

# A model of surround suppression through cortical feedback

Thomas J. Sullivan<sup>a,\*</sup>, Virginia R. de Sa<sup>b</sup>

<sup>a</sup> Department of Electrical Engineering, University of California, San Diego, La Jolla, CA, USA

<sup>b</sup> Department of Cognitive Science, University of California, San Diego, 9500 Gilman Dr MC 0515, La Jolla, CA 92093, USA

Received 3 January 2005; accepted 21 December 2005

## Abstract

Surround suppression occurs when a visual stimulus outside a neuron's classical receptive field causes a reduction in firing rate. It has become clear that several mechanisms are working together to induce center-surround effects such as surround suppression. While several models exist that rely on lateral connections within V1 to explain surround suppression, few have been proposed that show how cortical feedback might play a role. In this work, we propose a theory in which reductions in excitatory feedback contribute to a neuron's suppressed firing rate. We also provide a computational model that incorporates this idea.

© 2006 Elsevier Ltd. All rights reserved.

*Keywords:* Surround suppression; Contextual modulation; Feedback; Cortical feedback

## 1. Introduction

A visual stimulus can excite a neuron in the primary visual cortex if the stimulus is within the neuron's receptive field. Stimuli presented outside a neuron's receptive field would not by themselves excite the neuron, but these stimuli can modulate the response of a neuron that is already firing. These modulatory visual stimuli presented in the area outside the receptive field, called the surround, typically suppress the response of an active neuron (Cavanaugh, Bair, & Movshon, 2002a,b). Thus, this effect is called surround suppression.

Besides this suppressive effect, several other related experimental results have been reported (for reviews, see the papers by Seriés, Lorenceau and Frégnac (2003), Shapley (2004), Angelucci and Bullier (2003)). Under some conditions, the surround may cause facilitation as opposed to suppression (Polat, Mizobe, Pettet, Kasamatsu, & Norcia, 1998). The data suggest that this happens much more when the stimuli have low contrast as opposed to high contrast. An experiment by Polat et al. (1998) used localized Gabor patches in the surround of the same contrast as a center Gabor patch. They were able to show examples of neurons that were facilitated by the surround at low stimulus contrast, but suppressed by the same stimulus at higher contrast. Also, it seems that the very size of a neuron's

receptive field can change. If the stimuli are low contrast, the neuron's receptive field will be larger than if the stimuli are high contrast (Cavanaugh et al., 2002a; Kapadia, Westheimer, & Gilbert, 1999; Sceniak, Ringach, Hawken, & Shapley, 1999). (This result seems to hold for the various ways of measuring receptive field size (Cavanaugh et al., 2002a)).

Why is this surround influence on neuronal response important to investigate? One reason is that these effects may tell us something about the underlying cortical microcircuitry of primary visual cortex. The most straight-forward models of the primary visual cortex involve feedforward connections to linear Gabor-like filters. This type of model cannot describe these surround suppression findings. Understanding how various mechanisms contribute to these surround suppression effects will reveal what other mechanisms are at work, including lateral interactions and feedback from higher cortical areas. Also, center-surround interactions may play important functional roles in visual processing. It has been suggested that center-surround effects may be related to the perceptual integration of contours. In support of this idea many psychophysics studies have shown directly that surround stimuli can have an influence on how we perceive a visual input (Bakin, Nakayama, & Gilbert, 2000; Polat & Sagi, 1993).

Cavanaugh et al. (2002a) and Sceniak et al. (1999) proposed a model that largely reproduces the suppressive effects. This model is based on Gaussian center and surround components. The response from the center component is divided or subtracted by the response from the surround component. This is similar to the model proposed by Schwartz and

\* Corresponding author. Tel.: +1 858 373 8954; fax: +1 858 534 1128.  
E-mail address: [tom@sullivan.to](mailto:tom@sullivan.to) (T.J. Sullivan).

Simoncelli (2001). In this case, a filter's feedforward excitation is divided by a quantity that depends on the outputs of the nearby filter responses. These models show suppression as a response to surround stimuli.

None of these high-level models propose underlying neural mechanisms. Creating a realistic neural model will allow us to gain a deeper understanding and make more detailed predictions. Several compelling neuron-level models have been proposed that reproduce some of the center-surround effects (Dragoi & Sur, 2000; Somers et al., 1998; Stemmler, Usher, & Niebur, 1995; Ursino & Cara, 2004). These models propose that lateral connections connecting the neurons responding to the center and surround stimuli can play a major role in creating center-surround effects. This seems logical and is backed up by evidence matching the scale of facilitation in psychophysics studies to the distribution of lateral connections in V1 (Stettler, Das, Bennett, & Gilbert, 2002).

There is evidence that another mechanism plays a role in surround suppression: “the suppressive surrounds of V1 neurons are derived in part from feedback signals. It is a challenge for theorists to understand what is the function of this feedback in visual perception” (Shapley, 2004). This view is supported by the anatomical findings that the lateral connections within a layer do not extend far enough to account for all surround suppression effects (Angelucci et al., 2002; Cavanaugh et al., 2002a). Furthermore, it was shown in a separate study that lateral connections would not be fast enough to account for the time course of the suppression (Bair, Cavanaugh, & Movshon, 2003).

It is not obvious how feedback from higher cortical areas can contribute to surround suppression and the other center-surround effects. One clue comes from the finding in rat physiology that the feedback pathway is balanced more toward excitation than feedforward pathways (Johnson & Burkhalter, 1996; Johnson & Burkhalter, 1997; Shao & Burkhalter, 1996). In these results, excitation dominated inhibition in the cortical feedback pathway, much more so than in the feedforward pathway. This is consistent with the monkey MT and V2 results of Hupé et al. (1998), Hupé, James, Girard, and Bullier (2001), and Mignard and Malpeli (1991). They showed that many neurons showed a significant decrease in response when MT and V2 were inactivated. Of interest here is how is it possible that excitatory feedback can be responsible for a suppression of V1 response?

Motivated by the findings of Johnson and Burkhalter (1996), Johnson, Shao, and Burkhalter (1997), Shao and Burkhalter (1996), the theory proposed here begins by assuming that higher cortical areas like V2 feed back mostly excitation to V1. Groups of neurons in V1 and V2 that have been activated together often in the past facilitate each other's firing rate. What appears to be neuronal suppression is actually a decrease in this feedback excitation. A surround stimulus can trigger this decrease by causing a competition for representation among V2 neurons. With no surround stimulus, a center stimulus representation in V2 wins easily and feeds back excitation to the center neurons in V1. When a surround stimulus is present,

that stimulus fights for representation in V2 through lateral inhibition. The resulting competition decreases the excitation given to the original neurons.

A computational model incorporating these ideas is presented in the following sections. It is intended to test whether cortical feedback can play a role in suppression in the way just described. It is different from the feedback model offered by Grossberg and Raizada (2000). In that model, a folded feedback mechanism is proposed in which excitatory feedback from higher areas is directed through layer 6 and subsequently fed through excitatory and inhibitory connections to layer 4. The result is a modulatory pathway that has excitation in the center and inhibition in the surround. In our model, V2 neurons feedback directly with excitation into the feedforward V1 pathway (layer 4, effectively). Only center excitation is present. It is also different from the work of Rao and Ballard (1999). Their V2 neurons also learn associations between V1 outputs, but those neurons feed back inhibition to cancel out the responses of the V1 neurons. An important note to keep in mind is that it seems that multiple mechanisms are at work producing center-surround effects. The purpose for this model is to examine the role of excitatory cortical feedback. Future work should investigate the relative roles and interaction of all involved mechanisms.

## 2. Model

The model consists of six small regions of neurons (see Fig. 1). Five of the regions represent V1 hypercolumns and the other represents a hypercolumn in V2. Each of the five V1 regions receive feedforward input from limited, non-overlapping receptive fields in the visual input space. The V1 regions have no direct lateral interaction with each other since this model is limited to the investigation of the potential role of feedback in surround suppression. The V1 regions can only interact with each other indirectly through feedforward and feedback connections between them and the V2 region.

Each V1 region contains a grid of  $16 \times 16$  rate-based neurons. All of the neurons within a region share the same receptive field. The receptive fields contain a  $15 \times 15$  grid of pixels, so each neuron has 225 ( $15 \times 15$ ) incoming weights. The activity of one excitatory neuron is computed as a dot

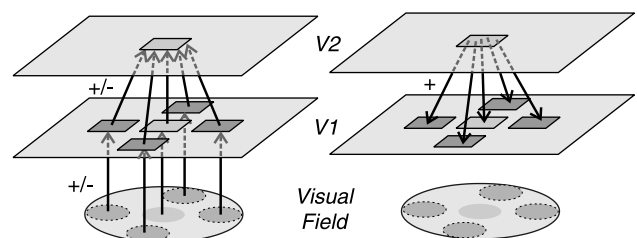


Fig. 1. The center-surround model. Five V1 hypercolumn-sized neuron regions are simulated along with one V2 region. For visual clarity the model is shown twice: on the left with only the feedforward connections shown and on the right with only feedback connections shown. The feedforward connections are excitatory and inhibitory, while the feedback connections are only excitatory. The V1 regions influence each other only indirectly through interaction with the V2 region.

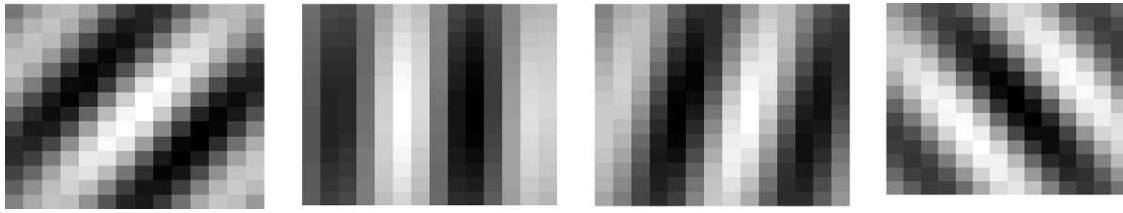


Fig. 2. Example input images for V1 training. The  $15 \times 15$  pixel images are sine-wave gratings of random orientation and phase offset.

product modified by a logistic function, given by:

$$A_k = f\left(\sum_J I_j w_{kj}\right),$$

where  $I_j$  is one of the  $J$  pixels in the receptive field, and there are  $J$  incoming weights,  $w_{kj}$ , into the  $k$ th neuron.  $f(x)$  is the logistic function given by:

$$f(x) = \frac{f_{\max}}{1 + Ce^{-bx}}.$$

$C$  and  $b$  are parameters that control the shape of the logistic function.  $C$  sets the offset and  $b$  sets the steepness of the curve. Their values are discussed later.

Before running the model the incoming weights of the V1 neurons of each region were trained using a standard self-organizing map (SOM) algorithm (Kohonen, 2001). This was done in order to create realistic patterns of activation in V1. The SOM Toolbox for Matlab (Alhoniemi, Himberg, Parhankangas, & Vesanto, 1999) was used for training. The input training images are sine-wave gratings of random orientation and phase offset, such as in Fig. 2. The network was trained on five runs through the data set, which included 2000 grating images. The topology of the network was set to a toroid with an initial learning rate,  $\alpha$ , of 1.5 (decreasing toward 0 throughout the training). One set of weights was generated for the center V1 region. A separate set of weights was generated with the same type of inputs but a different random assignment of initial weights. This second set of weights is used by all four surround V1 regions for computational efficiency. The resulting trained weights give an indication of the stimuli that excite the V1 neurons. Fig. 3 shows the generated weights of the center V1 region and the surround V1 regions.

It is proposed that lateral inhibition in the higher cortical area plays a key role in surround suppression. It is therefore necessary to include more detail in the V2 region model. The V2 region is constructed with a grid of  $16 \times 16$  excitatory neurons, superimposed on a grid of  $8 \times 8$  inhibitory neurons, all of which are rate-based. This gives a ratio of 4:1 excitatory to inhibitory neurons, which is consistent with physiology (Braitenberg & Schuz, 1998). Each neuron is connected to its neighbors with weights that drop off with distance in a Gaussian fashion (close neighbors have stronger connections than far-away neighbors). There are connections from excitatory neurons to excitatory neurons, excitatory to inhibitory, and inhibitory to excitatory. In all cases the weights drop off with distance, but the height and width of the Gaussian

shape are set to different values for each set of connections. The dimensions of these Gaussians are discussed later in Section 3. The neuron region is connected as a torus in order to avoid edge effects.

The activity of one excitatory neuron in V2 is calculated by summing the contributions of the feedforward connections, the lateral excitatory connections, and the lateral inhibitory connections, as in:

$$A_e = f\left(\sum \text{FF} + \sum \text{LAT}_{\text{EE}} - \sum \text{LAT}_{\text{IE}}\right)$$

More precisely, if there are  $L$  excitatory neurons and  $M$  inhibitory neurons in the V2 region and  $K$  neurons in each of

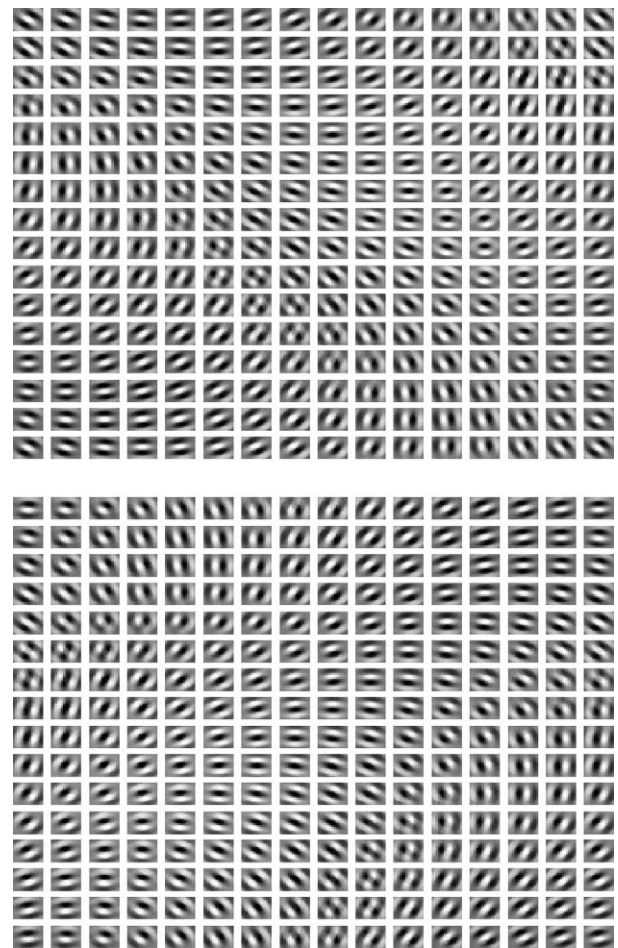


Fig. 3. Incoming weights of V1 center and surround regions after training. The top map shows the incoming weights of the V1 center region after training. Each square corresponds to the weights of one of the V1 neurons. The bottom map shows the incoming weights of one of the V1 surround regions. These weights are used for all the surround regions for computational efficiency.

the V1 regions, then

$$A_e = f\left(\left(\sum_{5\text{reg}} \sum_K A_k w_{ek}\right) + \sum_L A_l w_{el} - \sum_M A_m w_{em}\right),$$

where  $\sum_{5\text{reg}}$  indicates the summing of the contributions of all five V1 regions. In this equation,  $A_k$ ,  $A_j$ , and  $A_m$  are the neuron activities from the V1 regions, excitatory neurons in V2, and inhibitory neurons in V2, respectively. The  $w$ 's are the associated weights. Similarly, the activity of the inhibitory neurons is given by:

$$A_i = f\left(\left(\sum_{5\text{reg}} \sum_K A_k w_{ik}\right) + \sum_L A_l w_{il}\right).$$

$f(x)$  is the same logistic function given above. However, the excitatory and inhibitory populations have different offset, gain, and maximum firing rate values obtained from the literature, as discussed below.

All five V1 regions send feedforward projections to the V2 region. The feedforward projections originate from the V1 neurons and terminate on both excitatory and inhibitory neurons within the V2 region. The output weights extending from one V1 neuron take on a Gaussian pattern, which is identical for all the V1 neurons (except for a shift in location). The V2 contributions from all five V1 regions are summed together to form the feedforward input to the V2 region.

The feedback projections originate from the excitatory neurons of the V2 neurons. This feedback uses a similar Gaussian weight pattern with all positive weights. Since inhibitory neurons are not modelled explicitly in the V1 regions, the feedback is exclusively excitatory in nature. This is an exaggeration of the known anatomy discussed in the introduction. Using exclusively excitatory feedback serves to highlight one of the main themes of this model: that feedback can cause surround suppression by reducing the amount of excitatory feedback excitation produced. The results would be the same if the feedback were mostly excitatory, as opposed to exclusively excitatory.

With both the V1-to-V2 projections, as well as the V2-to-V1 projections, a more realistic model would include the learning of associations between groups of neurons. When a group of neurons is activated in one of the V1 regions, which neurons in the V2 region will be driven to excitation? In the cortex, this may be a matter decided by the learning process. Groups of neurons in V1 will converge onto V2 neurons to form V2 receptive field properties. These connections likely reflect the statistics of the inputs used during training (natural visual stimuli, in the case of the human visual system). For example, if certain neuronal activations in the V1 regions coincided often, we predict they would come to be strongly connected to the same V2 neurons and that this arrangement would lead to firing rate facilitation of V2 neurons. Center and surround stimuli that were not seen together during development would not converge onto the same V2 neurons. In this case, lateral inhibition in V2 would result. Measuring or deriving the connectivity patterns of V2 neurons is currently a big open

question. We will leave this problem for future study in order to concentrate on surround suppression mediated by cortical feedback.

In this model, the problem of learning connections between V1 and V2 (both feedforward and feedback) is avoided. The connections are created topographically such that each neuron in the V1 regions sends out a Gaussian weight pattern to the V2 region, and vice-versa. Since, the V1 center region and the V2 surround region are trained separately with an SOM, the topographical connections result in random orientation pairs (from center and surround) converging in V2 neurons.

The activity equations of the individual neurons are highly coupled within the V2 region and between V2 and V1 regions. To compute the overall activity, an iterative algorithm is used. Using the equations described above, an old set of neuron activations leads to a new set of activations. The simulation errors are set such that forcing the simulator to converge to lower error values made no significant difference in the population's activity.

### 3. Parameter assignment

The parameters that were built into the model, and the values used in the simulations, are specified in Table 1. Within the V2 region, we use a Gaussian distribution of weights leaving an excitatory neuron that terminate on the other excitatory neurons. Each excitatory neuron in the V2 region uses the same distribution of lateral output weights. The prototype for this distribution is:

$$W_{EE} = \text{SF}_{EE} \frac{1}{2\pi\sigma^2} e^{-d^2/2\sigma^2}$$

where  $\text{SF}_{EE}$  is a unit-less scale factor,  $d$  is the distance between the two neurons which share the weight, and  $\sigma$  is the weight distribution width. The same equation was used for the excitatory-to-inhibitory weights and the inhibitory-to-excitatory weights, but with different  $\sigma$  and scale factor values. The weight distribution widths are taken from anatomical studies (Douglas, Koch, Mahowald, Martin, & Suarez, 1995; Hirsch, Alonso, Reid, & Martinez, 1998; Stettler et al., 2002). For the scale factors on the other hand, it is difficult to find values from experimental work. These were set such that there is competition between neurons through lateral inhibition. To be sure that we did not just happen across a perfect set of model parameters, these scale factors were varied over substantial ranges with similar results. Those parameter range simulations and results are described in Appendix A.

Individual neuron behavior is governed by a logistic function with three parameters: the maximum firing rate, gain, and offset. In physiology experiments the maximum firing rate of the excitatory neurons and inhibitory neurons are hard to estimate and vary considerably from neuron to neuron. The maximum firing rates chosen from the reported distributions (Ringach, Shapley, & Hawken, 2002; Stemmler et al., 1995) are set at 50 Hz for excitatory neurons and 100 Hz for inhibitory neurons. The gains were obtained by McCormick, Connors, Lighthall and Prince (1985) and are set in the model

Table 1  
Model parameters and their values

Model parameter	Value	Reference
Hypercolumn width	900 $\mu\text{m}$	Stettler et al. (2002)
E-to-E weight distribution ( $1 \sigma$ )	110 $\mu\text{m}$	Hirsch et al. (1998)
E-to-I weight distribution ( $1 \sigma$ )	110 $\mu\text{m}$	Hirsch et al. (1998)
I-to-E weight distribution ( $1 \sigma$ )	500 $\mu\text{m}$	Stettler et al. (2002)
E-to-E efficacy scale factor	1.0	Range: 0.25–1.75
E-to-I efficacy scale factor	0.55	Range: 0.30–1.10
I-to-E efficacy scale factor	85	Range: 40–130
Maximum gain of E neurons	241 Hz/nA	McCormick et al. (1985)
Maximum gain of I neurons	549 Hz/nA	McCormick et al. (1985)
Spontaneous rate of E neurons ( $f_0$ )	2 Hz	Ringach et al. (2002)
Maximum firing rate of E ( $f_{\text{max}}$ )	50 Hz	Ringach et al. (2002)
Spontaneous rate of I neurons ( $f_0$ )	1 Hz	Stemmler et al. (1995)
Maximum firing rate of I ( $f_{\text{max}}$ )	100 Hz	McCormick et al. (1985)
V2-to-V1 feedback scale factor	3.75	Range: 2.40–6.00

These values used in the model simulations were taken or derived from the references given in the right-hand column. The function of the model parameters are described in the text. For model parameters without references, ranges for valid behavior are given. This is discussed further in Appendix A.

as 241 Hz/nA for excitatory neurons and 549 Hz/nA for inhibitory neurons. These gains are used to set the maximum slope of the logistic function by noting that

$$b = 4 \frac{(df/dx)_{\text{max}}}{f_{\text{max}}}$$

where  $b$  is the logistic equation parameter,  $(df/dx)_{\text{max}}$  is the maximum neuron gain obtained from the literature, and  $f_{\text{max}}$  is the neuron's maximum firing rate. The logistic parameter  $C$ , on the other hand, determines how much input drive is required to excite the neuron. To make an appropriate choice, the firing rate with no input was examined. This was compared to typical spontaneous firing rates and again this was chosen to be something reasonable from the distributions reported (Ringach et al., 2002; Stemmler et al., 1995): 2 Hz for excitatory neurons and 1 Hz for inhibitory neurons. The logistic function parameter,  $C$ , was set by noting that

$$C = \frac{f_{\text{max}}}{f_0} - 1$$

where  $f_0$  is the firing rate with no input current.

The gains of the neuron types are reported in units of Hertz per nanoAmphere. Neurons can be thought to take a current (with units of nanoAmphere) as input and output a firing rate

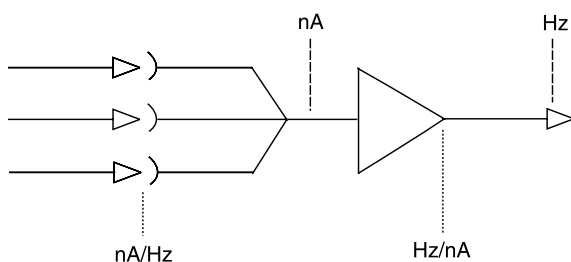


Fig. 4. Diagram showing the units used in the model. The incoming synapses of the neurons convert neuron firing rates (in Hertz) into current (in nanoAmperes). This current is summed in the dendrites and converted into a firing rate near the beginning of the axon. The output of the neuron is a firing rate (measured in Hertz).

(in Hertz). The weights in this model represent synapses with units nanoAmphere per Hertz. Therefore, an output firing rate multiplied by a weight will result in a current (the units are shown graphically in Fig. 4).

The V2-to-V1 scale factor is used to scale the weights leaving the V2 excitatory neurons and terminating on the V1 neurons. This model proposes that surround suppression is a result of lower feedback excitation. Thus, the feedback excitation sent from V2 to V1 should be roughly equivalent to the amount of suppression. The V2-to-V1 scale factor was set to a number that shows a reasonable suppression. As discussed in the Appendix, any value within a wide range works.

#### 4. Results

Computer simulations were performed with the model. Fig. 5 shows a typical visual input used during the testing.

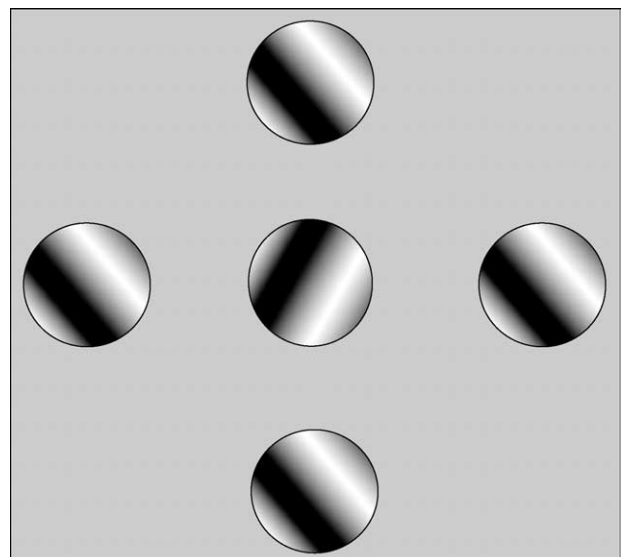


Fig. 5. Typical input test image. A sine-wave grating of varying contrast is placed in the center with a surround sine-wave grating.

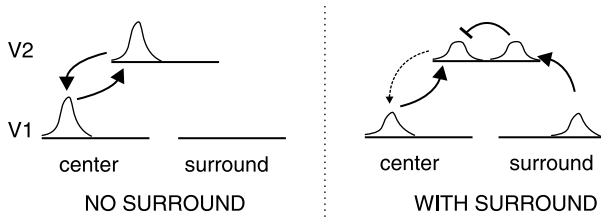


Fig. 6. Suppression mechanism in model. With no surround, the V1 center region neurons and the V2 region neurons facilitated each others firing rates. With a surround, lateral inhibition in the V2 region causes a decrease in excitatory feedback.

There is a sine-wave grating in the center surrounded by other sine-wave gratings of an independently chosen orientation. As described in Section 2, the receptive field of the center V1 neurons encompasses only the center area of the visual field. The surround V1 neurons have their receptive fields surrounding the center receptive field (see Fig. 1).

A two-dimensional cartoon drawing of typical neuron suppression is shown in Fig. 6. With only a center stimulus, a group of neurons in the V1 center region becomes active. This excites a group of neurons in the V2 region. These active neurons further enhance each others firing rates through the feedforward and excitatory feedback connections. When a surround stimulus is added, groups of neurons from the surround regions (only one surround region is shown) become active. They, too, excite some neurons in V2, but if these V2 neurons are different, competition through lateral inhibition takes place. This lateral inhibition diminishes the firing rate of all the V2 neurons, and the amount of excitatory feedback is reduced. This reduction in V1 activity is the suppression caused by the surround.

A typical example is shown in Fig. 7. A neuron’s response to its optimal stimulus is shown with increasing stimulus contrast.

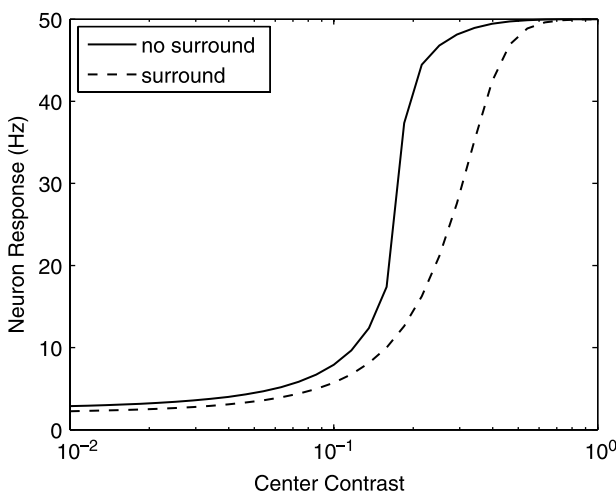


Fig. 7. Typical suppression by surround stimulus. (Solid line). A model neuron’s response to its optimal stimulus is shown with changing contrast. (Dashed line). When a surround is present, the neuron’s response is suppressed. In this case, the activity of the neurons in the center V1 region and surround V1 regions produced V2 activity in separate locations (about eight neuron grid spacings apart).

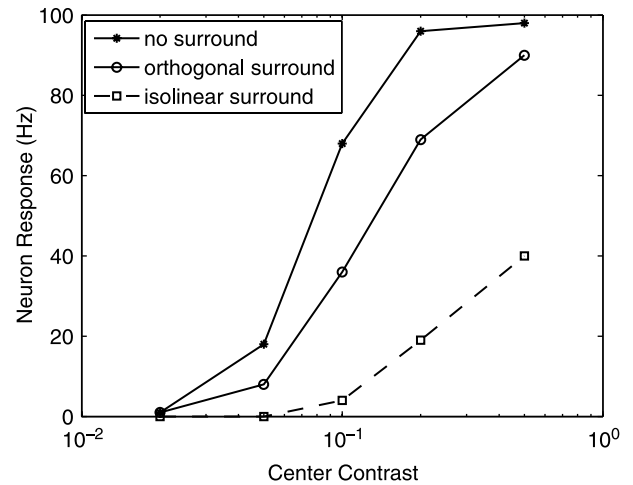


Fig. 8. Surround suppression in experiments with Macaque V1 (after Cavanaugh et al., 2002b). A neuron’s response with changing contrast is typically suppressed when a surround stimulus is introduced.

The top curve shows the response with no surround stimulus. The bottom curve shows the response of the same neuron after a surround is added. The neuron’s response is indeed suppressed. This is similar in nature to the experimental findings of Cavanaugh et al. (2002b) shown in Fig. 8.

In this case, the neurons in V2 inhibit each other and result in surround suppression. When activities from the center and surround V1 regions converge onto one group of neurons in V2, we would expect that facilitation could occur. Fig. 9 shows that this is indeed the case. In this figure, the response with no surround stimulus and changing center contrast is plotted. The neuron response is facilitated when a surround stimulus is present. While we have not seen a graph exactly like this in the literature, facilitation at low contrasts is consistent with

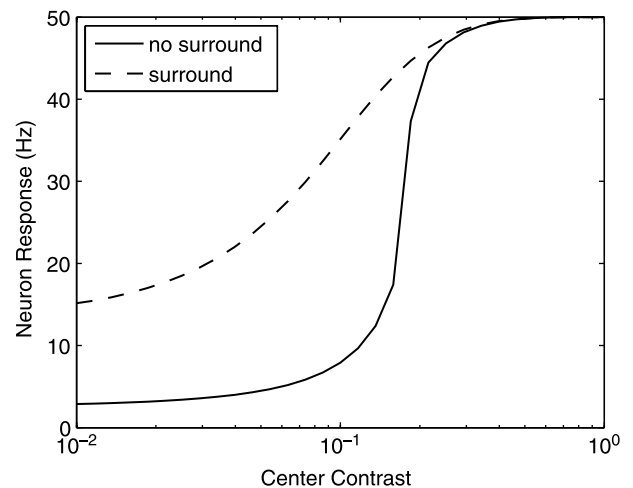


Fig. 9. Facilitation by surround stimulus. (Solid line). A model neuron’s response to its optimal stimulus is shown with changing contrast. (Dashed line). When a surround is present, the neuron’s response is facilitated. In this case, the activity of the neurons in the center V1 region and surround V1 regions converged onto the same location in V2.

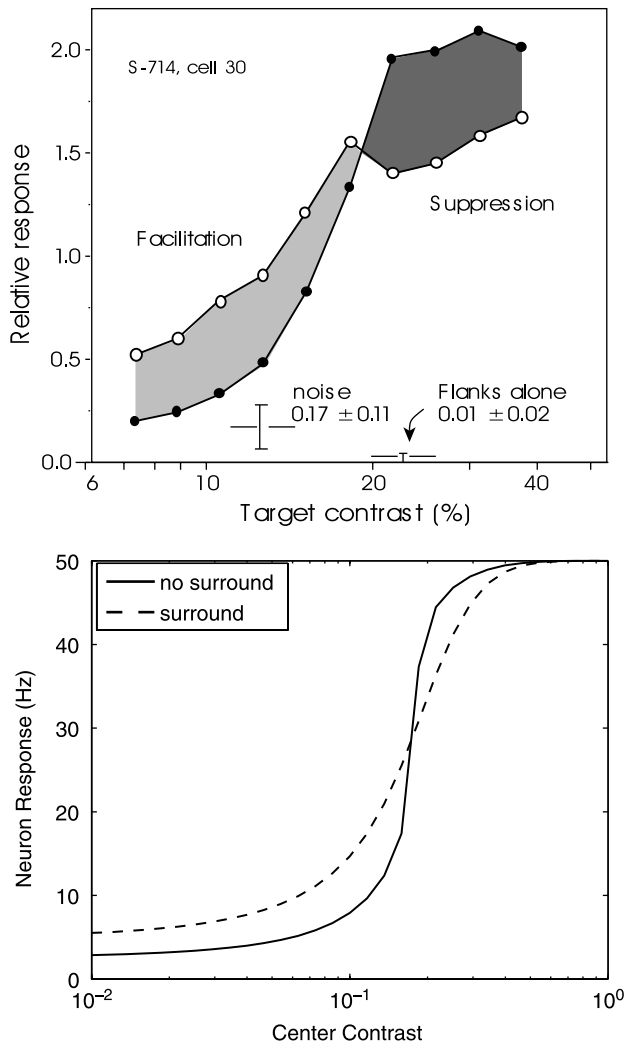


Fig. 10. Facilitation and suppression with changing contrast. (Top). Facilitation recorded in experiments on cats (taken from (Polat & Sagi, 1993) with permission). Facilitation and suppression both occur with the same surround stimulus (in this case small, collinear patches in the surround). (Solid line). A model neuron's response to its optimal stimulus is shown with changing contrast. (Dashed line). When a surround is present, the neuron's response is facilitated at low contrasts and suppressed at high contrasts. In this case, the activity of the neurons in the center V1 region and surround V1 regions converged onto the V2 region in locations being separated by five neurons.

previous experimental findings (Polat & Sagi, 1993). Our model predicts that some neurons could have contrast response functions like that shown in Fig. 9.

Fig. 10 shows an example of something in between facilitation and suppression. When the active neurons in the V1 center and surround regions do not converge exactly, but are slightly offset (with mean activities five neurons apart in V2, for example), the center-induced response in V2 will be facilitated at low contrasts but suppressed at higher contrasts. This is because at low contrasts, the inhibitory neurons in V2 are not sufficiently excited to produce appreciable inhibition, allowing facilitation through short-range excitatory connections. At higher contrasts, the inhibitory neurons (with their higher gains) overwhelm the facilitation that takes place. This

facilitation and suppression in the center-induced response of V2 is reflected back to V1 through the excitatory connections. The effect is similar to the results reported by Polat et al. (1998).

## 5. Discussion

It seems likely from experimental findings that feedback from higher cortical areas contributes to surround suppression. This model shows a plausible way in which this feedback can play a role. The model assumes that groups of neurons in V1 and V2 are linked through learning. A visual stimulus that has been seen before during training excites these linked neurons through feedforward and feedback connections. In the case of a stimulus that does not match with previous experience, lateral inhibition in V2 reduces activity and leads to less excitatory feedback. A reduction in excitatory feedback will give the appearance that V1 neuron firing rates have been suppressed. While at first glance, it is difficult to imagine how excitatory feedback from higher cortical areas can contribute to suppression in V1, this reduction in excitatory drive is an interesting candidate.

The linking together of groups of neurons through learning seems to be supported by findings of facilitation for iso-oriented, co-linear flanking stimuli (Polat et al., 1998). This is because long contours are a key component of natural scenes. While this effect is strongly influenced by lateral connections, it is not hard to imagine similar connections forming from groups of neurons in V1 to neurons in V2. At first glance, it is difficult to reconcile these facilitation findings with suppression caused by iso-oriented surround (Cavanaugh et al., 2002b). This could be explained by noting that a stimulus completely surrounding the center has both a component in line with the center and a component forming lines in parallel, but to the side, of the center. These parallel lines to the side of the center might create enough inhibition to counteract any facilitation that occurs. The facilitation would be consistent with being exposed to extended lines during visual development, while the inhibition from the sides would be consistent with not being exposed to many gratings during development.

Facilitation by orthogonal surround (Jones, Wang, & Sillito, 2002) is a trickier problem to reconcile with our model. Our model proposes that frequently observed stimuli should be facilitated and seldom observed stimuli suppressed. Developmental stimuli probably do not look much like the orthogonal surround stimuli (with the possible exception of corners). This may be due to a separate explanation. Indeed, Dragoi and Sur (2000) propose a lateral disinhibition mechanism to account for orthogonal surround facilitation. In any case, the idea that there is always suppression by iso-linear surrounds and facilitation by orthogonal surrounds is an overly simplified view. In reality, there are a wide variety of modulation patterns (see Levitt & Lund, 1997, for example).

The work of Hupé et al. (2001) was unable to report any influence of V2 on surround suppression effects and so does not support this model. This negative result, though, focuses only on near surround effects. Since far surround effects could be

dominated by cortical feedback (Angelucci et al., 2002), it would be interesting to see similar experiments using only far surrounds. This work certainly does not propose that cortical feedback is the only mechanism involved in surround suppression. It appears that several different mechanisms contribute simultaneously. Certainly, lateral connections have a strong influence on center-surround interactions, especially in the near surround. The role of feedforward connections from LGN has not yet been completely explored. Also, cortical feedback comes from several different areas to converge onto V1 (Hupé et al., 1998; Mignard & Malpeli, 1991). What this work does is propose a plausible model by which cortical feedback can contribute.

The great challenge now is to tease apart how (and how much) each mechanism contributes to center-surround interactions. One place to start may be to recognize that lateral connections may dominate the near surround interactions (Hupé et al., 2001), but the far surround influences may be dominated by cortical feedback (Angelucci et al., 2002). Using this observation, experiments may be able to show how the feedback and lateral interactions differ.

The model can be used to make a few predictions. The most straight-forward is that in some cases, neurons will be facilitated by the surround at all contrast levels, as shown in Fig. 9. This should happen for a far-surround (where feedback presumably dominates) that is very consistent with the center stimulus. Also, this model implies that the effects of facilitation/suppression from the far surround will disappear when feedback is disabled. Thus, an experiment could be done in the same style as the one by Hupé et al. (2001), but this time concentrating on the effects due only to the far surround (disabling MT as well as V2 would also be interesting). At first glance it seems that the model predicts that V2 alone can activate parts of V1, leading to hallucination or filling-in. However, this would only be the case if V2 were sufficiently activated, which might first require activation in V1. This feedback might therefore only play a modulatory role in most cases and specific predictions about inducing hallucination and filling-in are difficult to make.

Finally, we believe this is a particularly important topic for further research. Any light shed onto the roles of lateral and feedback interactions can be exploited for computer vision systems. These seem to be important mechanisms used in animal vision systems, but have not been employed much in our current efforts in man-made vision.

## Acknowledgements

The authors would like to thank Paul Sajda for helpful discussions and the anonymous reviewers for improving the manuscript. This material is based upon work supported by the National Science Foundation under NSF Career Grant no. 0133996 and an NIH Cognitive Neuroscience Predoctoral Training Grant. This work was also supported by NSF IGERT Grant no. DGE-0333451 to GW Cottrell.

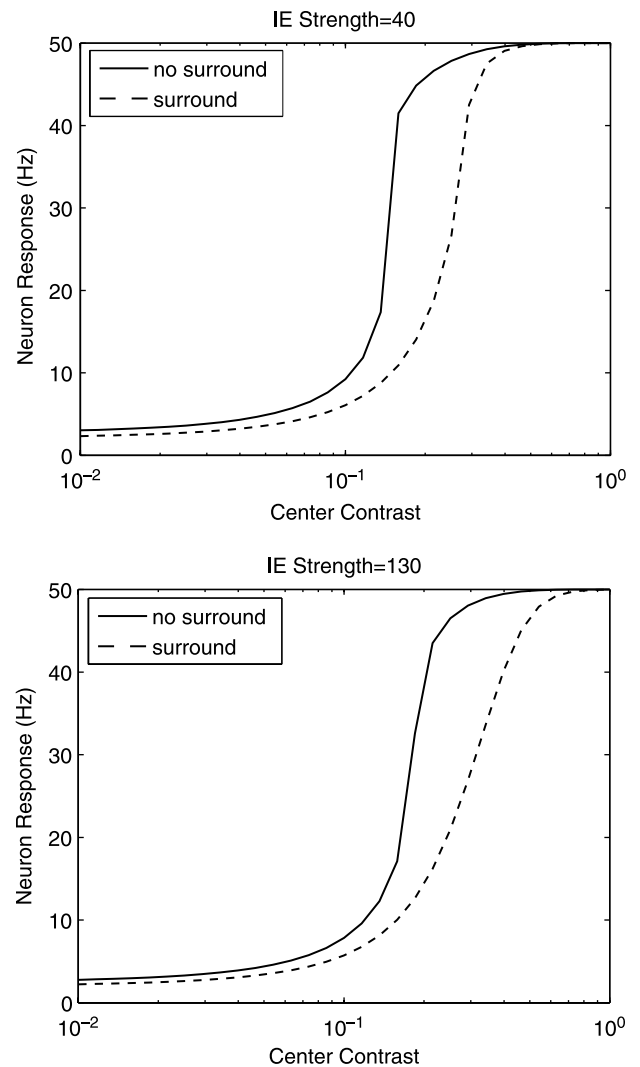


Fig. 11. Example parameter range test. The I-to-E efficacy scale factor (IE) was varied between 40 and 130. Across the range, surround suppression was qualitatively similar. (Top). IE=40. (Bottom). IE=130.

## Appendix A. Parameter ranges

In order to allay fears that an extremely narrow parameter set was found that accounts for the surround suppression findings, key parameters were systematically altered. The behavior at each of these new parameter settings was tested. For example, when the strength of the excitatory-to-inhibitory

Table 2  
Model parameter ranges

Model parameter	Low value	Nominal value	High value
E-to-E efficacy	0.25	1.00	1.75
E-to-I efficacy	0.30	0.55	1.10
I-to-E efficacy	40	85	130
V2-to-V1 scale factor	2.40	3.75	6.00

These model parameters were tested over the indicated ranges. The suppressive behavior was qualitatively similar over these parameter ranges.



neurons in the V2 model is changed from 85 to 40, a typical suppression curve still shows qualitatively the same suppression (as seen in the top of Fig. 11). The ranges that were verified to give the same qualitative results are shown in Table 2. When the parameters went beyond these ranges, either the behavior started to change or the model was not tested since the parameter range was deemed large enough to give a robust model.

## References

- Alhoniemi, E., Himberg, J., Parhankangas, J., & Vesanto, J. (1999). *Som toolbox*. Available at: <http://www.cis.hut.fi/projects/somtoolbox>.
- Angelucci, A., & Bullier, J. (2003). Reaching beyond the classical receptive field of V1 neurons: Horizontal or feedback axons? *Journal of Physiology Paris*, 97, 141–154.
- 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. *The Journal of Neuroscience*, 22(19), 8633.
- Bair, W., Cavanaugh, J. R., & Movshon, J. A. (2003). Time course and time–distance relationships for surround suppression in macaque v1 neurons. *The Journal of Neuroscience*, 23(20), 7690.
- Bakin, J. S., Nakayama, K., & Gilbert, C. D. (2000). Visual responses in monkey areas v1 and v2 to three-dimensional surface configurations. *The Journal of Neuroscience*, 20(21), 8188.
- Braitenberg, V., & Schuz, A. (1998). *Cortex: Statistics and geometry of neuronal connectivity*. Berlin: Springer.
- 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. *Journal of Neurophysiology*, 88, 2530.
- Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002b). Selectivity and spatial distribution of signals from the receptive field surround in macaque v1 neurons. *Journal of Neurophysiology*, 88, 2547.
- Douglas, R. J., Koch, C., Mahowald, M., Martin, K. A. C., & Suarez, H. H. (1995). Recurrent excitation in neocortical circuits. *Science*, 269(5226), 981–985.
- Dragoi, V., & Sur, M. (2000). Dynamic properties of recurrent inhibition in primary visual cortex: Contrast and orientation dependence of contextual effects. *Journal of Neurophysiology*, 83(2), 1019–1030.
- Grossberg, S., & Raizada, R. D. S. (2000). Contrast-sensitive perceptual grouping and object-based attention in the laminar circuits of primary visual cortex. *Vision Research*, 40(10–12), 1413–1432.
- Hirsch, J. A., Alonso, J.-M., Reid, R. C., & Martinez, L. M. (1998). Synaptic integration in striate cortical simple cells. *The Journal of Neuroscience*, 18(22), 9517–9528.
- Hupé, J., James, A., Payne, B., Lomber, S., Girard, P., & Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by v1, v2 and v3 neurons. *Nature*, 394, 784–787.
- Hupé, J.-M., James, A. C., Girard, P., & Bullier, J. (2001). Response modulations by static texture surround in area v1 of the macaque monkey do not depend on feedback connection from v2. *Journal of Physiology*, 85, 146–163.
- Johnson, R., & Burkhalter, A. (1996). Microcircuitry of forward and feedback connections within rat visual cortex. *Journal of Comparative Neurology*, 368(3), 383.
- Johnson, R. R., & Burkhalter, A. (1997). A polysynaptic feedback circuit in rat visual cortex. *The Journal of Neuroscience*, 17(18), 7129–7140.
- Jones, H. E., Wang, W., & Sillito, A. M. (2002). Spatial organization of magnitude of orientation contrast interactions in primate v1. *Journal of Neurophysiology*, 88, 2796–2808.
- Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (1999). Dynamics of spatial summation in primary visual cortex of alert monkeys. *Proceedings of the National Academy of Science*, 96(21), 12073–12078.
- Kohonen, T. (2001). *Self-organizing maps* (3rd ed.). Berlin: Springer.
- Levitt, J. B., & Lund, J. S. (1997). Contrast dependence of contextual effects in primate visual cortex. *Nature*, 387, 73–76.
- McCormick, D. A., Connors, B. W., Lighthall, J. W., & Prince, D. A. (1985). Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons of the neocortex. *Journal of Neurophysiology*, 54(4), 782–806.
- Mignard, M., & Malpeli, J. G. (1991). Paths of information flow through visual cortex. *Science*, 251, 1249–1251.
- Polat, U., Mizobe, K., Pettet, M. W., Kasamatsu, T., & Norcia, A. M. (1998). Collinear stimuli regulate visual responses depending on cell's contrast threshold. *Nature*, 391, 580–584.
- Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. *Vision Research*, 33(7), 993–999.
- Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive field effects. *Nature Neuroscience*, 2(1), 79–87.
- Ringach, D. L., Shapley, R. M., & Hawken, M. J. (2002). Orientation selectivity in macaque v1: Diversity and laminar dependence. *The Journal of Neuroscience*, 22(13), 5639.
- Sceniak, M. P., Ringach, D. L., Hawken, M. J., & Shapley, R. (1999). Contrast's effect on spatial summation by macaque v1 neurons. *Nature Neuroscience*, 2(8), 733–739.
- Schwartz, O., & Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control. *Nature Neuroscience*, 4(8), 819–825.
- Seriés, P., Lorenceau, J., & Frégnac, Y. (2003). The 'silent' surround of v1 receptive fields: Theory and experiments. *Journal of Physiology Paris*, 97, 453–474.
- Shao, Z., & Burkhalter, A. (1996). Different balance of excitation and inhibition in forward and feedback circuits of rat visual cortex. *The Journal of Neuroscience*, 16(22), 7353.
- Shapley, R. (2004). A new view of the primary visual cortex. *Neural Networks*, 17(5–6), 615–623.
- Somers, D. C., Todorov, E. V., Siapas, A. G., Toth, L. J., Kim, D.-S., & Sur, M. (1998). A local circuit approach to understanding integration of long-range inputs in primary visual cortex. *Cerebral Cortex*, 8, 204–217.
- Stemmler, M., Usher, M., & Niebur, E. (1995). Lateral interactions in primary visual cortex: A model bridging physiology and psychophysics. *Science*, 269, 1877–1880.
- Stettler, D. D., Das, A., Bennett, J., & Gilbert, C. D. (2002). Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron*, 36, 739–750.
- Ursino, M., & Cara, G. E. L. (2004). A model of contextual interactions and contour detection in primary visual cortex. *Neural Networks*, 17(5–6), 719–735.