

Modeling the Nature of Center-Surround Interactions in Early Visual Cortex

Sunday Oluwaseun Gbenro
University of Nice - Sophia Antipolis

24 July, 2009

Submitted in partial fulfillment of Mathmods Masters Programme at Unice

INTRODUCTION

The structure and nature of the of visual systems has been of great interest in computational neuroscience. Quite a number of authors have devoted work to understand and quantify the underlying principles of visual perception. The model we study here follow the steps proposed by Eero Simoncelli (see reference).

Contrast sensitivity, orientation selectivity and spatial frequencies are known to be characterises of center-surround modulations in the visual cortex. We examine how these properties affect suppression and facilitation of the response of the center to stimulus presented in the classical receptive field CRF. We capture in a quantitative model, the rich behavior of center-surround interactions at the level of primary visual cortex V1 through consideration of surround inhibition and excitation that occur at V1. We showed that a divisive normalization model with weights determined from optimization procedure can justify some proposition that for low spatial frequency we have weights that is more diffused than for higher spatial frequency.

Visual Perception

Visual images are formed from light reflected from different surfaces in the environment. The light from these surfaces is typically combined in a non-linear way. Image perception consists of two basic steps:

- capturing the image with the eye,
- recognizing and interpreting the image with the visual cortex in the brain.

We shall be primarily concerned with the second step in this report.

The Visual Field

More than one-quarter of the human cortex is thought to be involved in processing visual information, which shows the importance of visual sensation in human perception. The primary visual cortex V1, also known as the striate cortex, receives inputs from the lateral geniculate nucleus which is connected to the photoreceptors

in the retina. V1 is linked with area V2 by a very dense set of reciprocal connections. Beyond areas V1 and V2 which are the two largest cortical areas of the primate visual system is a large number of smaller cortical areas that contain neurons with selectivity to parameters such as color, depth, direction of motion, etc. Many of these areas during activation can now be observed directly in human by magnetic resonance imaging.

The visual areas with their corresponding functional specialization are: V1 -basic edge detection area , V2 stereo, V3 color, V3a is texture segmentation area, V3b for segmentation and grouping, V4 is Recognition area, V7 for face recognition, MT is general “motion area“, MST is working memory/mental imagery area.

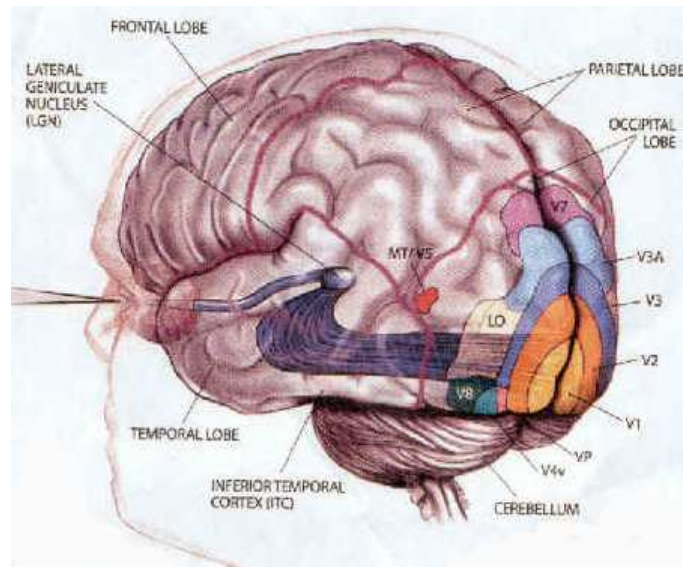


Figure 1: Human brain

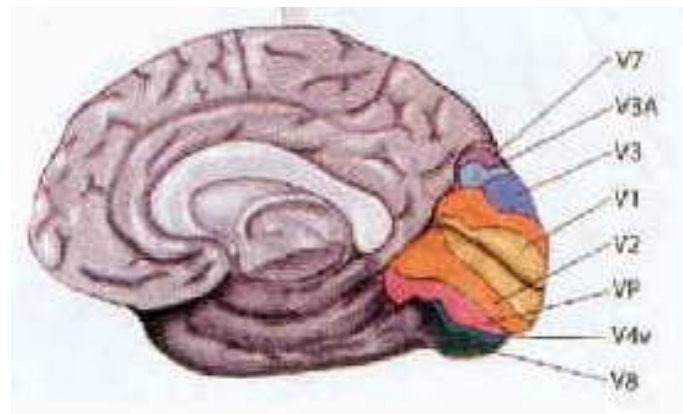


Figure 2: Visual field

Processing Visual information

Retina and lateral geniculate cells have center-surround receptive fields, while most neurons in the visual cortex respond to more complex stimuli. Uniform illumination of the whole receptive field gives, on average, little excitation; neurons in the visual cortex are stimulated best by bars or edges of a particular orientation or spatial frequency or specific color or movement information about disparity of images on the two retinas is used to give stereoscopic perception depth and distance. Neurons of the visual cortex with closely related functions are arranged anatomically in columns of cells responding to the same region of the retina, dominated by the same eye.

Both feedforward and feedback processing are ubiquitous in Visual system; the feedforward and the feedback models. The feedforward connections transfer information from the low order areas to higher order areas and the feedback connections carry information in the reverse direction.

According to the feedforward model, neurons in areas V1 and V2 perform local computations on a 2D representation of the visual image and global 3D representations are assumed to be achieved in higher order areas through a succession of filters corresponding to neurons with very sophisticated properties. Due to the difficulty of segmentation of a visual scene when the image is ambiguous, the need for a model in which the local analysis and global precept can be combined is necessary. Hence, the feedback model was develop, not to replace the feedforward model but, to compensate for the shortcomings.

Image Decomposition

By image decomposition, we mean splitting an image into the set of statistically and independent components, containing details of different information classes.

There are several methods of achieving this, we discuss four common image decomposition procedures and, show with examples, the advantages of each of them.

Wavelet Pyramid

Wavelets are multi-scale linear transforms for image representation. The basis functions of these representations are localized in spatial position, orientation and spatial frequency (scale). The coefficients resulting from projection of natural images onto these functions are essentially uncorrelated with non-Gaussian marginal statistics.

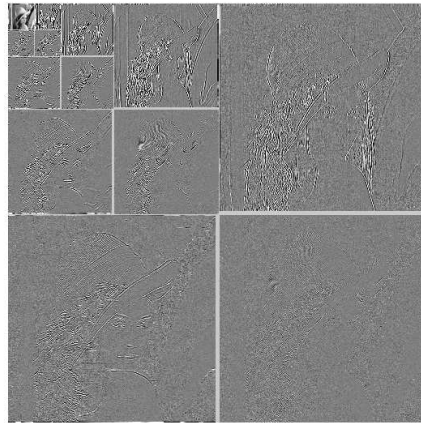


Figure 3: Wavelet pyramid.

Steerable Pyramid

Steerable Pyramid is a linear multi-scale, multi-orientation image decomposition that provides a useful front-end for image processing. This was developed in order to overcome the limitations of orthogonal wavelet decompositions (it does not represent oblique orientation well).

The basis functions of the steerable pyramid are directional derivative operators, that come in different sizes and orientations. Below is an example of decomposition of an image Lena.tif using a steerable pyramid containing 4 orientation subbands, at 2 scales.

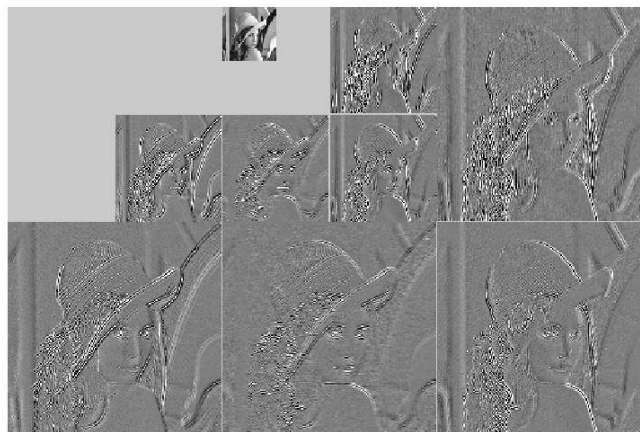
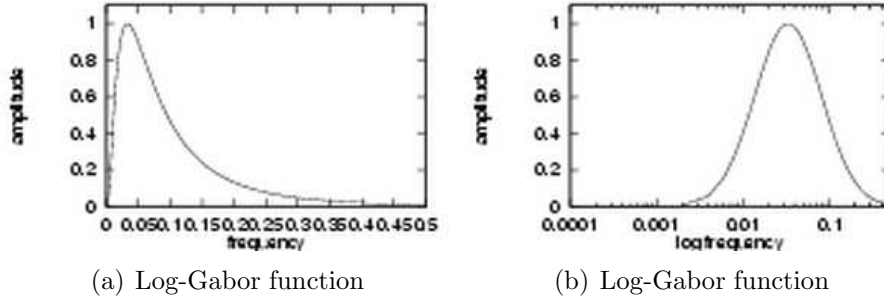


Figure 4: Steerable pyramid.

The Log-Gabor filters

Gabor filters are very good choice for obtaining best simultaneous localization of spatial and frequency information. Gabor filter have Gaussian transfer function when viewed on the linear frequency scale. It has Small bandwidth and so not optimal when we seek broad spectral information with maximal spatial localization.



The Log-Gabor function provides a better alternative to Gabor function. It was suggested by Field [1987], that natural images are better coded by filters that have Gaussian transfer function when viewed on the logarithmic frequency scale. Log-Gabor can be constructed arbitrary bandwidth and the bandwidth can be optimised to produce a filter with minimal spatial extent. The transfer function has a form

$$G(\omega) = e^{(-\log(\omega/\omega_0)^2)/(2(\log(k/\omega_0)^2))} \quad (1)$$

where ω_0 is the filter's center frequency.

The Log-Gabor functions have no DC component and it has an extended tail at the high frequency end. Also, Log-Gabor function is consistent with measurements on mammalian visual systems which indicate we have cell responses that are symmetric on the log frequency scale. The following are Gabor filters of different orientation with the corresponding filtered Lena images.

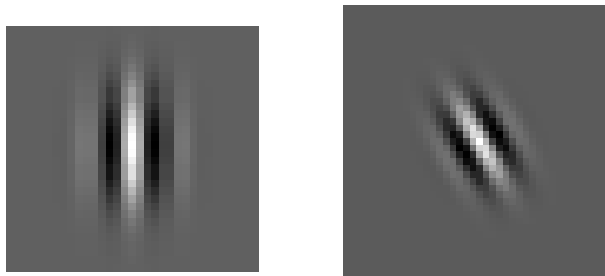


Figure 5: Gabor filter oriented at 0° & 30° .

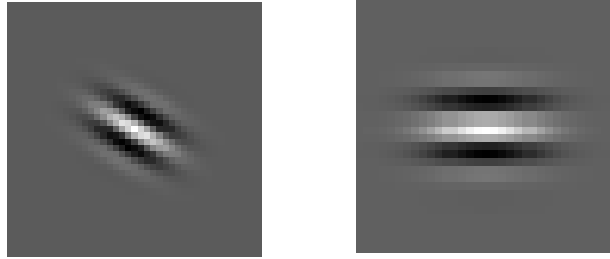


Figure 6: Gabor filter oriented at 60° & 90° .

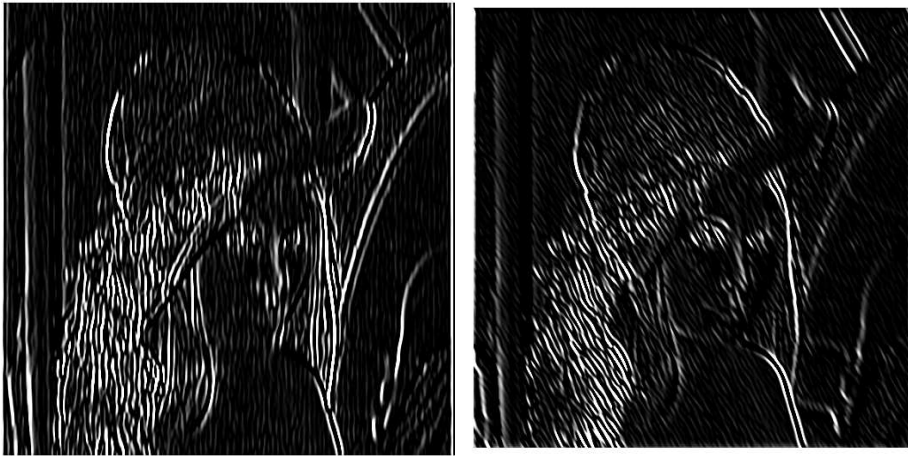


Figure 7: Gabor pyramid

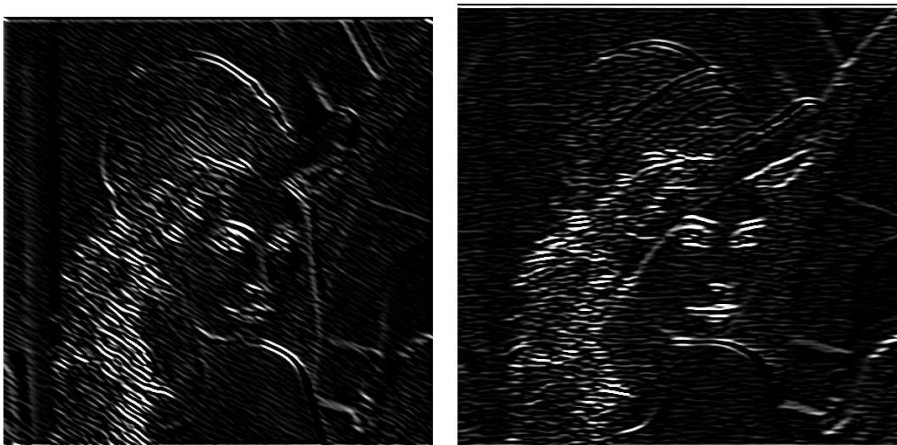


Figure 8: Gabor pyramid

Models on Center-Surround Modulations

Retina has center-surround receptive fields which size changes depending on the varying contrast of center-surround stimuli. Uniform illumination of the whole receptive field gives, on average, little excitation; neurons in the visual cortex are stimulated best by bars or edges of a particular orientation or spatial frequency or specific color or movement information about disparity of images on the two retinas is used to give stereoscopic perception depth and distance. Neurons of the visual cortex with closely related functions are arranged anatomically in columns of cells responding to the same region of the retina, dominated by the same eye..

Center-Surround Modulation

The region of the retina in which an introduction or removal of impulse-like stimuli evokes action potentials is called the classical receptive field of the visual neurons. We examine how stimuli placed outside the CRF modulates the activity the stimulus placed within the CRF. These interactions have been revealed to be of paramount importance in understanding the operation of visual neurons. Both experimental findings and theoretical work have shown results in this direction.

The center-surround modulation occur as a suppression or facilitation based on the properties of the stimuli in the center and/or surround of the receptive field. The properties of the stimuli that might induce suppression or facilitation include the relative orientation, contrast and spatial frequencies.

Contrast Response of Center and Surround

Given independent center and surround mechanism, suppression from the surround will manifest itself in the neuron's contrast response. Suppression might be either divisive or subtractive requiring responses at different stimulus contrast to differentiate the two. Characterizing changes in a neuron's contrast responses will inform us whether the influence from the surround should be modelled as a divisive or

subtractive suppression. To determine which of these three forms of suppression best characterized surround influences, we consider three models. The models are based on the well known Michaelis-Menten equation which well describes the contrast/response relationship in the visual cortical neurons. The first model that explains some behavior of surround suppression through a divisive change in the neuron's response gain.

This response gain model is

$$R = K(C_s) \left(\frac{C_s}{\sqrt{\sigma + C_c^2}} \right)^\beta \quad (2)$$

where R is the neuron's response, $K(C_s)$ is the scaling factor which is a function of the surround contrast, C_c is the center contrast, σ sets the neuron's contrast gain and β sets the slope of the neuron's contrast response function in a log-linear coordinates. The contrast response is scaled by a single factor $K(C_s)$ that depends on surround contrast.

The second model accounts for surround influence with divisive suppression but through a change in the contrast gain.

$$R = K \left(\frac{C_s}{\sqrt{\sigma(C_s) + C_c^2}} \right)^\beta \quad (3)$$

For the contrast gain model, the response scaling factor K is fixed but the contrast gain parameter $\sigma(C_s)$ depends on the surround contrast.

The third model assumes a subtractive influence from the surround

$$R = \max \left[o, K \left(\frac{C_s}{\sqrt{\sigma + C_c^2}} \right)^\beta - K_0(C_s) \right] \quad (4)$$

where $K_0(C_s)$ is an initial response which is a function of surround contrast.

Difference of Gaussian Model

We describe a receptive field model that assume independent center-surround mechanism in which the suppressive takes place through a subtractive mechanism. The difference of Gaussian model (DoG) composed of 2 overlapping mechanism, interacting subtractively.

In this model, first Gaussian (L_c) represent the excitatory contribution of the CRF center and is taken to correspond to the envelop of a Gabor function representing the CRF's spatial structure. The second Gaussian (L_s) centered at the same position,

describes the suppressive contribution of the surround. The response of a neuron to a circular grating of radius x is given by a function of the form :

$$R(x) = K_c L_c - K_s L_s \quad (5)$$

where K_c and K_s are the gains of the center and surround mechanism;

$$L_c = \int_0^x e^{-(y/\sigma_c)^2} dy \quad (6)$$

$$(7)$$

$$L_s = \int_0^x e^{-(y/\sigma_s)^2} dy \quad (8)$$

σ_c and σ_s represent the spatial extent of the center and surround components.

The DoG model assumes that the center and the surround mechanisms interact linearly. It has been argued by many authors that the relationship between neuron's contrast response and the surround stimulus is better explained by a vertical scaling of curve in log-linear coordinates which corresponds to a divisive interaction.

Consequently, Cavanaugh et al introduced an alternative model based on ratio of Gaussian (RoG).

Ratio of Gaussian Model

The receptive field model we present here assumes independent center and surround mechanism in which the surround influences response through a divisive gain control. We focus on how to explain the changes in the receptive field size through this model. The sensitivity of each of the center and surround is modeled with a one-dimensional Gaussian which corresponds to integrating a two-dimensional envelop of the form

$$\exp\left(\frac{-r^2}{2\sigma^2}\right) / r \quad (9)$$

where r is radius and σ is SD of the Gaussian envelop. This envelop determine the spatial extent of the receptive field. The activity of each mechanism is given by integrating the Gaussian envelop covered by a stimulus placed over their common centers. We divide the output of the center by the output of the surround. K_c and K_s control the gains of the center and surround independently. This divisive interaction of center-surround mechanism formed a ratio of Gaussian

$$R(x) = \frac{K_c L_c(x)}{1 + K_s L_s(x)} \quad (10)$$

where

$$L_c(x) = \left(\frac{2}{\sqrt{\pi}} \int_0^x e^{-(y/\sigma_c)^2} dy \right)^2, \quad (11)$$

$$(12)$$

$$L_s(x) = \left(\frac{2}{\sqrt{\pi}} \int_0^x e^{-(y/\sigma_s)^2} dy \right)^2 \quad (13)$$

where x is the stimulus diameter, K_c and K_s are the gains of the center and surround mechanisms, and L_c and L_s are summed squared activities of the center and surround mechanisms respectively. σ_c and σ_s represent the spatial extent of the center and surround respectively, with the constraint $\sigma_s < \sigma_c$.

Changes in receptive field structure with contrast

It has been shown that the spatial extent of the receptive field changes with stimulus contrast. The gain parameter in the model accounted for the changes in the mechanism activity with contrast and thus reflect each mechanism's sensitivity to contrast. At high contrast, the surround is relatively strong and suppresses weak response from the center mechanism. At lower contrasts surround is relatively weak and this suppression is relaxed allowing a higher response from the center. As the stimulus contrast decreases, responses decreases more for small stimulus than for large, resulting in a shift in diameter tuning.

Without suppression, the overall response to large (infinite) stimulus is proportional to $K_c/(1 + K_s)$. Expressing suppression as a fractional reduction in response, we obtain

$$S = 1 - \frac{1}{1 + K_s} \quad (14)$$

where S is the suppression value. But $K_s > 0$, thus we have $0 < S < 1$.

The influence of the surround was, on average, very weak at lower contrasts than at high and this difference in the contrast gain characteristics of the center-surround interaction causes the observed changes in the receptive field size with contrast.

Divisive Normalization Model

This model of cortical processing is based on the statistical characterization of the sensory signals. It is widely assumed that neurons in the sensory area of the brain are adapted, through evolution and development, to the signals to which they are exposed. An appealing hypothesis for the design of sensory signal is the statistically efficient representation of the visual information. The hypothesis states that the

role of early sensory processing is to remove redundancy or increase independence between neural responses to natural stimuli.

One of the methodologies to test this hypothesis, and which we shall follow, is to examine the statistical properties of the environmental signals and show that a transformation derived according to some statistical optimization criterion provides a good description of the response properties of a set of sensory neurons.

We decompose a natural image using some linear filter (such as wavelet basis, gabor, or steerable pyramids), when the higher-order statistical properties of the images are examined, they are found to show some statistical dependencies, even when the basis functions are chosen to optimize independence. The Principal component analysis and Independent component analysis were employed (reference 7 and 9 of simoncelli) to eliminate the statistical dependencies. But this two methods require some extra constraints such as spatial locality and/or symmetry, in order to obtain functions which adequately approximate cortical receptive field.

Considering the joint statistics of neural responses from a typical pair of adjacent basis functions, it is observed that pairs are decorrelated but not statistically independent. The strength of the dependency varies depending on the specific pair chosen. It is strongest for basis functions that are close in spatial position, orientation and scale.

This form of dependencies cannot be eliminated by linear processing. Indeed, a nonlinear form of cortical processing is needed which will ensure that the linear response of each basis function is rectified. By being rectified, we mean that the response of the neurons is squared and then divided by the weighted sum of the rectified responses of neighboring neurons.

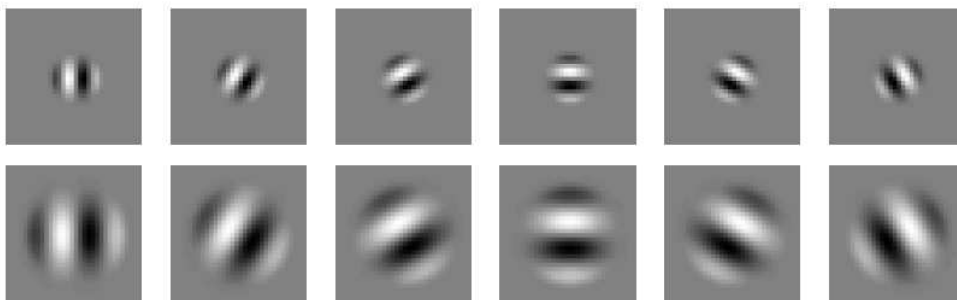


Figure 9: Basis function at two scales and 6 orientations steerable pyramids.

Divisive Normalization Model

The following gives the relationship between the center-surround mechanism through what we describe as divisive normalization.

$$R = P_k^2(x, y) / \left[\sum_{k(i,j)} \omega_{\bar{k}k} P_k^2(x + i, y + j) + \sigma^2 \right] \quad (15)$$

where R is the mean response, $P_k^2(x, y)$ is the center cell, $P_k^2(x + i, y + j)$ is the surround cells (are the values of coefficients at adjacent spatial positions, orientations and scales). $\omega_{\bar{k}k}$ and σ are chosen to minimize squared prediction error:

$$\min_{\omega_{\bar{k}k} \in \mathbb{R}^n, \sigma \in \mathbb{R}} \mathbb{E} \left[P_k^2(x, y) - \sum_{k(i,j)} \omega_{\bar{k}k} P_k^2(x + i, y + j) - \sigma^2 \right]^2 \quad (16)$$

where $\mathbb{E}[\cdot]$ indicates expected value (computed by integrating over the full spatial extent of a set of images).

The explicit values of for the weights $\omega_{\bar{k}k}$ and constant σ can be computed by

- maximum-likelihood (ML) estimation of the parameters, assuming a Gaussian form for the underlying conditional distribution.

$$\mathcal{P}(P_i | P_j, j \in N_i) = \frac{1}{\sqrt{2\pi(\sum_j \omega_{ji} P_j^2 + \sigma^2)}} \exp \left[\frac{-P_i^2}{2(\sum_j \omega_{ji} P_j^2 + \sigma^2)} \right] \quad (17)$$

- Least square estimation

We focus on the Least square estimation approach. Thus, Equation (16) can be re-written in the form:

$$\min_{\omega_{\bar{k}k} \in \mathbb{R}^n, \sigma \in \mathbb{R}} J(\omega_{\bar{k}k}, \sigma), \quad (18)$$

$$J(\omega_{\bar{k}k}, \sigma) = \sum_{x,y} \left[P_k^2(x, y) - \sum_{k(i,j)} \omega_{\bar{k}k} P_k^2(x + i, y + j) - \sigma^2 \right]^2 \quad (19)$$

The subscripts $\bar{k}k$ means that the weights $\omega_{\bar{k}k}$ depends on the relative positions \bar{k} of the center and k surround cells.

Optimization Approach

Here, we estimate the optimal values of the parameters ω and σ by doing a least square estimation. To estimate the minimum values of $\omega_{\bar{k}k}$ and σ , we compute the gradient of J as follows:

$$J'_{\omega_{\bar{k}k}} = -2 \sum_{x,y} \left[\sum_{k(i,j)} P_k^2(x+i, y+j) \right] \cdot \sum_{x,y} \left[P_k^2(x, y) - \sum_{k(i,j)} \omega_{\bar{k}k} P_k^2(x+i, y+j) - \sigma^2 \right] \quad (20)$$

Thus, by Euler equality we have

$$J'_{\omega_{\bar{k}k}} = -2 \sum_{x,y} \left[\sum_{k(i,j)} P_k^2(x+i, y+j) \right] \cdot \sum_{x,y} \left[P_k^2(x, y) - \sum_{k(i,j)} \omega_{\bar{k}k} P_k^2(x+i, y+j) - \sigma^2 \right] = 0 \quad (21)$$

so that we get

$$-2 \sum_{x,y} \left[\sum_{k(i,j)} P_k^2(x+i, y+j) \right] = 0 \quad (22)$$

or

$$\sum_{x,y} \left[P_k^2(x, y) - \sum_{k(i,j)} \omega_{\bar{k}k} P_k^2(x+i, y+j) - \sigma^2 \right] = 0 \quad (23)$$

$$\sum_{x,y} \left[P_k^2(x, y) - \sum_k \omega_{\bar{k}k} P_k^2(x+i, y+j) - \sigma^2 \right] = 0 \quad (24)$$

which implies that

$$\sum_{x,y} P_k^2(x, y) = \sum_{x,y} \sum_{k(i,j)} \omega_{\bar{k}k} P_k^2(x+i, y+j) - \sigma^2 \quad (25)$$

Since

$$\sum_{x,y} \sum_{k(i,j)} P_k^2(x+i, y+j) \neq 0$$

otherwise, we have that there is negative response such that the responses tends to cancel out each other. The optimum is then attained for $\omega_{\bar{k}k}$ and σ such that

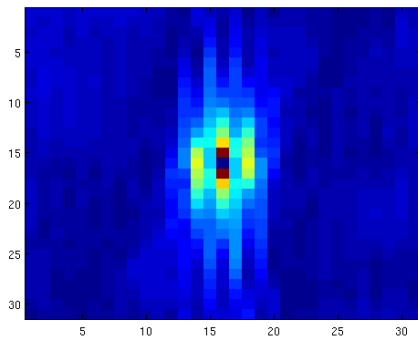
$$\sum_{x,y} P_k^2 = \sum_{x,y} \sum_{k(i,j)} \omega_{\bar{k}k} P_k^2 - \sigma^2 \quad (26)$$

Results

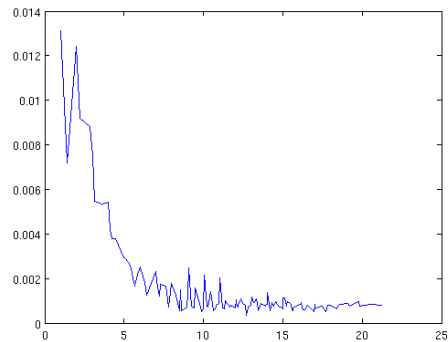
We carried out the optimization procedure for obtaining the minimum weights and we display the weights as greyscale images. Two different optimization algorithms were considered; Least square estimation and Gradient decent method. The two methods produced quite similar results. In both case, the tolerance for changes

in the optimization variable and the function output have to be set in a certain range for convergence to occur. In the Least square estimation approach, this was achieved by rescaling the input and the objective function was modified in the case of gradient decent.

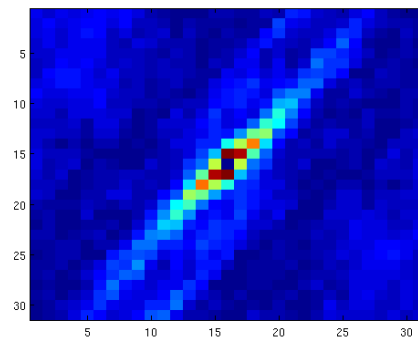
We used three set of images to ensure that the weights are independent of the images used. The images used are shown below. We decompose the images using different filters. The results for Log-Gabor filter, Steerable pyramid, Wavelet pyramid and Laplacian pyramid, are show below, for different subbands. The color corresponds to the weights observed with hotter color having higher weights



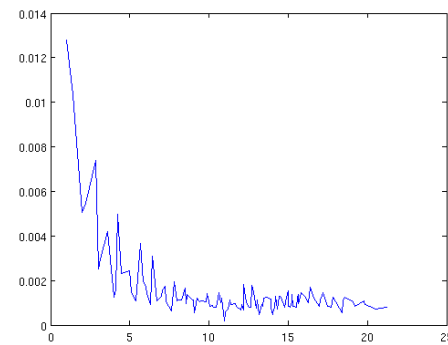
(a) Steerable Pyramid subband 6.



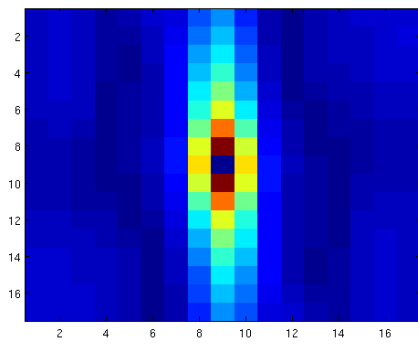
(b) Average weights against separation



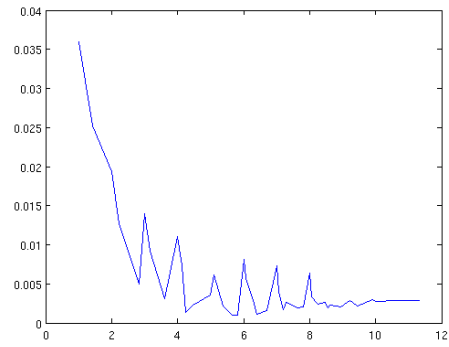
(c) Wavelet Pyramid subbands 6.



(d) Average weights against separation



(e) Log-Gabor subband 6



(f) Average weights against separation

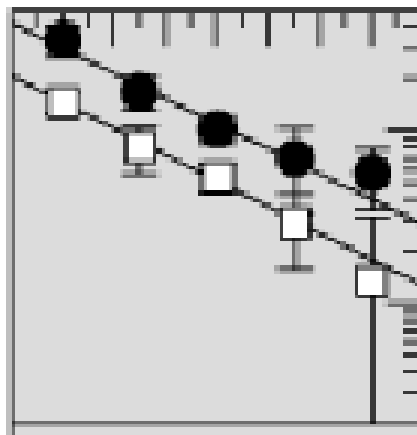


Figure 10: Center-Surround Separation

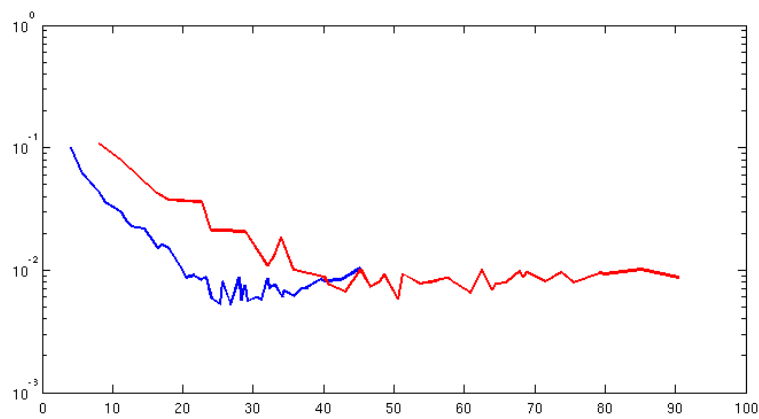


Figure 11: Center-Surround Separation

Discussion

In the case of steerable pyramid, we observe that the weights in the neighborhood of the center neuron are higher than those far away. This is observed for other filter

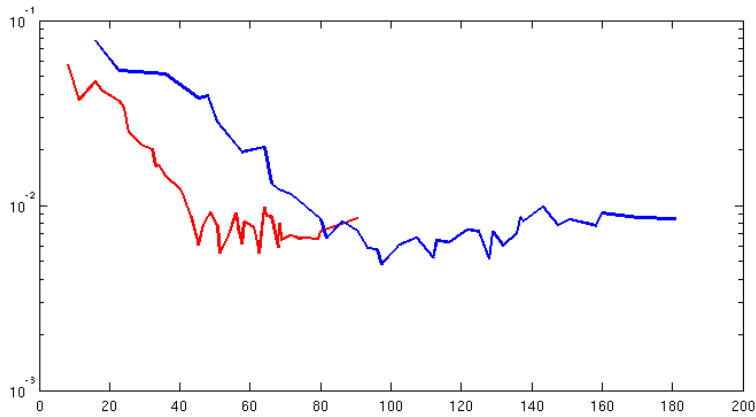


Figure 12: Center-Surround Separation

banks as well. Thus, the area of the proximity of the center is the region where the suppression is strongest. Similarly, in the case of Gabor-pyramid, we observe a form of symmetry in the weights. Thus, the suppressive effect of the surround cell which are about the same separation from the center neuron is the same.

In Petrov and Mckee (2006) (see reference) experiment on on center-surround separation, when the separation between the center and surround radius was kept at a constant value, it was found that the suppression was strongest for any masking. They suggested that the drop off in suppression overlap (see figure ??) for high spatial frequency and low spatial frequency of the surround stimulus. This idea contradicts what we have observed in our simulation. We have seen earlier in the results that the area of the surround in the proximity of the center is the region where there suppression is strongest. Our results, see figure 11 shows that the weights are more diffused for low spatial frequency than for high spatial frequency. This indicates that the drop off in suppression is faster for low spatial frequency than for high spatial frequency.

Conclusion

We have shown that the response of the center cell to a stimulus presented outside the classical receptive field is influenced by the surround contrast, orientation and spatial frequency. We obtained the set of weights that minimizes the squared prediction error for different subband and displayed these values as grayscale images. We have also established that the separation varies with the spatial frequency. What might be interesting to consider in a later work is to use a multiplicative model and try to investigate if the center-surround interactions preserves the kind of property we have established.

References

- E P Simoncelli & O Schwartz. Modeling surround suppression in V1 neurons with a statistically-derived normalization model. In *Adv. Neural Information Processing System*, volume 11, pages 153-159, Cambridge, MA, 1999. MIT Press. Presented at NIP-98, Denver Co, 1-3 Dec 1998.
- Schwartz, O., & Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control. *Nature Neuroscience*, 4, 819Y825. [PubMed] [Article]
- Yury Petrov & Suzanne P Mckee. The effects of spatial configuration on surround suppression of contrast sensitivity. *Journal of Vision* (2006) 6, 224-238 <http://www.journalofvision.org/6/3/4/>
- M. J. Wainwright, O Schwartz, & E. P. Simoncelli, Natural image statistics and divisive normalization: Modeling nonlinearities and Adaptation in cortical neurons“ , *Statistical Theory of Brain* , MIT Press, 2001
- Peter Kovesi Matlab and Octave functions for computer vision and image processing. <http://www.csse.uwa.edu.au/~pk/Research/Matlab>
- X. Zhang, J Park, J Salant, S Thomas, J Hirsch, D. Hood A multilicative model for spatial interaction in the human visual cortex *Journal of Vision* (2008) 8(8):4, 1-9 <http://www.journalofvision.org/8/8/4/>