Explaining Low-Level Brightness-Contrast Illusions Using Disinhibition

Yingwei Yu[1], Takashi Yamauchi[2], and Yoonsuck Choe[1]

[1] Department of Computer Science,
[2] Department of Psychology,
Texas A&M University,
College Station, TX 77843, USA,
yingwei@tamu.edu, tya@psyc.tamu.edu, choe@tamu.edu

Abstract. Conventional Difference of Gaussian (DOG) filter is usually used to model the early stage of visual processing. However, convolution operation used with DOG does not explicitly account for the effects of disinhibition. Because of this, complex brightness-contrast (B-C) illusions such as the White's effect cannot be explained using DOG filters. We discovered that a model based on *lateral disinhibition* in biological retinas allows us to explain subtle B-C illusions. Further, we show that a feedforward filter can be derived to achieve this operation in a single pass. The results suggest that contextual effects can be processed through recurrent disinhibition, and that a similar analysis may be applicable to higher brain functions. Another observation is that the feedback loop of the recurrent network structure may improve the overall stability of the system. Potential applications of the IDOG algorithm include new design of video capturing or display system with increased robustness and automatic detection and correction of perceived incoherences in luminance in video display panels, where accurate perception of intensity level is critical.

1 Introduction

Brightness-contrast (B-C) illusions allow us to understand the basic processes in the early visual pathway. B-C illusions can become very complex, and a complete explanation may have to be based on a multi-stage, multi-channel model, with considerations of top-down influences [1–3]. In this paper, however, we will focus on the very early stages of visual processing, and see how far we can exploit low-level mechanisms observed in biological vision systems toward explaining B-C illusions.

For example, the dark illusory spots at the intersections in the Hermann grid (Figure 1A) are due to lateral inhibition in the retina and the lateral geniculate nucleus (LGN) [4]. The visual signal in the eye is generated by the photoreceptor cells, and then it is passed through bipolar, horizontal, and amacrine cells and finally goes to LGN. Lateral inhibition is the effect observed in the receptive field where the surrounding inhibits the center area. When the stimulus is given

In Proceedings of the First International Workshop on Biologically Inspired Approaches to Advanced Information Technology (Bio-ADIT 2004), Lausanne, Switzerland, 2004.

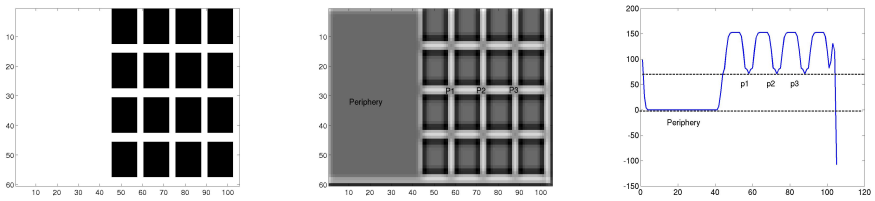

Fig. 1. The Hermann grid illusion **A**. The Hermann grid illusion. The intersections look darker than the streets. **B**. The output using a conventional DOG filter. **C**. The plot of brightness level prediction (from B). To measure the average response, we took the column-wise sum of rows 27 to 29. Note that the illusory spots (at positions p1, p2, and p3) have a brightness value much higher than the periphery. The conventional operation cannot explain why we perceive the periphery to be brighter than the dark illusory spots.

in the receptive field, the central receptors produce an excitatory signal, while the cells in the surrounding area send inhibition through the bipolar cells to the central area [5]. (Difference of Gaussian, or DOG, filter [6] is commonly used to simulate such a process.) Figure 1B and C show such an effect by using DOG filters. The plot on the right shows the brightness level of the middle row, and the dark illusory spots are clearly visible (p1, p2, and p3).

However, DOG filters alone cannot account for more complex visual B-C illusions. For example in the Hermann grid illusion, although the illusory spots get explained pretty well, the conventional DOG model cannot explain why the periphery (figure 1A, to the left) appears brighter than the illusory spots (figure 1A, to the right). This output is counter to our perceived experience. The reason for this failure is that the center of DOG in the peripheral area receives inhibition from all the directions which results in a weaker response than the intersections in the grid which only receive inhibition from four directions. Moreover, the White's effect [7] (figure 2A) cannot be explained using the conventional DOG filter. As shown in figure 2B, the output using conventional DOG filters gives an opposite result: The left gray patch on the black strip has a lower output value than the one on the white strip. On the contrary, we perceive that the left gray patch on the black strip as brighter than the one on the right.

Anatomical and physiological observations show that the center-surround property in early visual processing may not be strictly feed-forward, and it involves lateral inhibitions and, moreover, disinhibition. Hartline et al. used Limulus optical cells (figure 3) to demonstrate lateral inhibition and disinhibition effects in the receptive field [8]. (Note that disinhibition has also been found in vertebrates retinas such as in tiger salamanders [9] and in mice [10].) Disinhibition can effectively reduce the amount of inhibition in the case if we have a large

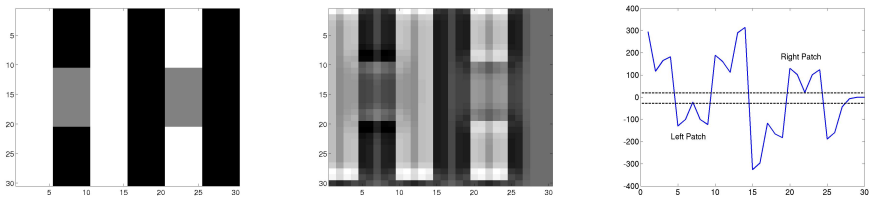

Fig. 2. The White's effect **A**. The White's effect. The gray patch on the left has the same gray level as the one on the right, but we perceive the left to be brighter than the right. **B**. The output using a conventional DOG filter. **C**. The brightness level of the two gray patches calculated using conventional DOG filter. As in the previous figure, we added up rows of 10 to 19 in the output to get the average response. Note that the left patch has a lower average value (below zero) than the right patch (above zero). The result contradicts our perceived brightness.

area of bright input, which might be the solution to the unsolved visual illusion problem.

In this paper, unlike DOG, we explicitly model disinhibition to derive a filter that is able to explain a wider variety of B-C illusions than the conventional DOG filters. In the following, we first review the model of Hartline et al. [8, 11, 12], and introduce our model which is called the Inversed DOG model (or IDOG) and show how it is derived. The next section shows the results to various illusions. Finally, the issues raised by our model is discussed, followed by the conclusion.

2 Hartline-Ratliff's Model of Disinhibition

Experiments on Limulus optical cells showed that the disinhibition effect is recurrent (figure 3). The final response of a specific neuron can be considered as the overall effect of the response from itself and from all other neurons. Conventional convolution operation using the DOG filter does not account for the effect of disinhibition which plays an important role in the final response. The final response of each receptor resulting from a light stimulus can be enhanced or reduced due to the interactions through inhibition from its neighbors, which may be important. It turns out that this effect can help solve some unsolved problems of B-C illusions, thus, it may be important to explicitly account for disinhibition.

The Hartline-Ratliff equation describing disinhibition in Limulus can be summarized as follows [8, 11, 12]:

$$r_m = \epsilon_m - K_s r_m - \sum k_{m \leftarrow n}(r_n - t_{m \leftarrow n}) \tag{1}$$

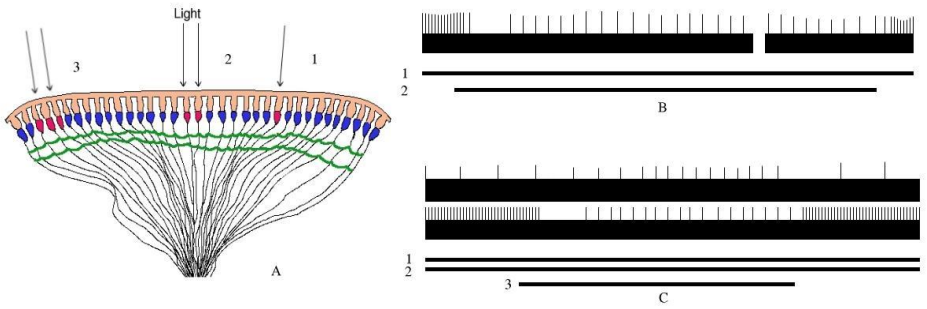

Fig. 3. Lateral inhibition in Limulus optical cells (Redrawn from [8]) The figure shows the disinhibition effect in Limulus optical cells. **A**. The retina of Limulus. Point light is presented to three locations (1, 2 and 3). **B**. The result of lighting position 1 and 2. The top trace shows the spike train of the neuron at 1, and the two bars below show the duration of stimulation to cell 1 and 2. When position 2 is excited, the neuron response of position 1 gets inhibited. **C**. Both 1 and 2 are illuminated, and after a short time, position 3 is lighted. The top two traces show the spike trains of cell 1 and cell 2. The three bars below are input duration to the three cells. As demonstrated in the figure, when position 3 is lighted, neurons at position 2 get inhibited by 3, so its ability to inhibit others get reduced. As a result, the firing rate of neuron at position 1 gets increased during the time neuron at position 3 is excited. This effect is called disinhibition.

where r_m is the response, K_s is the self-inhibition constant, ϵ_m is excitation of the m-th ommatidium, $k_{m \leftarrow n}$ is the inhibitory weight from other ommatidium, and $t_{m \leftarrow n}$ the threshold.

Brodie et al. extended this equation to derive a spatiotemporal filter, where the input was assumed to be a sinusoidal grating [13]. This model is perfect in predicting Limulus retina experiments as only a single spatial frequency channel filter, which means that only a fixed spatial frequency input is allowed [13]. Because of this reason, their model cannot be applied to a complex image (e.g., visual illusions such as the Hermann grid illusion), as various spatial frequencies could coexist in the input. In the following section, we will build upon the Hartline-Ratliff equation and derive a filter that can address these issues.

3 Simplified Disinhibition Model Using Single Matrix Inverse Operation: The IDOG Model

Rearranging equation 1 and generalizing to n inputs, the responses of n cells can be expressed in a simple matrix form as below by assuming the threshold and self-inhibitory constant to be zero (in this paper, we only care for spatial properties of visual illusion, so the assumption of zero self-inhibition rate is reasonable):

$$Kr = e, \tag{2}$$

where r is the output vector, e is the input vector and K is the weight matrix:

$$r = \begin{bmatrix} r_1 \\ r_2 \\ . \\ r_n \end{bmatrix}, e = \begin{bmatrix} e_1 \\ e_2 \\ . \\ e_n \end{bmatrix}, K = \begin{bmatrix} 1 & -w(1) & .. & -w(n-1) \\ -w(1) & 1 & .. & -w(n-2) \\ .. & .. & .. & .. \\ -w(n-1) & .. & .. & 1 \end{bmatrix}. \quad (3)$$

The above 1D model can be easily extended to 2D by serialization. We can serialize the input and output to 1D vectors to fit in the 1D model we have. The weight matrix K can be fitted into 2D by assigning the weight K_{ij} from neuron j to neuron i as the classic two-mechanism DOG distribution [6]:

$$K_{ij} = \begin{cases} -w(|i,j|) \text{ when } i \neq j \\ \\ 1 \qquad\qquad \text{when } i = j \end{cases}, \quad (4)$$

$$w(x) = \text{DOG}(x) = k_c e^{-(x/\sigma_c)^2} - k_s e^{-(x/\sigma_s)^2}, \quad (5)$$

where $|i,j|$ is the Euclidean distance between neuron i and j, k_c and k_s are the scaling constants that determine the relative scale of the excitatory and inhibitory distributions, and σ_c and σ_s their widths.

The response vector r can finally be derived from equation 2 as follows, and we can apply inverse serialization operation to get vector r back into 2D format:

$$r = K^{-1}e. \quad (6)$$

Figure 4 shows a single row (corresponding to a neuron in the center) of the weight matrix K, plotted in 2D. The plot shows that the neuron in the center can be influenced by the inputs from locations far away, outside of its own receptive field.

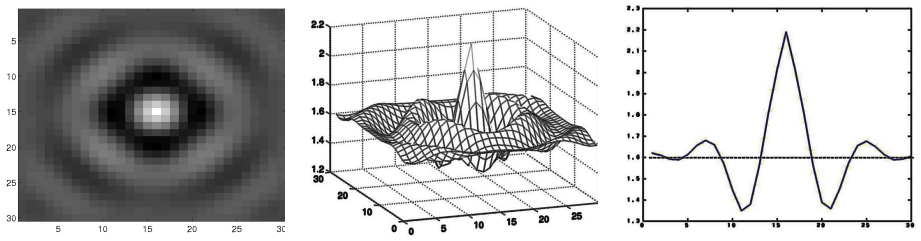

Fig. 4. An Inversed DoG filter The filter (i.e., the connection weights) of the central neuron is shown in log scale. **A**. A 2D plot of the filter. **B**. A 3D mesh plot of the filter. **C**. The plot of the central row of the filter. Note the multiple concentric rippling tails.

4 Results

In this section, we will test our IDOG model first with three Limulus cells and then with several B-C illusions (Hermann grid, the White's effect, and Mach band). Based on these experiments, we will demonstrate that disinhibition does play an important role in early visual processing.

4.1 Disinhibition in 1D: A model of the Limulus retinal cells

Reconsidering the Limulus experiments, let us suppose three Limulus cells have the same input, say 100. We assigned the weights based on the distance $w(1) = -0.5$ and $w(2) = -0.1$, which indicates that if the cells are near neighbors, their inhibition effect is 50%, while if they are remote neighbors, the effect is reduced to 10%. The response r is then calculated as follows:

$$r = K^{-1}e = \begin{bmatrix} 1 & 0.5 & 0.1 \\ 0.5 & 1 & 0.5 \\ 0.1 & 0.5 & 1 \end{bmatrix}^{-1} \times \begin{bmatrix} 100 \\ 100 \\ 100 \end{bmatrix} = \begin{bmatrix} 83.333 \\ 16.667 \\ 83.333 \end{bmatrix}.$$

If we increase the input a little bit (5%) to the neuron e_1, the result becomes different as shown below:

$$r = K^{-1}e = \begin{bmatrix} 1 & 0.5 & 0.1 \\ 0.5 & 1 & 0.5 \\ 0.1 & 0.5 & 1 \end{bmatrix}^{-1} \times \begin{bmatrix} 105 \\ 100 \\ 100 \end{bmatrix} = \begin{bmatrix} 90.227 \\ 12.500 \\ 84.722 \end{bmatrix}.$$

The third neuron increased the response from 83.333 to 84.722, since the second neuron was further inhibited by the first neuron (the response decreased from 16.667 to 12.5000), which had its input increased from 100 to 105. This result matches the experimental results from Hartline et al. [8], clearly demonstrating the disinhibition effect in our model.

4.2 Disinhibition in 2D: the Hermann grid illusion

In the Hermann grid, the illusory spots can be modeled quite well using conventional DOG filters. However, conventional DOG filters cannot explain why the periphery area appears brighter than the dark illusory spots. Convolving with conventional DOG filters results in more inhibition to the white peripheral area than the intersections in the grid, because the periphery gets inhibition from all radial directions while the intersection only get inhibition from four directions.

Our IDOG filter which explicitly models disinhibition provides a plausible explanation to this problem. Figure 5 shows the result of applying our filter to the Hermann grid image: C is the plot of the middle row of the filter response in B. The periphery is indeed brighter than the dark illusory spots, showing that disinhibition (and hence IDOG) can account for the perceived brightness in this particular example.

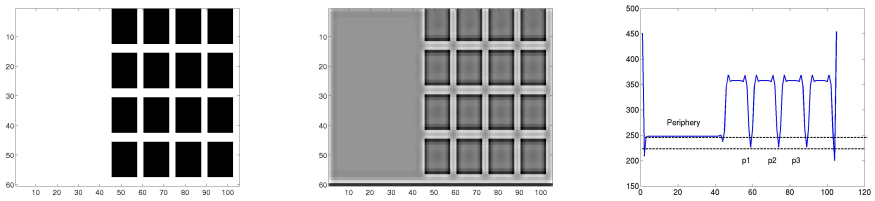

Fig. 5. The Hermann grid illusion and prediction **A**. Part of the Hermann grid which we used to test the response of the periphery and the illusory spots. **B**. The output response of IDOG. **C**. The prediction using the IDOG filter (from B). The illusory spots are at position p1, p2 and p3, which have a brightness value lower than the periphery. (The curve shows the column-wise sum of rows 27 to 29.)

4.3 Disinhibition in 2D: the White's effect

The White's effect [7] is shown in figure 6A: The gray patch on the black vertical strip appears brighter than the gray patch on the right. As shown in figure 2, DOG cannot explain this illusion. However, disinhibition plays an important role in this illusion: While the gray patch on the black strip receives inhibition from the two surrounding white strips, compared to the gray patch on the right side, disinhibition is relatively stronger. Because of this, the gray patch on the right side appears darker than the left side patch (C in figure 6).

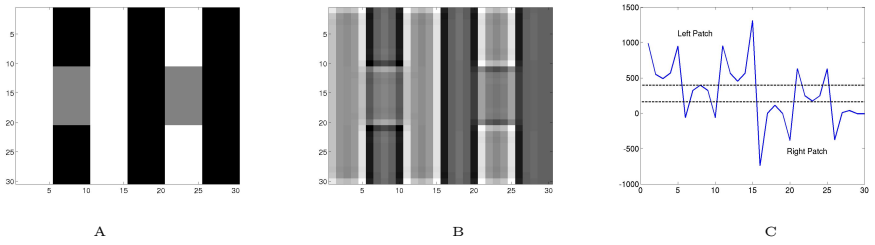

Fig. 6. The White's effect and prediction **A**. The White's effect stimulus. **B**. The output using IDOG. **C**. The prediction using the IDOG model (from B). The gray patch on the left results in a higher value than the right patch. (The curve shows the column-wise sum of rows 11 to 19.)

4.4 The Mach band

Comparing with the conventional DOG filter, one advantage of the IDOG model is that it preserves the different level of brightness as well as enhances the contrast at the edge. As demonstrated in figure 7, the four shades of gray are clearly

separated using IDOG. These different shades are not preserved using a conventional DOG filter. Note that this can be simply because the sum of the DOG matrix equals zero, and scaling up k_c in equation 5 can correct the problem. However, there is one subtle point not captured in the conventional DOG approach: the wrinkle (figure 7E) near the Mach bands observed in Limulus experiments [14]. Compared to the IDOG result, we can clearly see that this wrinkle is absent in the DOG output (figure 7C) .

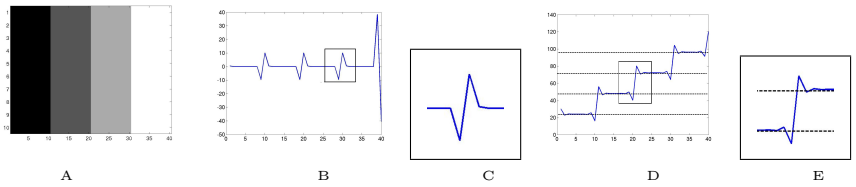

Fig. 7. The Mach band **A**. The Mach band input image. **B**. The output using a conventional DOG filter. The different brightness levels are not preserved. **C**. An expanded view of the inset in B. **D**. The output using IDOG. The different brightness levels are preserved **E**. An expanded view of the inset in D.

5 Discussion and Future Work

We have shown that by explicitly modeling disinhibition, we can more accurately explain various B-C illusions. Although there are many other improved DOG filter models, such as the oriented DOG (ODOG) filter proposed by Blakeslee and McCourt [15], they still cannot (under our analysis) explain certain problems like the phenomenon related to the periphery area of the Hermann grid (figure 1).

Our model is strongly motivated by biological facts as well as computational considerations. First, experimental evidence shows that the inhibition in the retinal receptive fields can be explained by the isotropic amacrine and horizontal cells. Second, we utilize the classical two-mechanism DOG distribution. Third, as the experiments demonstrated by Hartline and colleagues using Limulus cells, disinhibition is a natural effect of recurrent lateral inhibition, which does not work well with a single-pass convolution operation. Another interesting observation is that the IDOG filter has a similar shape as the circular Gabor filter [16]. Circular Gabor filters have been successfully used in rotation-invariant texture discrimination, and it would be interesting to see if IDOG can also be used in such a domain. Also, there is further psychophysical evidence [17] suggesting that early visual processing can be modeled by filters similar to our disinhibition-based IDOG filters.

Another observing is that the loop back feature of disinhibition network may increase the overall robustness. Because disinhibition is recurrent inhibition, each individual cell has output connected with some other cells and at the same time

those cells send their output to this individual as feedback controls to this cell. Comparing with a group of cell, one single cell usually has less stability. In the case, one individual cell becomes less stable, the other cells in the loop could adjust the cell by the feedback. Thus, this loop back feature may increase the stability of the individual. To summarize , this recurrent network structure can keep the feature of common feedforward lateral inhibition as to sharpen the edge, perserve the brightness contrast information as much as perceived better than feedforwards, and increase the stability of the system.

One limitation to our approach is that the inverse weight matrix results in a non-local operation, thus it can be computationally inefficient. To overcome this issue, we can use an approximated algorithm. Based on our observation, the IDOG filter usually converges to a value near zero at a distance twice that of the DOG-based receptive field. We can use the IDOG filter which is twice the original receptive field size and still use a local convolution operation to process larger images.

Potential applications of IDOG algorithm include designing new robust visual capturing or display systems and automatic detection and correction of perceived incoherences in luminance in video display panels, where accurate perception of intensity level is critical. Such applications will be important in mission-critical domains such as aircraft display panel design. The concept of disinhibition can also be applied to higher brain functions such as categorization and memory (e.g., Vogel [18] proposed a model of associative memory based on disinhibition). We believe a close analysis of cortical horizontal connections and their physiology under the disinhibition framework can provide us with new insights on their functions. This in turn will allow us to apply the general concept of disinhibition in advanced intelligent systems, firmly based on biological observations.

6 Conclusion

We have shown that certain limitations of DOG filters can be overcome by explicitly modeling disinhibition, and that a simple feedforward filter can be derived. Using the IDOG filter, we were able to successfully explain several B-C illusions that were not sufficiently explained in previous models. Our work also shows that complicated recursive effects can be explicitly calculated or approximated using a single matrix multiplication. The results suggest that contextual effects can be processed through recurrent disinhibition, and a similar analysis may be applicable to higher brain functions. Such an analysis will allow us to apply disinhibition in building advanced intelligent systems based on a firm grounding on biology.

Acknowledgments

The authors wish to thank Jyh-Charn Liu for helpful discussions. This research was supported in part by the Texas Higher Education Coordinating Board ARP/ATP program grant 000512-0217-2001.

References

1. Adelson, E.H.: Lightness perception and lightness illusions. In Gazzaniga, M., ed.: The New Cognitive Neurosciences. 2nd edn. MIT Press, Cambridge, MA (2000) 339–351
2. Fiorentini, A., Baumgartner, G., Magnussen, S., Schiller, P.H., Thomas, J.P.: The perception of brightness and darkness: Relations to neuronal receptive fields. In Spillmann, L., Werner, J.S., eds.: Perception: The Neurophysiological Foundations. Academic Press (1990) 129–161
3. Kelly, F., Grossberg, S.: Neural dynamics of 3-D surface perception: Figure-ground separation and lightness perception. Perception and Psychophysics **62** (2000) 1596–1618
4. Spillmann, L.: The Hermann grid illusion: a tool for studying human perceptive field organization. Perception **23** (1994) 691–708
5. Goldstein, E.B.: Sensation and Perception. Sixth edn., Wadsworth-Thomson Learning (2000)
6. Marr, D., Hildreth, E.: Theory of edge detection. Proc. R. Soc. Lond. B **207** (1980) 187–217
7. White, M.: A new effect of pattern on perceived lightness. Perception **8** (1979) 413–416
8. Hartline, H.K., Ratliff, F.: Inhibitory interaction of receptor units in the eye of Limulus. Journal of General Physiology **40** (1957) 357–376
9. Roska, B., Nemeth, E., Werblin, F.: Response to change is facilitated by a three-neuron disinhibitory pathway in the tiger salamander retina. Journal of Neuroscience **18** (1998) 3451–3459
10. Frech, M.J., Perez-Leon, J., Wassle, H., Backus, K.H.: Characterization of the spontaneous synaptic activity of amacrine cells in the mouse retina. Journal of Neurophysiology **86** (2001) 1632–1643
11. Hartline, H.K., Ratliff, F.: Spatial summation of inhibitory influences in the eye of Limulus, and the mutual interaction of receptor units. Journal of General Physiology **41** (1958) 1049–1066
12. Stevens, C.F.: A Quantitative Theory of Neural Interactions: Theoretical and Experimental Investigations. PhD thesis, The Rockefeller Institute (1964)
13. Brodie, S., Knight, B.W., Ratliff, F.: The spatiotemporal transfer function of the limulus lateral eye. Journal of General Physiology **72** (1978) 167–202
14. Ratliff, F., Hartline, H.K.: The responses of limulus optic nerve fibers to patterns of illumination on the receptor mosaic. Journal of General Physiology **42** (1959) 1241–1255
15. Blakeslee, B., McCourt, M.E.: A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. Vision Research **39** (1999) 4361–4377
16. Zhang, J., Tan, T., Ma, L.: Invariant texture segmentation via circular Gabor filters. 16 th International Conference on Pattern Recognition (ICPR'02) **2** (2002)
17. Adini, Y., Sagi, D., Tsodyks, M.: Excitatory-inhibitory network in the visual cortex: Psychophysical evidence. Proc. Natl. Acad. Sci. **94** (1997) 10426–10431
18. Vogel, D.: A biologically plausible model of associative memory which uses disinhibition rather than long-term potentiation. Brain and Cognition **45** (2001) 212–228