

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

Xueping Yao
Universite de Nice Sophia Antipolis

23 July 2009

Preface

Nowadays, more and more attention are being paid on the exploration of the human vision research from mathematical and computational perspectives. The goal is to unveil the principles that govern the functioning of neurons and assemblies thereof and to use the results to bridge the gap between biological and computational vision. Depending on the scale of analysis, a variety of mathematical frameworks to model the brain have been taken into consideration.

The first goal of this paper is to establish a review of recent literature relevant to center-surround interactions within early cortical areas (with the focus to begin on V1 and LGN). Giving a coherent description of the stimulus factors that may influence suppression and enhancement prepares us the necessary biological background for mathematical modeling. The second goal is to present the result of implementing a quantitative model towards explaining the nature and dynamics of center-surround interactions, which provides an evidence that contradicts with Petrov and McKee (2006)'s conclusion in [1].

Chapter 1 firstly introduces fundamental concepts and mechanisms in human vision research, including the pathway of visualizing and perceiving natural images in visual system in brain. Secondly, it explains what is center-surround interaction and related factors which may account for such phenomenon.

Chapter 2 investigates several primary mathematical models being studied in vision research. Each of them has both improved performance and drawback in fitting data collected by neurophysiological and psychological researches comparing to other models. This chapter also gives a brief introduction of some popular processing tools in image research.

Chapter 3 gives a brief introduction of some popular processing tools in image research, including Gabor filter, steerable pyramid and wavelet pyramid. It summarizes both the advantage and the drawback of each method.

Chapter 4 is concerned with the implementation of a quantitative model proposed in a specific paper. It describes the numerical optimization and its numerical result. It also discusses the meaning what the result indicates and the relationship between the results in this paper and from other researches.

Chapter 5 focus on continuing some concepts firstly referred in the Chapter 1 with more details to give a more comprehensive picture about the center-surround interaction phenomenon. In addition, this chapter provides some hints about where further interest in this topic may lie in.

Contents

Abstract	1
1 Introduction	3
1.1 An overall introduction of v1	3
1.2 Center-surround interaction	5
2 Mathematical models	7
2.1 The normalization model	7
2.2 The multiplicative model	8
2.3 Other models	8
3 Image decomposition	11
3.1 Gabor filter	11
3.2 The steerable pyramid	12
3.3 The wavelet pyramid	13
4 Implementation results	14
4.1 Numerical optimization	16
4.2 Results analysis	18
5 Discussion	20
References	21

Chapter 1

Introduction

1.1 An overall introduction of v1

The image captured by each eye is converted into nerve impulses by retina and transmitted to the brain by the optic nerve. This nerve terminates on the cells of the lateral geniculate nucleus (LGN), the first relay in the brain's visual pathways. The cells of LGN then project to their main target, the primary visual cortex (V1). It is in the primary visual cortex that the brain begins to reconstitute the image from the receptive fields of the cells of the retina. Fig.1.1 and 1.2 illustrate both the structure and the pathway. A large portion of V1 is mapped to the Fovea, a spot located in the center of the retina responsible for sharp vision, this is known as cortical-magnification. V1 neurons have strong tuning to a small set of stimuli, the neuronal responses can discriminate small changes in visual orientations, spatial frequencies and colors, furthermore, these neurons have Ocular Dominance, tuning to one of the two eyes, and they tend to cluster together as cortical columns.

Another important notation is receptive field (RF) of each cell in the visual cortex. It is a discrete area in space relative to the fovea where the presentation or removal of a visual stimulus will cause cellular activation. Cells tend to respond to simple patterns (such as oriented bars) and code for a specific region of visual space. By definition, stimuli presented outside of this receptive field will neither increase nor decrease the ongoing activity of that individual cell. The location and size of RF could be also discussed based on center-surround interaction in V1.

There are three types of cells or neurons in the Primary Visual Cortex (V1):

1. Simple Cells: Respond to bars of light, excited to a specific line of a particular orientation placed in the center of its receptive field, and stops firing if moved away from its center.
2. Complex Cells: Respond to line orientation in or out of its excitatory/ inhibitory zones, and particularly so to movement.
3. Hyper Complex Cells: Respond to moving corners or angles.

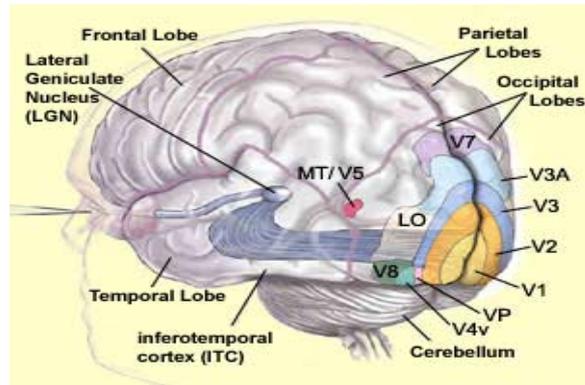


Figure 1.1: The structure of the brain

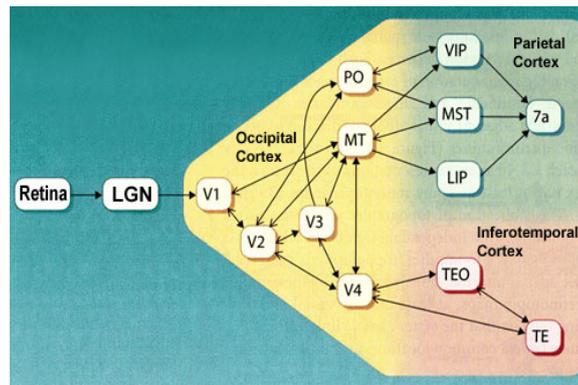


Figure 1.2: The relationship between different parts in the brain

V1 consists of tiled sets of selective spatio-temporal filters, together they can carry out neuronal processing of spatial frequency, orientation, motion, direction and speed (thus temporal frequency). The visual information relayed to V1 is not coded in terms of spatial (or optical) imagery, but rather as local contrast. For example, an image with one half black and one half white, the separating line has the strongest local contrast and is encoded, while few neurons code the brightness information (black or white). As information is sent further to other visual areas, it is coded as increasingly non-local frequency/phase signals, at these early stages of visual processing, spatial location of visual information is well preserved amid the local contrast encoding.

V1 is the first site where strong orientation and direction selectivities are observed in the macaque monkey (Hubel and Wiesel, 1968) in [2]. While the vast majority of V1 cells show some degree of orientation selectivity, only approximately 25-35% of V1 cells are strongly directionally selective (Schiller et al., 1976 [3]; DeValois et al., 1982 [4]). The classic method for testing orientation and direction selectivity is to measure the spike rate of a single cell in response to drifting oriented luminance bars and/or drifting luminance spots as Fig.1.3. The orientations of center and surround stimuli are orthogonal in the left sub-image. And the center and surround stimuli are of the same orientation in the right

sub-image.

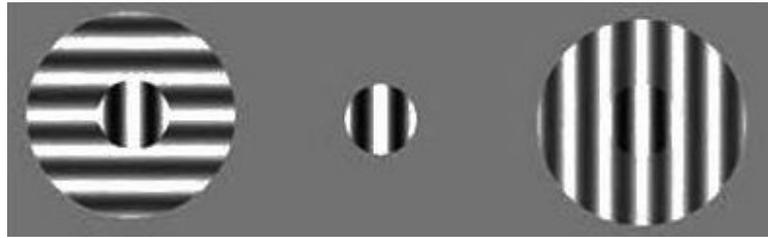


Figure 1.3: The drifting oriented luminance spots.

1.2 Center-surround interaction

The spiking response of a primary visual cortical cell to a stimulus placed within its receptive field can be up- and down-regulated by the simultaneous presentation of objects or scenes in the "silent" regions which surround the receptive field. Such phenomenon in perceiving is called center-surround interaction. For some stimulus conditions, the apparent contrast is suppressed and for other conditions the apparent contrast is enhanced. Some 40 years ago, Hubel and Wiesel (1965) [5] noticed that neurons in Visual areas 18 and 19 of cats responded much more weakly if the otherwise optimal stimulus was extended beyond the neuron's receptive field. Since then, this type of inhibition termed surround suppression has been commonly observed in early visual areas of both cats and primates. In recent psychophysical study, (Petrov, Carandini and McKee, 2005) found the same properties of the surround suppression in human observers [6]. However, the explanations suggested so far (e.g., Schwartz and Simoncelli, 2001) [7] tend to focus on one aspect of the suppression (usually, its orientation tuning), while ignoring others. In addition, many relevant properties have not been studied.

Among all relevant factors, spatial aspects are of particular interest because the proposed models make specific assumptions about how the location of the surround mask affects suppression. Contrast detection and contrast matching are two primary kinds of experiments to collect data. For instance, in contrast matching tasks, the contrast dependence of center-surround interactions was measured by systematically varying the suprathreshold contrasts of the central and surround gratings (Xing and Heeger, 2001). The effects of surround on the perceived contrast of the target were, overall, in agreement with the cat and monkey neurophysiological data: (a) the effect of the surround was to suppress perceived contrast, (b) the suppression was strongest when the target and the surround carriers had the same orientation, and (c) the strength of the suppression did not change significantly between collinear and flanking surround layouts (Cannon and Fullenkamp, 1991 [8]; Ejima and Takahashi, 1985; Xing and Heeger, 2001).

The observed phenomenon that the center-surround interaction could be either inhibitory or facilitatory are addressed by many researches. The effect of surrounds on contrast detection thresholds varies from facilitation (Polat and Sagi, 1993 [9], 1994 [10]; Yu, Klein,

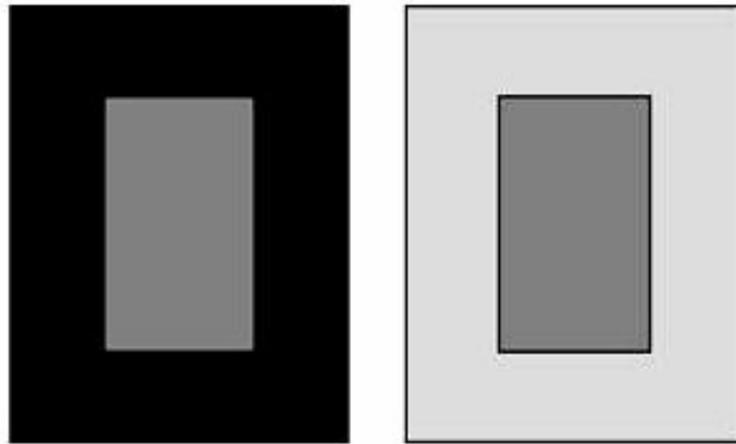


Figure 1.4: The center-surround interaction, the right center seems to be "stronger" than the left center, while actually they are of same strength. The apparent contrast of center stimulus is affected by its surround.

and Levi, 2002 [11]; Zenger-Landolt and Koch, 2001 [12]) to suppression (Solomon and Morgan, 2000 [13]; Williams and Hess, 1998 [14]), depending on the surround area, orientation, and phase with respect to the target, as well as the stimulus eccentricity (Andriessen and Bouma, 1976 [15]; Petrov et al., 2005 [16]; Snowden and Hammett, 1998 [17]; Zenger-Landolt and Koch, 2001). Recently, Petrov, Verghese, and McKee (2006) showed that the facilitation observed in contrast detection thresholds is primarily due to a reduction in uncertainty about target location. Mahmoodi and Young (2005) suggested that the center-surround interaction in area V1 is sensitive to the higher-order structures of natural scene images, such as image contours[18].

Chapter 2

Mathematical models

The responses of cells in the primary visual cortex are influenced in interesting and complex ways by the presence of patterned stimuli appearing in the visual space surrounding the receptive field of a given cell. And there exist a variety of different models that consider the nature of these interactions. When a neighbor stimulus is present, mutual inhibition of the responses to both target and neighbor stimuli is often observable. Two kinds of models, a simple multiplicative model and the divisive normalization model, have been used to describe the spatial interaction in the visual cortex.

2.1 The normalization model

The normalization model was originally proposed to describe the interaction among cortical channels, such as orientation and spatial frequency channels. Note that interactions among cortical channels differ from the spatial interaction: the visual stimuli used for studying interactions among cortical channels often overlap each other, while the stimuli used for studying spatial interaction should be separated. Nonetheless, the success of the normalization model indicates that features of the normalization model, such as mutual inhibition and response normalization, are important in describing spatial interaction.

The normalization model assumes that the responses to multiple stimuli are pooled to generate a divisive inhibition. Although it appears in various forms in different papers, the normalization model can be expressed as (Xing and Heeger, 2001) in [19]

$$R = \frac{R_{max} C_t^\alpha}{\sigma^\beta + C_t^\beta + k C_n^\beta} \quad (1)$$

where C_n is the contrast of a neighbor stimulus, k is factor that determines the strength of the inhibitory effect, R_{max} is the asymptotic amplitude of the response, C_t is the contrast of the stimulus, α and β are related to the excitatory and inhibitory modulation respectively, σ is the semi-saturation contrast.

The normalization model has been shown to be fairly consistent with a wide range of single cell recordings and psychophysical data. It captures both the nonlinearity of the visual

system and the reduction in response amplitude by the neighbor stimulus. In addition, it has been shown, with reducing information redundancy, to allow the visual system to code nature images more efficiently. However, the normalization model does not appropriately describe the responses when the neighbor stimulus has a high contrast and the target stimulus has a low contrast. For instance, Ejima and Takahashi (1985) reported that the inhibitory effect of the neighbor stimulus approaches an asymptotic level when the contrast of the neighbor stimulus is much higher than the contrast of the target contrast in [20].

2.2 The multiplicative model

To overcome the primary drawback of normalization model, a multiplicative spatial interaction model, which fits data better when the neighbor contrast is high, is proposed in [21].

$$R = \frac{AC_t^\alpha}{\sigma^\beta + C_t^\beta} \left(1 + \frac{B}{1 + \left(\frac{qC_n}{C_t}\right)^\gamma}\right) \quad (2)$$

where, R is the amplitude of the response, B is the factor describing the strength of the spatial interaction, γ is a power term that describes nonlinearity of the spatial interaction, q is a factor that describes the effective contrast of the neighbor stimulus. When C_n is zero, $A(1 + B)$ is just the R_{max} in normalization model (1).

The spacial interaction is described as a multiplicative process here. Firstly, the spatial interaction term and the physical contrast of the target stimulus are separate terms that are multiplied together to determine the amplitude of the target response. Secondly, the spatial interaction mechanism is nonlinear. the γ term here is larger than 1. Therefore, when C_n/C_t deviates slightly from $(C_n/C_t)^\gamma$, and the spacial interaction term will change dramatically. This reflects the mutual inhibition between target and neighbor, where the stimulus with the slightly larger contrast exhibits a much stronger influence than predicted by the difference in contrasts of the two stimuli. Consequently, the difference between target contrast and the neighbor stimuli is amplified. Thirdly, the multiplicative model emphasizes the saturation of the spatial interaction when two stimuli have very difference contrasts. Therefore, a weak target stimulus among strong neighbor stimuli can remain visible because the spatial inhibition from the neighbor response is limited.

2.3 Other models

Besides of above two primary categories of mathematical models, a divisive model ($R \circ G$) is proposed by assuming independent center and surround mechanisms in which the surround influences responses through a divisive gain control. It is based on the ratio of two Gaussian sensitivity distributions to study responses to circular patches of grating[22]:

$$R(x) = \frac{K_c L_c(x)}{1 + k_s L_s(x)} \quad (3)$$

where

$$L_c(x) = \left(\frac{2}{\sqrt{\pi}} \int_0^x e^{-(y/w_c)^2} dy \right)^2 \quad (4)$$

and

$$L_s(x) = \left(\frac{2}{\sqrt{\pi}} \int_0^x e^{-(y/w_s)^2} dy \right)^2 \quad (5)$$

This model is intended to provide a simple explanation of changes in receptive field size by using mechanisms with spatially constant dimensions. The sensitivity distribution of each mechanism is model with a one-dimensional Gaussian envelope. It is important to understand that the Gaussian envelopes do not describe the spatial weighing function of the receptive field but only the envelope of that function. So for a linear approximation to a simple cell, the center envelope would correspond roughly to the Gaussian envelope of a Gabor filter. x is the stimulus diameter, k_c and k_s are the gains of the center and surround mechanisms, L_c and L_s are the summed squared activities of the center and surround mechanisms. The spatial extents of the center and surround components are represented by w_c and w_s , $w_c < w_s$.

Another kind of model ($D \circ G$) is based on the difference of two Gaussians that describes the mechanisms of the center and the surround respectively .

The response of simple cell is modeled as a convolution of a Gabor on an input image $I(x, y)$:

$$R(x, y, \theta, F, \sigma, \phi) = I(x, y) * g(x, y, \theta, F, \sigma, \phi) \quad (6)$$

$$g(x, y, \theta, F, \sigma, \phi) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x'^2 + \lambda y'^2}{2\sigma^2}\right) \cos(2\pi F_0 x' + \phi) \quad (7)$$

where $x' = x \cos(\theta) + y \sin(\theta)$, $y' = -x \sin(\theta) + y \cos(\theta)$.

And the response of complex cell is modeled as the responses of a pair of simple cells which are with $\frac{\pi}{2}$ phase difference:

$$E_{F,\sigma,\theta}(x, y) = \sqrt{R(x, y, \theta, F, \sigma, \phi)^2 + R(x, y, \theta, F, \sigma, -\frac{\phi}{2})^2} \quad (8)$$

where $\theta_i = \frac{i\pi}{8}$, ($i = 0, \dots, 7$) is the orientation of the normal to the parallel stripes of the Gabor function, and the optimal response could be expressed as:

$$\hat{E}_{F,\sigma} = \max\{E_{F,\sigma,\theta}(x, y) | (i = 0, \dots, 7)\} \quad (9)$$

Define $A(x, y) = \theta_k$ and $k = \arg \max\{E_{F,\sigma,\theta}(x, y) | (i = 0, \dots, 7)\}$

The impact of nonclassical inhibitory field is modeled by a $D \circ G$ function:

$$D \circ G_{\sigma,\theta}(x, y) = \left[\frac{1}{2\pi(4\sigma)^2} \exp\left(-\frac{x'^2 + \lambda y'^2}{2(4\sigma)^2}\right) - \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x'^2 + \lambda y'^2}{2\sigma^2}\right) \right]^+ \quad (10)$$

And the weighting coefficient of position is as follows:

$$W_{\sigma,\theta}^+(x, y) = \frac{D \circ G_{\sigma,\theta}(x, y) \cdot V(x' - \sigma)}{\iint D \circ G_{\sigma,\theta}(x, y) \cdot V(x' - \sigma) dx dy} \quad (11)$$

$$W_{\sigma,\theta}^-(x, y) = \frac{D \circ G_{\sigma,\theta}(x, y) \cdot V(-x' - \sigma)}{\iint D \circ G_{\sigma,\theta}(x, y) \cdot V(-x' - \sigma) dx dy} \quad (12)$$

$$V(\xi) = \begin{cases} 1, & \text{for } x \geq 0 \\ 0, & \text{for } x < 0 \end{cases} \quad (2.1)$$

$$(2.2)$$

The weighting coefficient of phase difference is as follows, where σ_Δ is the Gaussian variance.

$$A'_{\theta,\sigma_\Delta}(x, y) = \exp\left(-\frac{\min(|A(x, y) - \theta|, \pi - |A(x, y) - \theta|)^2}{2\sigma_\Delta^2}\right)$$

Finally, the inhibitory impact is modeled as:

$$H_\theta(x, y) = \min\{[\hat{E}_{F,\sigma} * W_{\sigma,\theta}^+ * A'_{\theta,\sigma_\Delta}](x, y), [\hat{E}_{F,\sigma} * W_{\sigma,\theta}^- * A'_{\theta,\sigma_\Delta}](x, y)\}$$

with α modulates the inhibitory strength, the final Gabor energy response is :

$$C_\sigma(x, y) = [\hat{E}(x, y) - \alpha H((x, y))]^+$$

This $(D \circ G)$ model is according to the biological mechanism of non-classical receptive field inhibition in the visual cortex. With Gabor energy as a response of receptive field, a non-classical receptive field of region is modeled by two ellipsoid semi-circle with direction based on the difference of two Gaussian functions. The weight function has been designed by difference of phase between center receptive field and the surrounding inhibitive region, to stimulate the non-classical receptive field inhibition mechanism. Modeling the non-classical receptive field as two ellipsoid semi-circle is the important aspect of this quantitative model.

Chapter 3

Image decomposition

The visual stimuli used for studying interactions among cortical channels often overlap each other, while the stimuli used for studying spatial interaction should be separated. Thus, it is necessary to apply image decomposition to take a close look at details at different channels.

Multi-scale linear transforms such as wavelets have become popular for image representation. Typically, 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. In addition, a number of authors have noted that wavelet coefficients have significantly non-Gaussian marginal statistics. Because of these properties, it is believed that wavelet bases provide a close approximation to the independent components decomposition for natural images. There are several widely used image decomposition methods, including the Gabor filter, the wavelet pyramid (separable) and the steerable pyramid.

3.1 Gabor filter

The Gabor filter is a linear filter whose impulse response is defined by a harmonic function multiplied by a Gaussian function as follows:

$$g(x, y, \lambda, \theta, \phi, \sigma, \nu) = \exp\left(-\frac{x'^2 + \nu y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \phi\right)$$

where $x' = x \cos(\theta) + y \sin(\theta)$, $y' = -x \sin(\theta) + y \cos(\theta)$. λ represents the wavelength of the cosine factor, θ represents the orientation of the normal to the parallel stripes of a Gabor function, ϕ is the phase offset, σ is the sigma of the Gaussian envelope and ν is the spatial aspect ratio, and specifies the ellipticity of the support of the Gabor function.

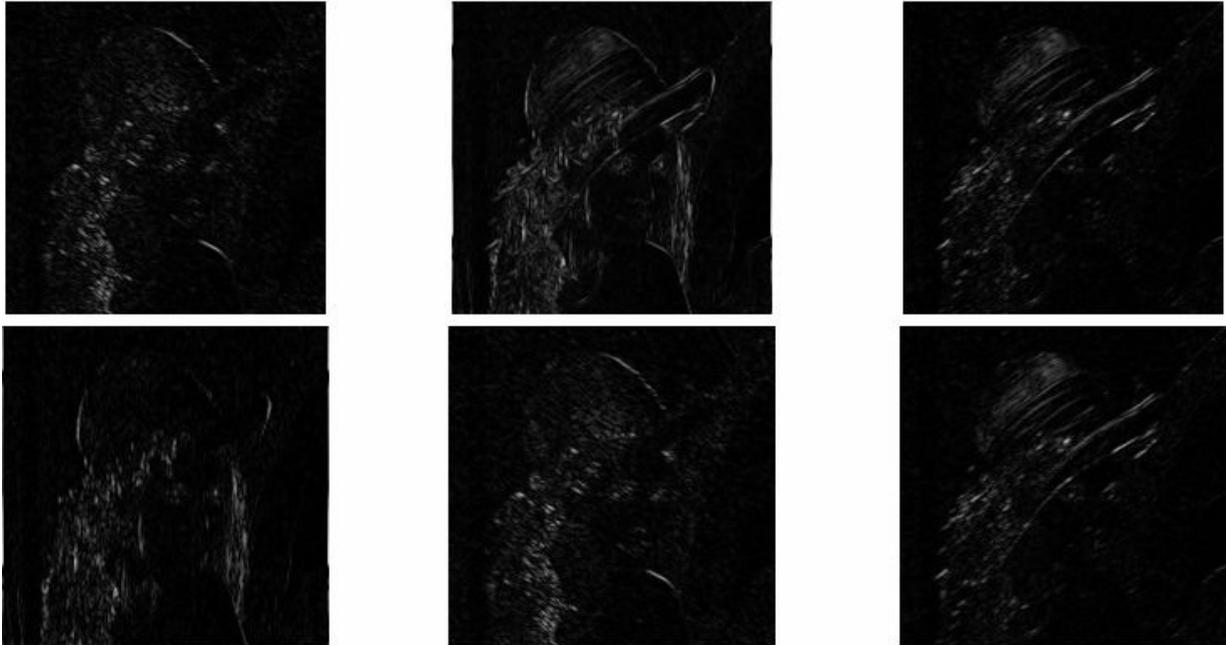


Figure 3.1: Gabor filter, Wavelength=8, orientation=0, phase offset=90, aspect ratio=0.5, bandwidth=1, number of orientation=6

Applying Gabor filter to the test image "Lena" to get Fig. 3.1, the result illustrates the idea that one can view the details of an image through extracting each subband at different directions. The Gabor filter is popular in image processing since its orientated kernel models the orientational inclination of human vision well.

3.2 The steerable pyramid

The steerable pyramid is a linear multi-scale, multi-orientation image decomposition that provides a useful front-end for image-processing and computer vision applications. It overcomes the limitations of orthogonal separable wavelet decompositions that those representations are heavily aliased, and do not represent oblique orientations well. It performs a polar-separable decomposition in the frequency domain, thus allowing independent representation of scale and orientation. Since it is a tight frame (self-inverting), it obeys the generalized form of Parseval's Equality: The vector-length (L2-norm) of the coefficients equals that of the original signal. More importantly, the representation is translation-invariant and rotation-invariant. This can make a big difference in applications that involve representation of position or orientation of image structure. The primary drawback is that the representation is overcomplete by a factor of $4k/3$, where k is the number of orientation bands.

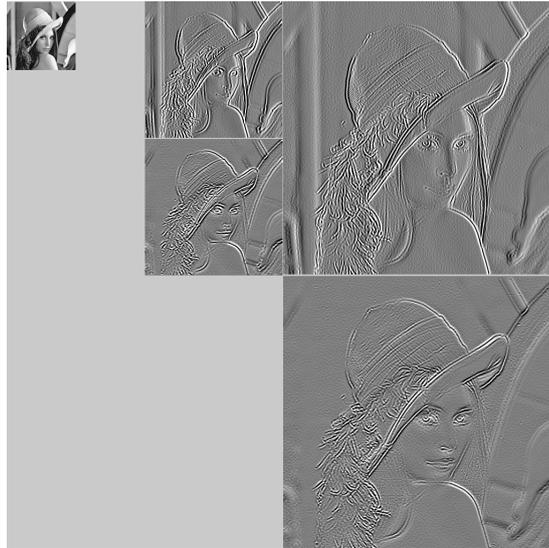


Figure 3.2: Steerable pyramid decomposition

The basis functions of the steerable pyramid are directional derivative operators, that come in different sizes and orientations. An example decomposition of the same image "Lena" which is used in the explanation of Gabor filter is shown above. This particular steerable pyramid contains 2 orientation subbands, at 2 scales. The number of orientations may be adjusted by changing the derivative order (for example, first derivatives yield two orientations).

3.3 The wavelet pyramid

The wavelet pyramid will be explained in more details in the later chapter with some example. An overall picture of comparing the difference between each of them could be viewed:

1. Gabor filter: oriented kernels, not inverse jointly-localized, not translation-invariant, not rotation-invariant, not self-inverting
2. Wavelet decomposition: no oriented kernels (not diagonals), jointly-localized, not translation-invariant, not rotation-invariant, self-inverting
3. Steerable pyramid: oriented kernels, jointly-localized, approximately translation-invariant, approximately rotation-invariant, approximately self-inverting

Chapter 4

Implementation results

When reading through relevant literatures of the centre-surround interaction, we found an interesting question: It is proposed that surround suppression should be locally anisotropic (Schwartz and Simoncelli, 2001). In the contrast, Petrov and McKee (2006) gave a conclusion that surround suppression is locally isotropic based on psychological experiments. Since the first proposal is not consistent with the latter one, we implement the model adopted by Schwartz and Simoncelli (2001) to test which conclusion is more reasonable. Besides, centre-surround interaction in the primary visual cortex (area V1) has been studied extensively using artificial, abstract stimulus patterns, such as bars, gratings and simple texture patterns. In this report, we extend the study of centre-surround interaction by using natural scene images and the divisive normalization model proposed by Schwartz and Simoncelli (1999) in [23]:

$$R = C^2 / [\sum_k \omega_k P_k^2 + \sigma^2]$$

The parameters $\{\omega_k\}$ and σ are chosen to minimize squared prediction error through extracting the coefficient of corresponding subband in the wavelet pyramid decomposition:

$$\{\hat{\omega}, \hat{\sigma}\} = \arg \min \mathbf{E}[C^2 - \sum \omega_k P_k^2 - \sigma^2]^2 \quad ,$$

Where C is the value of center coefficient, P_k are the values of coefficients at adjacent spatial positions, orientations and scales, and $\mathbf{E}[\cdot]$ indicates expected mean value.

Using a multi-scale wavelet basis to decompose natural test images, we examine their statistics from decomposition coefficients. Although the coefficients of this representation are nearly decorrelated, they exhibit important higher-order statistical dependencies that cannot be eliminated with purely linear processing. In particular, rectified coefficients corresponding to basis functions at neighboring spatial positions, orientations and scales are highly correlated. A method of removing these dependencies is to divide each coefficient by a weighted combination of its rectified neighbors to have a divisive normalization model as above.

To study the model, we adopt the Wavelet decomposition pyramid to have a weighted combination of squared coefficients at two scales, all three orientations. The weights

parameters $\{\omega_k\}$, representing the interaction between center and surround neurons, of the final normalization signal are optimized for the statistics of a set of three 512×512 images as follow:



Figure 4.1: Boat



Figure 4.2: Lena



Figure 4.3: Goldhill

For instance, applying the wavelet pyramid to the second image "Lena", we get the coefficients at different scales and orientations as Fig. 4.4.

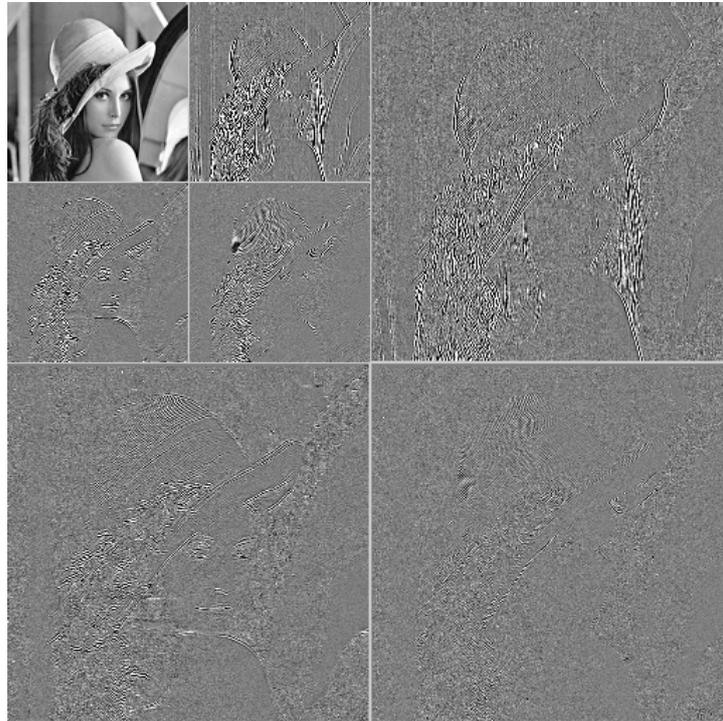


Figure 4.4: Wavelet pyramid

4.1 Numerical optimization

The primary steps to analyze the optimized weights of the details from 2nd recursive level at vertical direction could be summarized as:

1. Extract the corresponding coefficients matrix which is 128×128
2. Choose a 13×13 window with the center pixel modeling the center neuron
3. Link $\{\omega_k\}$ to each pixel in this window
4. Move the window over the whole extracted area to establish the objective function
5. Apply optimization algorithm to find the weights $\{\omega_k\}$ by minimizing the objective function

The optimized weights $\{\omega_k\}$ of each subband could be visually represented as follows. Fig. 4.5 and Fig. 4.6 represent the value of optimized weights corresponding to horizontal details and diagonal details respectively. For each image, one can observe that the weights are symmetric with respect to the center, which verifies that surround neurons whose position are symmetric on one direction will have the same impact on the center neuron. Besides, it also suggests that the center-surround interaction declines as the distance between the center neuron and the surrounding neuron increases.

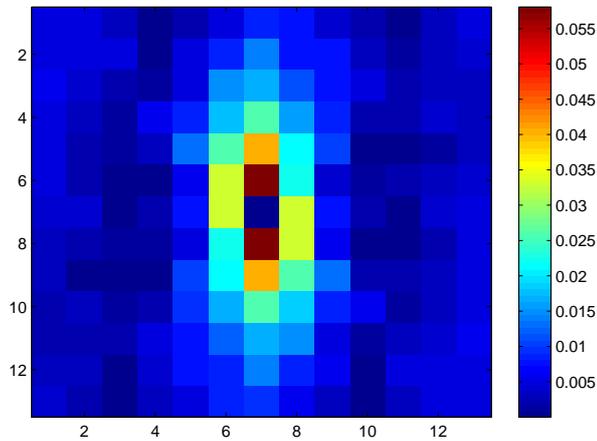


Figure 4.5: Optimized weights, 2nd level, vertical

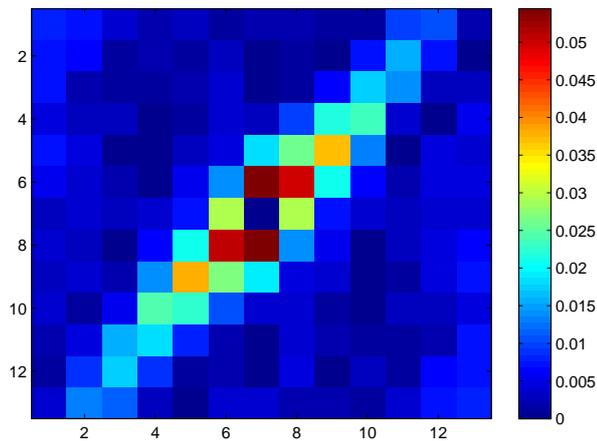


Figure 4.6: Optimized weights, 2nd level, diagonal

4.2 Results analysis

Petrov and McKee (2006) suggested in [1] that surround suppression should be locally isotropic. In their experiment, the five different layouts illustrated in Fig.4.7b-f were used to test for both surround collinearity with the target and surround symmetry (unilateral vs. bilateral). To test for collinearity effects, suppression from the bow-tie mask collinear with the target, C (Fig.4.7e), was compared with suppression from the bow-tie mask flanking the target, F (Fig.4.7f). Also, suppression from the half-annulus mask shown near the end of the target, E (Fig.4.7c), was compared with that from the half-annulus mask shown near the side of the target, S (Fig.4.7d). To test the effects of symmetry, suppression from the two unilateral masks, E and S, was compared with that from the two bilateral masks, C and F. The results for four subjects are shown in Fig.4.8. They concluded that the suppression was unaffected by either the surround collinearity (C vs. F and E vs. S) or its symmetry (E and S vs. C and F) based on the final average data.

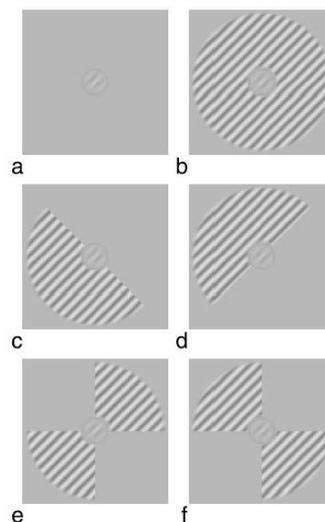


Figure 4.7: (a) No surround condition. A cosine phase Gabor target with 45 deg orientation and a thin localizer circle is shown. (b) The Gabor target surrounded by the full-annulus mask. The same-phase surround condition is shown. (c) The half-annulus mask positioned at the end of the Gabor target. (d) The half-annulus mask positioned at the side of the target. (e) The bow-tie mask collinear with the target. (f) The bow-tie mask flanking the target.

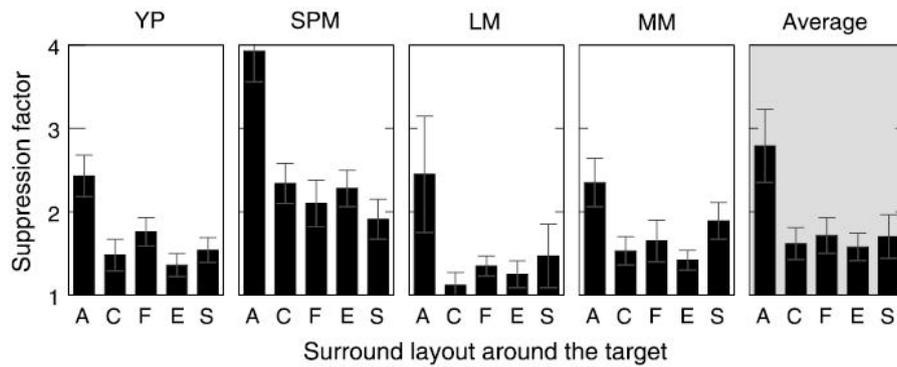


Figure 4.8: A:full annulus; C:collinear bow-tie mask; F:flanking bow-tie mask; E:end half-annulus mask; S:side half-annulus mask

On the contrary, from the optimal weights values of Fig. 4.5 and Fig. 4.6 being displayed above, we could observe an orientated and symmetric characteristic, which indicates that nearby neurons in different positions are in charge of responding to stimuli in different orientations in the center-surround interaction. It also implies that surround suppression acts as a fine-tuned divisive normalization, with larger weights assigned to surround signals having the same orientation as the center. It illustrates that the suppression has the orientational selectiveness along every direction of the channel.

These quantitative findings verifies the conclusion that surround suppression is locally anisotropic based on their mathematical model given by Schwartz and Simoncelli (2001). The conclusion has only been qualitatively suggested before. Our results quantitatively give a support to it. Meanwhile, it raises another interesting problem. Concerning to the question that whether the inhibitory impact is locally anisotropic or isotropic, both qualitative and quantitative analysis based on the divisive normalization model indicate the anisotropy, while the records of human subject tests support isotropicity. Mathematically, from which aspect should the model be modified to fit data better? Biologically, what factors related to center-surround interactions have been ignored or underestimated during modeling?

Chapter 5

Discussion

The model we have studied and implemented captures many important characters of center-surround interaction in early visual cortex. And it fits and explains sets of data collected by biological experiments and psychophysical studies well. However, This simplified model also contradicts with some conclusions drawn from other experiments. To have a conclusion with more credibility, we may take into consideration of more aspects, such as the size of receptive field, stimulus contrast, recent stimulation history, phase difference, etc..

For instance, the single-cell search for feature-trigger specificity has fostered a view of visual neurons as static and localized windows of the visual world historically, functioning independently of one another. However, The observation of center-surround modulations shows that this assumption is inaccurate. Classical receptive field of neurons in V1, their size and functional selectivities, are found to be dynamically altered by the spatial and temporal context of the visual stimulation. When multiple objects or natural scenes are shown, they interact non-linearly over extended cortical regions and periods of time. Another aspect is the influence of stimulus contrast. A large part of investigations implicitly or explicitly considered the extent of cortical spatial summation and, therefore, the size of the classical receptive field to be fixed and independent of stimulus characteristics or surrounding context. On the contrary, some investigators found that the extent of spatial summation in V1 neurons depended on contrast, and was on average 2.3-fold greater at low contrast. This adaptive increase in spatial summation at low contrast was seen in cells throughout V1 and was independent of surround inhibition.

The precise relationship between the response of a cell and its surrounding pattern is a phenomenon that merits further investigation. Moreover, new psychophysical and neurophysiological results have emerged in this domain in the past several years providing new observations regarding the manner in which these interactions vary over space and time. This allows the possibility of attaining a clearer picture of the nature and dynamics of these interactions.

Acknowledgements

I would like to thank Dr. Pierre Kornprobst and Dr. Neil Bruce for the great help and support during my internship at INRIA Sophia Antipolis - Mditerrane.

Thanks are also due to Dr. Chiara Simeoni and Prof. Victorita Dolean for their organizational help.

References

- [1] Y. Petrov and S. P. McKee. The effect of spatial configuration on surround suppression of contrast sensitivity. *Journal of Vision*, 6:224–238, 2006.
- [2] D. H. Hubel and T. N. Wiesel. Receptive fields and functional architecture of monkey striate cortex. *J. Physiol. London*, pages 215–243, 1968.
- [3] P. Schiller, B. Finlay, and S. Volman. Quantitative studies of single-cell properties in monkey striate cortex. *Journal of Neurophysiology*, 39:1288–1351, 1976.
- [4] R. De Valois, E. Yund, and N. Hepler. The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research*, 22:531–559, 1982.
- [5] D. H. Hubel and T. N. Wiesel. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, 28:229–289, 1965.
- [6] Y. Petrov, M. Carandini, and S. McKee. Two distinct mechanisms of suppression in human vision. *The Journal of Neuroscience*, 25:8704–8707, 2005.
- [7] O. Schwartz and E. P. Simoncelli. Natural signal statistics and sensory gain control. *Nature Neuroscience*, 4:819–825, 2001.
- [8] M. W. Cannon and S. C. Fullenkamp. Spatial interactions in apparent contrast: Inhibitory effects among grating patterns of different spatial frequencies, spatial positions and orientations. *Vision Research*, 31:1985–1998, 1991.
- [9] U. Polat and D. Sagi. Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. *Vision Research*, 33:993–999, 1993.
- [10] U. Polat and D. Sagi. The architecture of perceptual spatial interactions. *Vision Research*, 34:73–78, 1994.
- [11] C. Yu, S. A. Klein, and D. M. Levi. Facilitation of contrast detection by cross-oriented surround stimuli and its psychophysical mechanisms. *Journal of Vision*, 2(3):243–255, 2002.
- [12] B. Zenger-Landolt and C. Koch. Flanker effects in peripheral contrast discrimination: psychophysics and modeling. *Vision Research*, 41:3663–3675, 2001.

- [13] J. A. Solomon and M. J. Morgan. Facilitation from collinear flanks is cancelled by non-collinear flanks. *Vision Research*, 40:279–286, 2000.
- [14] C. B. Williams and R. F. Hess. Relationship between facilitation at threshold and suprathreshold contour integration. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, 15:2046–2051, 1998.
- [15] J. J. Andriessen and H. Bouma. Eccentric vision: Adverse interactions between line segments. *Vision Research*, 16:71C78, 1976.
- [16] Y. Petrov, M. Carandini, and S. McKee. Two distinct mechanisms of suppression in human vision. *The Journal of Neuroscience*, 25:8704C8707, 2005.
- [17] R. J. Snowden and S. T. Hammett. The effects of surround contrast on contrast thresholds, perceived contrast and contrast discrimination. *Vision Research*, 38:1935C1945, 1998.
- [18] Robertson R. Mahmoodi S. Guo, K. and M. Young. Centre-surround interactions in response to natural scene stimulation in the primary visual cortex. *European Journal of Neuroscience*, 21:536–548, 2005.
- [19] J. Xing and D. J. Heeger. Measurement and modeling of center-surround suppression and enhancement. *Vision Research*, 41:571–583, 2001.
- [20] Y. Ejima and S. Takahashi. Apparent contrast of a sinusoidal grating in the simultaneous presence of peripheral gratings. *Vision Research*, 25:1223–1232, 1985.
- [21] X. Zhang, J. C. Park, J. Salant, S. Thomas, J. Hirsch, and D. C. Hood. A multiplicative model for spatial interaction in the human visual cortex. *Journal of Vision*, 8(8):4, 1–9, 2008.
- [22] J. R. Cavanaugh, W. Bair, and J. A. Movshon. Nature and interaction of signals from the receptive field center and surround in macaque v1 neurons. *Journal of Neurophysiology*, 88:2530–2546, 2002.
- [23] E.P. Simoncelli and O. Schwartz. Modeling surround suppression in v1 neurons with a statistically-derived normalization model. *Advances in Neural Information Processing Systems*, 11:153–166, 1999.