compromise of all these values:  $v^{\text{SUR}} = 0.5$ .

The strength of the non-linearity  $p^{\text{SUR}}$  varies as well. In the data sets of Sundberg et al. (2009) and Cavanaugh et al. (2002a),  $p^{\text{SUR}} = 2.5$  and  $p^{\text{SUR}} = 2.0$  results in the best

	1) Amplification	2) Normalization	3) Spatial pooling	4a) Feature-based suppression	4b) Surround suppression
Biased competition with spatial att.	+		+	+	
Biased comp. with feature-based att.	+			+	
Contrast gain	+	+			
Response gain	+	+			+
Scaling	+				
Sharpening	+			+	
Surround suppression	+				+

Table 3: Overview about the primarily involved mechanisms in each experiment. The related explanation of a mechanism in a particular experiment can be found in Section 3.

fit. However, in experiments with suppression from close, adjacent neurons, disabling the non-linearity by setting  $p^{\text{SUR}} = 1$  results in the best fit. The diversity might depend on factors such as if the suppression originated from neurons in the far (Fig. 9a, 9b) or close surround (Fig. 3b, 6, 7, 8a). We have chosen as standard value  $p^{\text{SUR}} = 1$  as this value allows to fit more data sets. Unfortunately, this standard value prevents that the model can reproduce the significant increase of receptive field at lower contrasts in the data of Cavanaugh et al. (2002a), as discussed in the supplementary materials. Yet, the model still shows suppression from the surround at all contrasts. As a side effect, high values for  $p^{\text{SUR}}$  decrease the influence of the suppression as well.

Therefore, the fits of Sundberg et al. (2009) and of Williford & Maunsell (2006) differ notable as the suppression has an incorrect strength. The standard values of  $v^{\text{SUR}} = 0.5$ ,  $p^{\text{SUR}} = 1.0$  results in a suppression which is too strong for the experiment of Sundberg et al. (2009), which is optimally fitted by  $v^{\text{SUR}} = 0.55$ ,  $p^{\text{SUR}} = 2.5$ , and is too weak for the experiment of Williford & Maunsell (2006), which is optimally fitted by  $v^{\text{SUR}} = 1.0$ ,  $p^{\text{SUR}} = 1.0$ .

In summary, the standard model can replicate the main effects of all data sets with minor deviations regarding surround suppression effects. Thus, all parameters could be fixed for a qualitative fit except the surround suppression parameters  $v^{\text{SUR}}$ ,  $p^{\text{SUR}}$ . Therefore, we conclude that primarily the model mechanisms, and not the free parameters are responsible to account for the data sets.

## 5. Conclusion

In this work, we proposed a new mechanistic model of the cortical microcircuit to explain the neuronal response modulation caused by visual attention. For this purpose, we unify existing proposals of attention into a new neurocomputational model and underline psychophysical and physiological concepts of attention with constraints from neuroanatomy. We found that many visual attention effects can be explained by a few mechanisms: 1) amplification of the response of a single neuron, 2) divisive normalization of this response, 3) spatial pooling within the receptive field, and 4) suppression between neurons encoding different features or different locations. By these effects

(Tab. 3), we are able to explain the following broad variance of physiological single-unit observations by a single model: biased competition with spatial or feature-based attention; modulation of the contrast response function resulting in contrast or response gain; modulation of the neuronal tuning curve resulting in scaling or sharpening; and modulation of surround suppression. Moreover, our model can predict attention effects and mechanisms in novel paradigms, which we demonstrated by combining the paradigms of biased competition and surround suppression.

The current model could be used as a core element in larger models of visual attention that include multiple areas in the visual cortex and control structures like FEF/LIP and PFC/7a. Such a system-level model could simulate very well psychophysical experiments as already demonstrated by a previous system-level model of attention of our group (Hamker, 2005b,a) for receptive field dynamics (Hamker & Zirnsak, 2006), spatial compression (Zirnsak & Hamker, 2010), or the split of spatial attention (Zirnsak et al., 2011). Besides psychophysics, our model may also inspire computational vision models for robots, e.g. Antonelli et al. (2014).

**Acknowledgements:** This work has been supported by a grant from the German Research Foundation (grant no. DFG HA2630/6-1) within the Research Network “Neuro-Cognitive Mechanisms of Conscious and Unconscious Visual Perception” (grant no. PAK 270/1 and 2), in part by the European Project FP7-NBIS “Spatial Cognition” (grant no. 600785), and in part by the Research Training Group “Cross-worlds” (grant no. GRK1780) founded by the German Research Foundation.

## References

- Albrecht, D., & Hamilton, D. (1982). Striate cortex of monkey and cat: Contrast response function. *J Neurophysiol*, *48*.
- Allman, J., Miezin, F., & McGuinness, E. (1985). Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception*, *14*, 105–26.
- Angelucci, A., Levitt, J. B., Walton, E. J. S., Hupe, J.-M., Bullier, J., & Lund, J. S. (2002). Circuits for local and global signal integration in primary visual cortex. *J Neurosci*, *22*, 8633–46.
- Antonelli, M., Gibaldi, A., Beuth, F., Duran, A. J., Canessa, A., Chessa, M., Hamker, F., Chinellato, E., & Sabatini, S. P. (2014). A hierarchical system for a distributed representation of the peripersonal space of a humanoid robot. *IEEE Trans Auton Mental Develop*, *6*, 259–273.
- Ardid, S., Wang, X.-J., & Compte, A. (2007). An integrated microcircuit model of attentional processing in the neocortex. *J Neurosci*, *27*, 8486–95.
- Bair, W., Cavanaugh, J. R., & Movshon, J. A. (2003). Time course and time-distance relationships for surround suppression in macaque V1 neurons. *J Neurosci*, *23*, 7690–701.

- Barone, P., Batardiere, A., Knoblauch, K., & Kennedy, H. (2000). Laminar distribution of neurons in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and indicates the operation of a distance rule. *J Neurosci*, *20*, 3263–81.
- Bonin, V., Mante, V., & Carandini, M. (2005). The suppressive field of neurons in lateral geniculate nucleus. *J Neurosci*, *25*, 10844–56.
- Boynton, G. M. (2009). A framework for describing the effects of attention on visual responses. *Vision Res*, *49*, 1129–43.
- Brosch, T., & Neumann, H. (2014). Interaction of feedforward and feedback streams in visual cortex in a firing-rate model of columnar computations. *Neural Net*, *54*, 11–6.
- Buia, C. I., & Tiesinga, P. H. (2008). Role of interneuron diversity in the cortical microcircuit for attention. *J Neurophysiol*, *99*, 2158–82.
- Busse, L., Wade, A. R., & Carandini, M. (2009). Representation of concurrent stimuli by population activity in visual cortex. *Neuron*, *64*, 931–42.
- Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Rev Neurosci*, *13*, 51–62.
- Carrasco, M. (2011). Visual attention: the past 25 years. *Vision Res*, *51*, 1484–525.
- Cass, J. R., & Spehar, B. (2005). Dynamics of collinear contrast facilitation are consistent with long-range horizontal striate transmission. *Vision Res*, *45*, 2728–39.
- Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002a). Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *J Neurophysiol*, *88*, 2530–46.
- Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002b). Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *J Neurophysiol*, *88*, 2547–56.
- Chelazzi, L., Duncan, J., Miller, E. K., & Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *J Neurophysiol*, *80*, 2918–40.
- Chelazzi, L., Miller, E. K., Duncan, J., & Desimone, R. (2001). Responses of neurons in macaque area V4 during memory-guided visual search. *Cerebral cortex*, *11*, 761–72.
- Chen, Y., & Martinez-Conde, S. (2008). Task difficulty modulates the activity of specific neuronal populations in primary visual cortex. *Nat Neurosci*, *11*, 974–982.
- Compte, A., & Wang, X. (2006). Tuning curve shift by attention modulation in cortical neurons: a computational study of its mechanisms. *Cerebral Cortex*, *16*, 761–78.



- David, S. V., Hayden, B. Y., Mazer, J. a., & Gallant, J. L. (2008). Attention to stimulus features shifts spectral tuning of V4 neurons during natural vision. *Neuron*, *59*, 509–21.
- Desimone, R. (1998). Visual attention mediated by biased competition in extrastriate visual cortex. *Philosophical transactions of the Royal Society of London. - Series B, Biological sciences*, *353*, 1245–55.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective attention. *Annu Rev Neurosci*, *18*, 193–222.
- Douglas, R. J., & Martin, K. a. C. (2004). Neuronal circuits of the neocortex. *Annu Rev Neurosci*, *27*, 419–51.
- Falkner, A. L., Krishna, B. S., & Goldberg, M. E. (2010). Surround suppression sharpens the priority map in the lateral intraparietal area. *J Neurosci*, *30*, 12787–97.
- Fallah, M., Stoner, G. R., & Reynolds, J. H. (2007). Stimulus-specific competitive selection in macaque extrastriate visual area V4. *Proc Natl Acad Sci USA*, *104*, 4165–9.
- Gilbert, C. D. (1998). Adult cortical dynamics. *Physiol Rev*, *78*, 467–85.
- Gregoriou, G., Gotts, S., Zhou, H., & Desimone, R. (2009). High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science*, *324*, 1207–1210.
- Grossberg, S., & Versace, M. (2008). Spikes, synchrony, and attentive learning by laminar thalamocortical circuits. *Brain Res*, *1218*, 278–312.
- Hamker, F. (2004). Predictions of a model of spatial attention using sum-and max-pooling functions. *Neurocomputing*, *56*, 329–343.
- Hamker, F. (2005a). The emergence of attention by population-based inference and its role in distributed processing and cognitive control of vision. *J Comput Vis Image Underst*, *100*, 64–106.
- Hamker, F. (2005b). The reentry hypothesis: the putative interaction of the frontal eye field, ventrolateral prefrontal cortex, and areas V4, IT for attention and eye movement. *Cerebral cortex*, *15*, 431–47.
- Hamker, F. H., & Zirnsak, M. (2006). V4 receptive field dynamics as predicted by a systems-level model of visual attention using feedback from the frontal eye field. *Neural Net*, *19*, 1371–82.
- Heuer, H. W., & Britten, K. H. (2002). Contrast dependence of response normalization in area MT of the rhesus macaque. *J Neurophysiol*, *88*, 3398–408.
- Hugues, E., & José, J. V. (2010). A unified and quantitative network model for spatial attention in area V4. *J Physiol*, *104*, 84–90.

- Hunt, J. J., Bosking, W. H., & Goodhill, G. J. (2011). Statistical structure of lateral connections in the primary visual cortex. *Neural Syst Circuits, 1*, 3.
- Hunter, J. N., & Born, R. T. (2011). Stimulus-dependent modulation of suppressive influences in MT. *J Neurosci, 31*, 678–86.
- Ito, M., & Gilbert, C. (1999). Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron, 22*, 593–604.
- Jensen, O., Kaiser, J., & Lachaux, J.-P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends Neurosci, 30*, 317–24.
- Jones, H. E., Wang, W., & Sillito, a. M. (2002). Spatial organization and magnitude of orientation contrast interactions in primate V1. *J Neurophysiol, 88*, 2796–808.
- Kastner, S., & Ungerleider, S. (2000). Mechanisms of visual attention in the human cortex. *Annu Rev Neurosci, 23*, 315–341.
- Lee, B.-T., & McPeck, R. M. (2013). The effects of distractors and spatial precues on covert visual search in macaque. *Vision Res, 76*, 43–9.
- Lee, J., & Maunsell, J. H. R. (2009). A normalization model of attentional modulation of single unit responses. *PLoS One, 4*, e4651.
- Lee, J., & Maunsell, J. H. R. (2010a). Attentional modulation of MT neurons with single or multiple stimuli in their receptive fields. *J Neurosci, 30*, 3058–66.
- Lee, J., & Maunsell, J. H. R. (2010b). The effect of attention on neuronal responses to high and low contrast stimuli. *J Neurophysiol, 104*, 960–71.
- Li, X., Lu, Z.-L., Tjan, B. S., Doshier, B. a., & Chu, W. (2008). Blood oxygenation level-dependent contrast response functions identify mechanisms of covert attention in early visual areas. *Proc Natl Acad Sci USA, 105*, 6202–7.
- Ling, S., Liu, T., & Carrasco, M. (2009). How spatial and feature-based attention affect the gain and tuning of population responses. *Vision Res, 49*, 1194–1204.
- MacEvoy, S., Tucker, T., & Fitzpatrick, D. (2009). A precise form of divisive suppression supports population coding in the primary visual cortex. *Nat Neurosci, 12*, 637–645.
- Martínez-Trujillo, J., & Treue, S. (2002). Attentional modulation strength in cortical area MT depends on stimulus contrast. *Neuron, 35*, 365–70.
- Martínez-Trujillo, J., & Treue, S. (2004). Feature-based attention increases the selectivity of population responses in primate visual cortex. *Curr Bio, 14*, 744–751.
- Maunsell, J. H. R., & Treue, S. (2006). Feature-based attention in visual cortex. *Trends Neurosci, 29*, 317–22.
- McAdams, C. J., & Maunsell, J. H. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J Neurosci, 19*, 431–41.

- Miller, E., Gochin, P., & Gross, C. (1993). Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque by addition of a second stimulus. *Brain Res*, *616*, 25–29.
- Miller, E. K., & Buschman, T. J. (2013). Cortical circuits for the control of attention. *Curr Opin Neurobiol*, *23*, 216–22.
- Mitchell, J., Stoner, G., & Reynolds, J. (2004). Object-based attention determines dominance in binocular rivalry. *Nature*, *429*, 2–5.
- Mitchell, J. F., Sundberg, K. a., & Reynolds, J. H. (2007). Differential attention-dependent response modulation across cell classes in macaque visual area V4. *Neuron*, *55*, 131–41.
- Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J Neurophysiol*, *70*, 909–19.
- Murray, S. (2008). The effects of spatial attention in early human visual cortex are stimulus independent. *J Vision*, *8*, 1–11.
- Ni, A. M., Ray, S., & Maunsell, J. H. R. (2012). Tuned normalization explains the size of attention modulations. *Neuron*, *73*, 803–13.
- Ninomiya, T., Sawamura, H., Inoue, K.-I., & Takada, M. (2012). Segregated pathways carrying frontally derived top-down signals to visual areas MT and V4 in macaques. *J Neurosci*, *32*, 6851–8.
- Orban, G. (2008). Higher order visual processing in macaque extrastriate cortex. *Physiol Rev*, *88*, 59–89.
- Pack, C. C., Hunter, J. N., & Born, R. T. (2005). Contrast dependence of suppressive influences in cortical area MT of alert macaque. *J Neurophysiol*, *93*, 1809–15.
- Potjans, T. C., & Diesmann, M. (2012). The Cell-Type Specific Cortical Microcircuit: Relating Structure and Activity in a Full-Scale Spiking Network Model. *Cerebral cortex*, .
- Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *J Neurosci*, *19*, 1736–53.
- Reynolds, J. H., & Heeger, D. J. (2009). The normalization model of attention. *Neuron*, *61*, 168–85.
- Reynolds, J. H., Pasternak, T., & Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron*, *26*, 703–14.
- Saenz, M., Buracas, G. T., & Boynton, G. M. (2002). Global effects of feature-based attention in human visual cortex. *Nat Neurosci*, *5*, 631–2.

- Spratling, M. W. (2008). Predictive coding as a model of biased competition in visual attention. *Vision Res*, *48*, 1391–408.
- Spratling, M. W. (2010). Predictive coding as a model of response properties in cortical area V1. *J Neurosci*, *30*, 3531–43.
- Spratling, M. W. (2012). Unsupervised learning of generative and discriminative weights encoding elementary image components in a predictive coding model of cortical function. *Neural Comput*, *24*, 60–103.
- Spratling, M. W., & Johnson, M. H. (2004). A feedback model of visual attention. *J Cognitive Neurosci*, *16*, 219–37.
- Steenrod, S. C., Phillips, M. H., & Goldberg, M. E. (2013). The lateral intraparietal area codes the location of saccade targets and not the dimension of the saccades that will be made to acquire them. *J Neurophysiol*, *109*, 2596–605.
- Sundberg, K. A., Mitchell, J. F., Gawne, T. J., & Reynolds, J. H. (2012). Attention influences single unit and local field potential response latencies in visual cortical area V4. *J Neurosci*, *32*, 16040–50.
- Sundberg, K. A., Mitchell, J. F., & Reynolds, J. H. (2009). Spatial attention modulates center-surround interactions in macaque visual area v4. *Neuron*, *61*, 952–63.
- Todd, J. J., Fougner, D., & Marois, R. (2005). Visual short-term memory load suppresses temporo-parietal junction activity and induces inattentive blindness. *Psychol Sci*, *16*, 965–72.
- Treue, S., & Trujillo, J. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, *399*, 575–579.
- Wagatsuma, N., Potjans, T. C., Diesmann, M., Sakai, K., & Fukai, T. (2013). Spatial and feature-based attention in a layered cortical microcircuit model. *PLoS One*, *8*, e80788.
- Whitney, D., & Levi, D. M. (2011). Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends Cogn Sci*, *15*, 160–8.
- Williford, T., & Maunsell, J. H. R. (2006). Effects of spatial attention on contrast response functions in macaque area V4. *J Neurophysiol*, *96*, 40–54.
- Willmore, B. D. B., Prenger, R. J., & Gallant, J. L. (2010). Neural representation of natural images in visual area V2. *J Neurosci*, *30*, 2102–14.
- Zanos, T. P., Mineault, P. J., Monteon, J. a., & Pack, C. C. (2011). Functional connectivity during surround suppression in macaque area V4. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 3342–5). volume 2011.
- Zhou, H., & Desimone, R. (2011). Feature-Based Attention in the Frontal Eye Field and Area V4 during Visual Search. *Neuron*, *70*, 1205–17.

- Zirnsak, M., Beuth, F., & Hamker, F. H. (2011). Split of spatial attention as predicted by a systems-level model of visual attention. *Eur J Neurosci*, *33*, 2035–45.
- Zirnsak, M., & Hamker, F. H. (2010). Attention Alters Feature Space in Motion Processing. *J Neurosci*, *30*, 6882–6890.