Modeling Disinhibition in the Early Visual Pathway

Yingwei Yu and Yoonsuck Choe
Department of Computer Science
Texas A&M University
College Station, TX 77843-3112
{yingwei,choe}@tamu.edu

Abstract

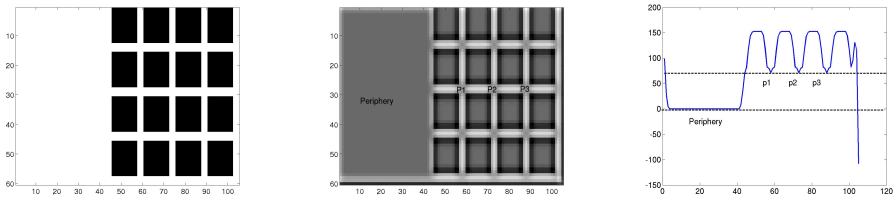
Conventional Difference of Gaussians (DoG) filter is usually used to model the early stage of visual processing. However, the convolution operation used with DoG does not explicitly account for the effects of disinhibition. Because of this, complex brightness contrast illusions such as the White effect cannot be explained using DoG filters. We discovered that a new model explicitly accounting for disinhibition allows us to explain why these brightness contrast illusions arise, and show that a feedforward filter can be derived to achieve this in a single pass.

1 Introduction

Brightness contrast illusions allow us to understand the basic process of the early visual pathway. 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 LGN [1]. 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 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 [2]. DoG filter [3] 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).

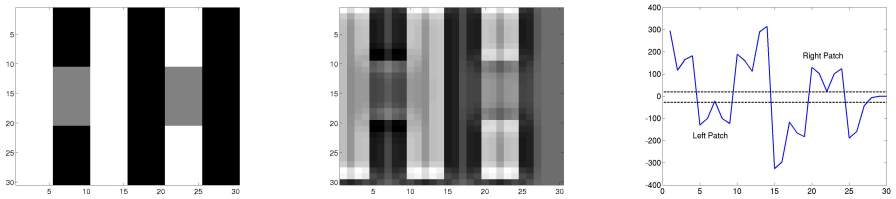
However, DoG filters alone cannot account for more complex visual brightness-contrast illusions. For example in the Hermann grid illusion, although the illusory spots get explained pretty well, the conventional DoG model cannot explain why the periphery appears brighter than the illusory spots. This output contradicts with our perception. The reason for this failure is that the peripheral area receives inhibitions from all the directions which produce a darker results than the intersections which only receive inhibitions from 4 directions. Moreover, the White effect [4] (figure 2A) cannot be explained using the conventional DoG filter. As shown in figure 2B, the output of the conventional DoG filter 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.

Yingwei Yu and Yoonsuck Choe. Modeling disinhibition in the early visual pathway. Technical Report 2003-8-6, Department of Computer Science, Texas A&M University, 2003.



A B C

Figure 1: **The Hermann grid illusion.** A.The Hermann grid. The intersections look darker than the streets. B.The output of conventional DoG filter. C.The plot of brightness level prediction (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.



A B C

Figure 2: **The White effect.** A. The White effect image. 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 of 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.

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 the limulus optical cells (figure 3) to demonstrate the lateral inhibition and disinhibition effects in the receptive field [5]. Disinhibition can reduce the amount of inhibition in the case if we have a large area of light input, which might be the solution to the unsolved visual illusion problem. However, disinhibition is not explicitly accounted for in the conventional DoG model.

In this paper, we explicitly model the disinhibition mechanism to derive a new filter that is able to explain a wider variety of brightness-contrast illusions than the conventional DoG filters. The next section introduces our model which is called the inverted DoG model (or IDOG) and shows how this model is derived. Section 3 shows the results to various illusions. Section 4 discusses the issues raised by our model, followed by the conclusion.

2 Disinhibition in Early Visual Processing

Experiments on the 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

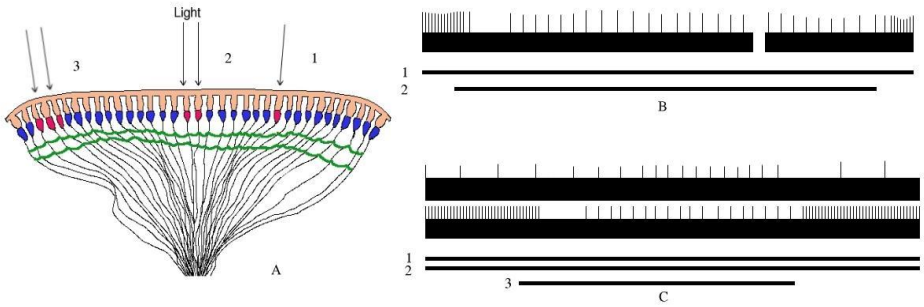


Figure 3: **Lateral inhibition in the limulus optical cells (Redrawn from [5]).** The figure shows the disinhibition effect in limulus optical cells. A. The retina of limulus. Point light is presented to three locations (1, 2 and 3). B. The result of lighting position 1 and 2. The top trace shows the 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.

unsolved problems of brightness contrast illusions, thus, it may be important to explicitly account for disinhibition.

3 Derivation of IDOG in 1D

Based on the observation in the previous section, we built a model utilizing disinhibition. Figure 4 shows an overview of our model based on the limulus retina discussed in [5]. The weights among these neurons are determined by the distance d between them, as expressed by the function $w(d)$ in the figure (the weights are symmetric).

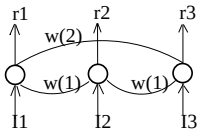


Figure 4: **Three fully connected optical-cell model based on the limulus experiment.** According to the experimental results with limulus cells, the three cells are fully connected so that inhibition can take effect recurrently. The weights between the neurons are based on their distance d, which is expressed as a function of distance $w(d)$.

We can write an activation equation from this network (the formulation is similar to [5]).

$$\begin{cases} r_1 = I_1 + w(1)r_2 + w(2)r_3 \\ r_2 = I_2 + w(1)r_1 + w(1)r_3 \\ r_3 = I_3 + w(1)r_2 + w(2)r_1 \end{cases}, \tag{1}$$

where I_1, I_2, and I_3 are the input values; and r_1, r_2, and r_3 are the output values for the three neurons as shown in figure 4.

Rearranging equation 1 and generalizing to n inputs, the responses of n cells can be expressed in matrix form as below:

$$D \times R = I, \tag{2}$$

where R is the output vector, I is the input vector and D is the weight matrix:

$$R = \begin{bmatrix} r_1 \\ r_2 \\ . \\ r_n \end{bmatrix}, I = \begin{bmatrix} I_1 \\ I_2 \\ . \\ I_n \end{bmatrix}, D = \begin{bmatrix} 1 & -w(1) & .. & -w(n-1) \\ -w(1) & 1 & .. & -w(n-2) \\ .. & .. & .. & .. \\ -w(n-1) & .. & .. & 1 \end{bmatrix}. \quad (3)$$

To get the value of the weight D_{ij} from neuron j to neuron i, we can apply the classic two-mechanism DoG distribution [3]:

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

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

where 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:

$$R = D^{-1} \times I. \quad (6)$$

4 Derivation of IDOG in 2D

To make our model more general to handle 2-dimensional images, we serialize our input and output matrix into a vector. The calculation of the weight matrix D is a bit more complicated. For example, let us first look at a simple case of 3×3 neuron array (figure 5).

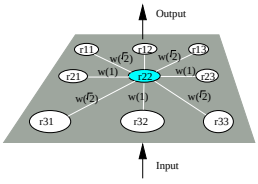


Figure 5: **Nine fully connected retinal cells: A small example.** The figure shows the connections of the neuron in the center: other neurons have the same full connections to each others in a similiar manner. The weights are determined by the spatial distance between the neurons. For example, the distance from r_{11} to r_{22} is $\sqrt{2}$, so the weight is set to $w(\sqrt{2})$.

For the neuron r_{22} shown in 5, the activity is defined as below, and consequently, we can write similar equations for all the other cells:

$$\begin{aligned} r_{22} \quad = \quad & I_{22} + w(\sqrt{2})r_{11} + w(1)r_{12} + w(\sqrt{2})r_{13} + w(1)r_{21} \\ & + w(1)r_{23} + w(\sqrt{2})r_{31} + w(1)r_{32} + w(\sqrt{2})r_{33} \end{aligned} \quad (7)$$

The weight matrix D can then be derived from the system of equations in the form of equation 7. Each row corresponds to one cell in the network, and it represents the collective effect of the afferent and lateral inputs.

For example, for a network as shown in figure 5, the D matrix is defined as

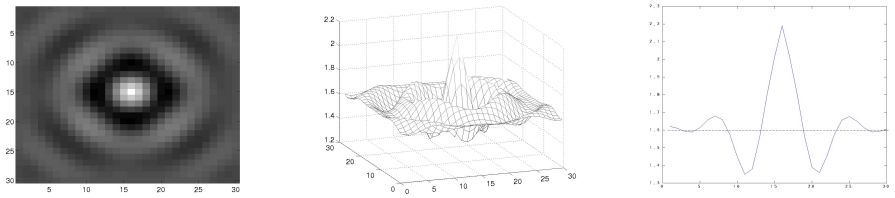
$$D = \begin{bmatrix} 1 & -w(1) & -w(2) & -w(1) & -w(\sqrt{2}) & -w(\sqrt{5}) & -w(2) & -w(\sqrt{5}) & -w(2\sqrt{2}) \\ -w(1) & 1 & -w(1) & -w(\sqrt{2}) & -w(1) & -w(\sqrt{2}) & -w(\sqrt{5}) & w(2) & -w(\sqrt{5}) \\ -w(2) & -w(1) & 1 & -w(\sqrt{5}) & -w(\sqrt{2}) & -w(1) & -w(2\sqrt{2}) & -w(\sqrt{5}) & -w(2) \\ -w(1) & -w(\sqrt{2}) & -w(\sqrt{5}) & 1 & -w(1) & -w(2) & -w(1) & -w(\sqrt{2}) & -w(\sqrt{5}) \\ -w(\sqrt{2}) & -w(1) & -w(\sqrt{2}) & -w(1) & 1 & -w(1) & -w(\sqrt{2}) & -w(1) & -w(\sqrt{2}) \\ -w(\sqrt{5}) & -w(\sqrt{2}) & -w(1) & -w(2) & -w(1) & 1 & -w(\sqrt{5}) & -w(\sqrt{2}) & -w(1) \\ -w(2) & -w(\sqrt{5}) & -w(2\sqrt{2}) & -w(1) & -w(\sqrt{2}) & -w(\sqrt{5}) & 1 & -w(1) & -w(2) \\ -w(\sqrt{5}) & w(2) & -w(\sqrt{5}) & -w(\sqrt{2}) & -w(1) & -w(\sqrt{2}) & -w(1) & 1 & -w(1) \\ -w(2\sqrt{2}) & -w(\sqrt{5}) & -w(2) & -w(\sqrt{5}) & -w(\sqrt{2}) & -w(1) & -w(2) & -w(1) & 1 \end{bmatrix}. \tag{8}$$

To further generalize our model to an $m \times n$ input matrix, the size of the weight matrix D becomes $mn \times mn$. The weight D_{ij} can then be calculated as below:

$$D_{ij} = \begin{cases} -w(E(Pi, Pj)) & \text{When } i \neq j \\ 1 & \text{When } i = j \end{cases}, \tag{9}$$

where Pi and Pj are the corresponding location of neurons i and j in the 2D network. The E function simply calculates the Euclidean distance between the points Pi and Pj. Same as our previous model in one dimension, the output R can then be calculated using equation 6.

Figure 6 shows a single row (the neuron in the center) of the weight matrix D, 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.



A B C

Figure 6: **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 ring-shaped area with positive weights.

5 Results

In this section, we will test our IDOG model first with the 3 limulus cells and then with several brightness contrast illusions (Hermann grid, the White effect, and Mach band). Based on these experiments, we will demonstrate that disinhibition does play an important role in early visual processing.

5.1 Disinhibition in 1D: A model of the limulus retinal cells

Reconsidering the limulus cells experiments, let us suppose three limulus cells have the same input, say 100. We assign arbitrary value to the weight 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 calculated as follows:

$$R = D^{-1} \times I = \begin{bmatrix} 1 & 0.5 & 0.1 \\ 0.5 & 1 & 0.5 \\ 0.1 & 0.5 & 1 \end{bmatrix}^{-1} \times \begin{bmatrix} 100 \\ 100 \\ 100 \end{bmatrix} = \begin{bmatrix} 83.333 \\ 16.667 \\ 83.333 \end{bmatrix}.$$

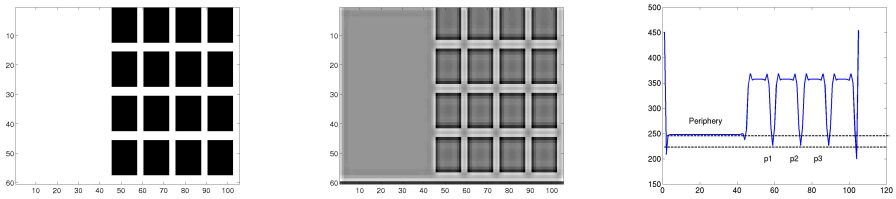
If we increase the input a little bit (5%) to the neuron r_1, the result becomes different as shown below:

$$R = D^{-1} \times I = \begin{bmatrix} 1 & 0.5 & 0.1 \\ 0.5 & 1 & 0.5 \\ 0.1 & 0.5 & 1 \end{bmatrix}^{-1} \times \begin{bmatrix} 105 \\ 100 \\ 100 \end{bmatrix} = \begin{bmatrix} 90.227 \\ 12.500 \\ 84.722 \end{bmatrix}.$$

Obviously, the third neuron increases the firing rate from 83.333 to 84.722, since the second neuron gets more inhibition (changes from 16.667 to 12.5000) from the first neuron, which has its input increased from 100 to 105. This result matches that of Hartline et al. [5] on experiment on limulus eye cells, which clearly demonstrates the disinhibition effect.

5.2 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 peripheral white area than the intersections, because the periphery gets inhibition from all directions while the intersection only get inhibition from four directions.



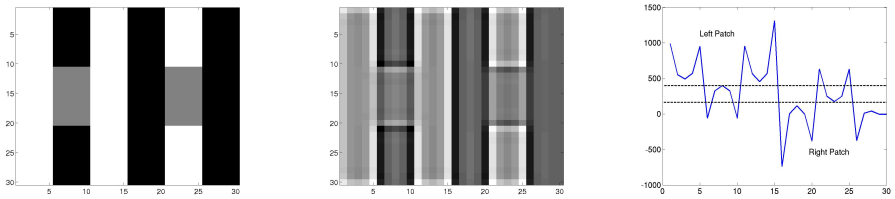
A B C

Figure 7: **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. 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.)

Our Inversed DoG filter which explicitly models disinhibition provides a plausible explanation to this problem. Figure 7 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. Obviously, the periphery is brighter than the dark illusory spots, and this result shows that disinhibition is important in early visual processing.

5.3 Disinhibition in 2D: the White effect

The White effect [4] is shown in figure 8A: 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 8).

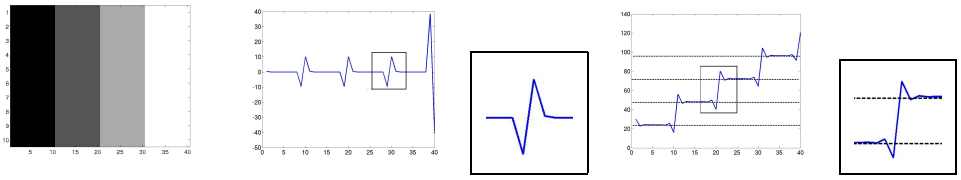


A B C

Figure 8: **The White effect and prediction.** A. The White effect stimulus. B. The output of IDOG. C. The prediction using the IDOG model. 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.)

5.4 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 9, 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 [**?**] can correct the problem. However, there is one subtle point not captured in the conventional DoG approach: the wrinkle (figure 9E) near the Mach bands observed in limulus experiments [6]. Compared to the IDOG result, we can clearly see that this wrinkle is absent in the DoG output (figure 9C) .



A B C D E

Figure 9: **The Mach band.** A. The Mach band input image. B. The output using conventional DoG filter. The different brightness levels are not preserved (the inset is expanded in C). D. The output using IDOG. The different brightness levels are preserved (the inset is magnified in E).

6 Discussion and Future Work

We have shown that by explicitly modeling disinhibition, we can more accurately explain various brightness-contrast illusions. Although there are many other improved DoG filter models, such as the oriented DoG filter proposed by Blakeslee and McCourt (ODOG[7]), they still cannot 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 classic two-mechanism DoG distribution. Third, as the experiments demonstrated by Hartline using limulus cells, disinhibition is a natural effect of lateral inhibition and it is recurrent, which does not work well with a local convolution operation. An interesting observation is that the inversed DoG filter has a similar shape as the circular Gabor filter [8]. Circular Gabor filters have been successfully used in rotation invariant texture discrimination, and it would be interesting to see if IDOG can be used in such a domain. Also, there is further psychophysical evidence [9] suggesting that early visual processing can be modeled by filters similar to our disinhibition-based IDOG filters. One limitation to our approach is that

the inversed 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 inversed DoG filter usually converges to a value near zero at a distance twice that of the DoG-based receptive field. We can use the inversed DoG filter which is twice the original receptive field size and still use a local convolution operation to process larger images.

We are currently working on other types of brightness-contrast illusions, and the results are very promising. Another interesting topic to explore in the future is what kind of computational goal does the IDOG filter fulfill [10]? What kind of computational benefit can filters like IDOG bring to an organism?

7 Conclusion

We have shown that certain limitations of DoG filters can be overcome by explicitly modeling disinhibition, and that a simple feedforward filter (IDOG) can be derived. Using the IDOG filter, we were able to successfully explain several brightness-contrast 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 pass operation.

Acknowledgments

We would like to thank Takashi Yamauchi 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] L. Spillmann. The Hermann grid illusion: a tool for studing human perceptive field organizatio. *Perception*, 23:691–708, 1994.

[2] E. B. Goldstein. *Sensation and Perception*. Wadsworth-Thomson Learning, sixth edition, 2000.

[3] D. Marr and E.C. Hildreth. Theory of edge detection. *Proc. R. Soc. Lond. B*, 207:187–217, 1980.

[4] M. White. A new effect of pattern on perceived lightness. *Perception*, 8:413–416, 1979.

[5] H. K. Hartline and F. Ratliff. Inhibitory interaction of receptor units in the eye of Limulus. *Journal of General Physiology*, 40:357–376, 1957.

[6] F. Ratliff and H. K. Hartline. The responses of limulus optic nerve fibers to patterns of illumination on the receptor mosaic. *Journal of General Physiology*, 42:1241–1255, 1959.

[7] B. Blakeslee and M. E. McCourt. A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. *Vision Research*, 39:4361–4377, 1999.

[8] J. Zhang, T. Tan, and L. Ma. Invariant texture segmentation via circular Gabor filters. *16 th International Conference on Pattern Recognition (ICPR'02)*, 2, 2002.

[9] Y. Adini, D. Sagi, and M. Tsodyks. Excitatory-inhibitory network in the visual cortex: Psychophysical evidence. *Proc. Natl. Acad. Sci.*, 94:10426–10431, 1997.

[10] H. Barlow. What is the computational goal of the neocortex? In C. Koch and J.L. Davis, editors, *Large Scale Neuronal Theories of the Brain*, pages 1–22. MIT Press, Cambridge, MA, 1994.