Chapter 6

Visual Information Processing with Excitatory-Inhibitory Networks

Upto now we have considered in detail only neural networks comprising a small number of processing elements, N - viz., the largest network examined so far, contains three excitatory and three inhibitory neurons (i.e., N = 6). This simplification has made it possible to investigate the properties of such networks as a function of several parameters, without making the attendant complexity too forbidding to allow any analysis. We shall now introduce relatively larger networks (N $\sim 10^4$) for the specific purpose of applying such systems to visual information processing. The specific tasks considered are those of segmentation and adaptive smoothing (followed by edge extraction), which form part of the phenomenon known as "early vision", the first stage in the visual processing path in the brain. At this level, the raw sensory data (intensity value) is the source from which the primitive features are obtained, to construct representations of objects present in the visual field. These are used by higher stages for further processing, ultimately leading to the act of "seeing". Several efforts have been made by the neural network community to model various aspects of vision. These models have been formulated at various levels of abstraction - from a detailed representation of the actual neurobiological structures employed for vision, to a phenomenological explanation. While low level processing can be investigated retaining a degree of fidelity to the biological apparatus, in the case of higher level processing, the enormous complexity of the neurobiological architecture (and the lack of complete understanding of the processes employed) forces one to invoke various simplification to model features such as object recognition [17] and visual memory [173].

In this chapter we have investigated certain features of early vision, using a couple of excitatory-inhibitory neural network models. The process of image segmentation, in particular, object-background segregation, has been examined with a network of

excitatory-inhibitory pairs, coupled over a local neighborhood. The other model introduced, has been inspired by the architecture of the retina and consists of three layers, comprising excitatory, inhibitory and excitatory elements respectively. It is used to implement adaptive smoothing on images - the resulting enhanced image being used for extracting edges.

The rest of the chapter is organized as follows: In Section 1, we have briefly reviewed some of the existing computational techniques used to implement "early vision" tasks, such as, segmentation and edge detection. In the next section, the network model used for dynamical segmentation of images is introduced. A theoretical analysis is presented for the case of an uncoupled network and results for different types of images are presented. In Section 3, we describe the three-layer network model used for adaptive smoothing. The results of applying the proposed method on a gray-level image are presented and compared with some existing techniques. Finally, we conclude with some comments on the relation of the models presented here to existing work reported in the image processing literature.

6.1 The problem of early vision

Early vision is the name given to that part of visual information processing where the principal features of an image are extracted (e.g., information concerning form, motion, color, etc.) and sent to higher brain areas for further processing. Although identification of the portion of the brain which is concerned with early vision is somewhat arbitrary (depending on an individual investigator's interpretation of the word 'early'), for our purpose we shall be concerned with the processing done in the *retina*. Note that, a considerable amount of processing appears to be involved before information from the photoreceptors in the eye reach higher visual centers in the brain, for further interpretation. It appears that this processing is done in subsequent stages so as to use the fewest possible number of active neurons to achieve an adequate representation of the stimulus. Such selective representations have the advantage of reducing considerably the volume of information that has to be propagated through the visual pathway, at any given time.

In this chapter, we shall be concerned with two specific tasks performed during early vision: (*i*) the detection of edges, and (*ii*) the segmentation of similar regions, in an image. We now briefly review a few existing techniques for achieving these objectives through computational means.

6.1.1 Edge detection

Edges are points of discontinuity in the gray level intensity values of an image - and hence, are local features, determined on the basis of local information. A large variety of edge detection methods are available in the image processing literature (e.g.

[118, 29, 23]). A good edge detector should be a filter with the following features:

- it should be a differential operator, taking either a first or second order spatial derivative of the image, and,
- it should be capable of being tuned to act at any desired scale (large filters for blurry edges and small ones to detect sharply focussed fine details).

The second requirement is very useful, as intensity changes occur at different scales in an image.

The simple differential operators used for edge detection are the first difference operators like Roberts gradient, Sobel gradient and Prewitt gradient and second difference operators like the Laplacian operator [64]. These not only respond to edges but also to isolated points. For a noisy picture, the Laplacian operator gives a higher response for noise than for a true edge, unless the noise has a low contrast. Marr and Hildreth [119] proposed the Laplacian of Gaussian (LOG) operator to alleviate this problem. To find intensity changes at a given scale, the image is first convolved with a two-dimensional Gaussian distribution with an appropriate standard deviation σ. This blurs the image, smoothing out all structures at scales much smaller than σ. The image is then filtered through a Laplacian operator and the zero-crossings obtained. The space described by the scale parameter σ and the zero-crossing curves is called the scale space.

Canny [29] has proposed an "optimal" edge detector in terms of good detection, good localization and single response. Good detection, i.e., low probability of detecting non-edges or not detecting the actual edges, is achieved by maximizing the signal-to-noise ratio (SNR), while, good localization (i.e. points marked as edges should be as close as possible to true edges) is achieved by maximizing the reciprocal of the (approximately) standard deviation of the displacement of edge points. Single response, i.e. one and only one response to a single edge point, is achieved by subjecting the optimal filter to a constraint that eliminates multiple responses. The Canny detector has been found to be extremely effective for a wide class of images. A neural network model implementation of Canny's method has also been reported [184].

All the edges produced by the operators discussed above, are normally not significant or relevant edges, when viewed by human beings. Therefore, one needs to extract prominent edges from the output of such edge detectors. Kundu and Pal [107] have suggested a method of thresholding, to extract the prominent edges, based on psychovisual phenomena.

6.1.2 Segmentation

Segmentation is an essential and important step of early vision [64]. It is a process of partitioning the image into some non-intersecting regions, such that, each region is

homogeneous and the union of no two adjacent regions is homogeneous. Segmentation may be formally defined as follows [139]:

> If F is the set of all pixels and $P(\quad)$ is a uniformity/ homogeneity predicate defined on groups of connected pixels, then segmentation is a partitioning of the set F into a set of connected subsets or regions $(S_1, S_2, ..., S_n)$ such that
>
> $$\cap_{(i=1)}^{n} S_i = F$$
>
> with $S_i \cap S_j = \emptyset, i \neq j$. The uniformity predicate $P(S_i)$ = true for all regions S_i, and $P(S_i \cup S_j)$ = false, when S_i, S_j are adjacent regions.

There are several approaches available for image segmentation. In addition to techniques based on histogram thresholding, edge detection, relaxation and semantic & syntactic approaches, several attempts have been made to develop segmentation methods using neural network models, particularly the Hopfield and Kohonen networks. Neural network based methods have the advantages of producing reasonable outputs even in highly noisy environments, as well as, generating results in real time.

Thresholding is an old and simple technique for image segmentation. It can be done based, either on global information (e.g. gray level histogram of the entire image), or on local information of the image. If only one threshold is used for the entire image, it is called global thresholding. On the other hand, when the image is partitioned into several sub-regions and a threshold is determined for each sub-region, it is referred to as local thresholding or adaptive thresholding. Thresholding techniques can also be classified into bi-level and multi-level thresholding. In the former case, the image is partitioned into two regions - object (say, black) and background (white). Thus, here thresholding can be viewed as a classification problem, with the pixels of the image being assigned to one of two classes: object and background. When the image is composed of several objects with different characteristics, one needs several thresholds for segmentation. This is multi-level thresholding.

Segmentation can also be obtained through detection of edges of various regions. A good strategy to produce meaningful segments would be to fuse region segmentation results and edge detector outputs.

Neural network based methods are attempts to achieve robustness with respect to random noise (or failure of processors) and to have real time output. Massive connectionist architecture makes the system robust while the parallel processing enables real time operation. Ghosh *et al* [61] used a Hopfield type network for extraction of objects from highly noise-corrupted scenes. The energy function of the network has been constructed in such a manner that in a stable state of the net, it extracts compact regions from a noisy scene.

6.2 The dynamical segmentation network

In this section, we investigate a model for performing bi-level image segmentation, comprising excitatory and inhibitory neurons, that are coupled to each other over a local neighborhood. The basic module of the proposed network is an excitatory-inhibitory neural pair. If x and y be the activities of the excitatory and the inhibitory elements respectively, then they evolve in time according to:

$$\begin{aligned} x_{n+1} &= F_a(w_{xx}x_n - w_{xy}y_n + I_n) \\ y_{n+1} &= F_b(w_{yx}x_n - w_{yy}y_n + I'_n), \end{aligned} \tag{6.1}$$

where, w_{ij} is the weight of synaptic coupling between elements i and j, F is the activation function defined by Eqn. (2.9) and I, I' are the external stimuli. By imposing the following restriction on the values of the synaptic weights:

$$\frac{w_{xy}}{w_{xx}} = \frac{w_{yy}}{w_{yx}} = k,$$

and absorbing w_{xx} and w_{yy} within a and b (respectively), we can simplify the dynamics to that of the following one-dimensional map:

$$z_{n+1} = F_a(z_n + I_n) - kF_b(z_n + I'_n). \tag{6.2}$$

Without loss of generality, we can take $k = 1$. In the following account we will be considering only time-invariant external stimuli, so that, for our purposes:

$$I_n = I'_n = I.$$

6.2.1 Analysis of response to constant magnitude external stimulus

We shall now consider how the dyanmics of the excitatory-inhibitory neural pair changes in response to external stimulus. Let us consider the isolated neural pair, whose time evolution is given by Eqn. (6.2). On replacing the expression of the transfer function from (2.9), we get

$$z_{n+1} = \exp(-b(z_n + I)) - \exp(-a(z_n + I)). \tag{6.3}$$

Now,

$$z_{n+1} = z_n = z^*$$

for a fixed point. It is stable if

$$\frac{dz_{n+1}}{dz_n} \geq -1,$$

i.e.,

$$(a - b)\exp(-a(z_n + I)) - bz_{n+1} \geq -1.$$

Therefore, for the fixed point to be marginally stable (i.e. $\frac{dz_{n+1}}{dz_n} = -1$), it must satisfy the following condition:

$$(a-b)\exp(-a(z^* + I_c)) = bz^* - 1, \tag{6.4}$$

where, I_c is the critical external stimulus for which z^* just attains stability. Let us define a new variable, α, as

$$\alpha = \frac{bz^* - 1}{a - b}. \tag{6.5}$$

Therefore, from (6.4), we get

$$\exp(-a(z^* + I_c)) = \alpha. \tag{6.6}$$

Also from (6.5),

$$z^* = \frac{1}{b} + \frac{1-\mu}{\mu}\alpha \tag{6.7}$$

where $\mu = b/a$. Now, from (6.3), a fixed point can be expressed as

$$z^* = -\exp(-a(z^* + I_c)) + \exp(-b(z^* + I_c)).$$

Therefore, from (6.6) and (6.7), the above expression can be written as

$$\alpha^\mu - \alpha = \frac{1}{b} + (\frac{1-\mu}{\mu})\alpha. \tag{6.8}$$

By simple algebraic manipulation, we get

$$\alpha = \frac{1}{b^{1/\mu}}(1 + a\alpha)^{1/\mu}. \tag{6.9}$$

Assuming $a\alpha << 1$, we need to consider only the first order terms in α in the right hand side, so that

$$b^{1/\mu}\alpha = 1 + \frac{a\alpha}{\mu}, \tag{6.10}$$

which gives the following expression for α:

$$\alpha = \frac{1}{b^{1/\mu} - \frac{a}{\mu}}. \tag{6.11}$$

For a real solution of z^* to exist, we must have $bz^* - 1 > 0$, since, otherwise, z^* will have an imaginary component (from (6.14)). In other words, $\alpha > 0$ (from (6.4)). Therefore, from (6.11), we must have

$$a < \mu b^{1/\mu}. \tag{6.12}$$

Since $b = \mu a$, we get

$$a > \mu^{\frac{\mu+1}{\mu-1}}. \tag{6.13}$$

For example, if $\mu = 0.5$, then $a > 8$, for z^* to be real. From (6.4) we get

$$\exp(aI_c) = \frac{a-b}{bz^* - 1} \exp(-az^*). \tag{6.14}$$

Taking logarithms on both sides, we have,

$$I_c = -z^* - \frac{1}{a} \log(\alpha).$$

Therefore replacing z^* from (6.7),

$$I_c = (\frac{\mu - 1}{\mu})\alpha - \frac{1}{\mu a} - \frac{1}{a} \log(\alpha). \tag{6.15}$$

The equation (6.15), together with (6.11), provides the critical value of the external stimulus, which leads the oscillatory neuron pair to a fixed stable state, subject to the restriction (6.13).

This expression can be further simplified. From (6.9), we can write

$$\mu \log(\alpha) = -\log(b) + \log(1 + a\alpha).$$

As before, assuming $a\alpha << 1$, we need to consider only the first order terms in α in the right hand side of the logarithmic expansion, which gives us

$$\log(\alpha) = \frac{a\alpha}{\mu} - \frac{1}{\mu} \log(b). \tag{6.16}$$

From (6.11), (6.15), and (6.16), the critical magnitude of the external stimulus is given as

$$I_c = \frac{1 - \frac{2}{\mu}}{(\mu a)^{1/\mu} - \frac{a}{\mu}} + \frac{1}{\mu a} \log(\frac{\mu a}{e}), \tag{6.17}$$

where $e = \exp(1)$. Fig. 6.1 shows the a vs. I_c curves for different values of μ, viz. μ=0.1, 0.25 and 0.5.

To make the network segment regions of different intensities ($I_1 < I_2$, say), one can fix μ and choose a suitable a, such that $I_1 < I_c < I_2$. So elements, which receive input of intensity I_1, will undergo oscillatory behavior, while elements receiving input of intensity I_2, will go to a fixed-point solution. Notice that, the curves obtained from Fig. 6.1 gives two values of a for the same I_c. This gives rise to an operational question: given a certain I_c, which value of a is more appropriate. Notice that the region of the a vs I_c curve, to the left of the maxima, has a very high gradient. This implies that, in the presence of wide variation in the possible value of I_c, choice of a from this region, will show very small variation - i.e., the system performance will be robust with respect to uncertainty in the determination of the appropriate value of I_c. This is possible in the case of any gray image, with a bimodal intensity distribution, having a long, almost uniform valley in between.

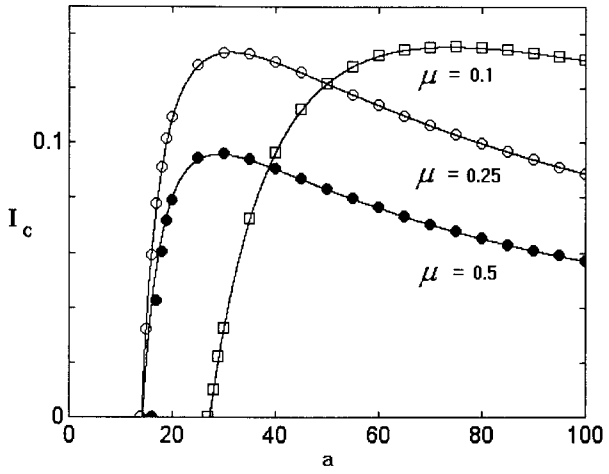


Figure 6.1: Critical magnitude (I_c) of the external stimulus, at which transition from periodic to fixed point behavior occurs. The circles (filled and blank) and squares represent the values obtained exactly through numerical procedures for $b/a = \mu =$ 0.5, 0.25 and 0.1, respectively. The curves indicate the theoretically predicted values.

On the other hand, the region of the curve to the right of the maxima has a very low gradient (almost approaching zero for high values of a). This implies structural stability in network performance, as wide variation in choice of a will give almost identical results. So, choice of a from this region is going to make the network performance stable against parametric variations. As both robustness against uncertain input, and, stability against parametric variations, are highly desirable properties in network computation, a trade-off seems to be involved here. The nature of the task in hand is going to be the determining factor of which value of a we should choose for a specific I_c.

6.2.2 The two-dimensional network

The introduction of spatial interactions over a local neighborhood in the above model produces some improvement in the segmentation performance. We have considered discrete approximations of circular neighborhoods [22] of radii r_{ex}, r_{in} $(r = 1, 2)$ in our simulations (Fig. 6.2).

There is an important feature to consider when choosing the neighborhoods of the excitatory and inhibitory neurons. Unless $r_{ex} < r_{in}$, the network activity becomes unstable owing to the unbounded increase in the activity of the excitatory elements. This is shown by looking at the averaged activity of the network $< z >_n = \frac{1}{N}\Sigma_{i=1}^{N} z_n(i)$, where $z(i)$ indicates the i-th excitatory-inhibitory neural pair (Fig. 6.3). For $r_{ex} \geq r_{in}$, $< z >$ shows an oscillatory behavior whose amplitude increases with n (Fig. 6.3 (b)). However, for the network to have stable behavior, the amplitude of the

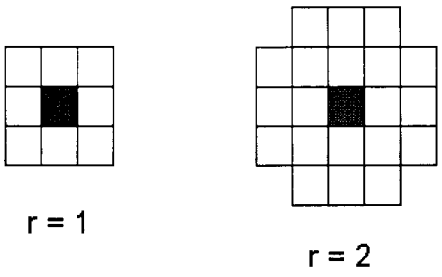


Figure 6.2: Discrete circular neighborhoods of radii $r = 1$ and $r = 2$.

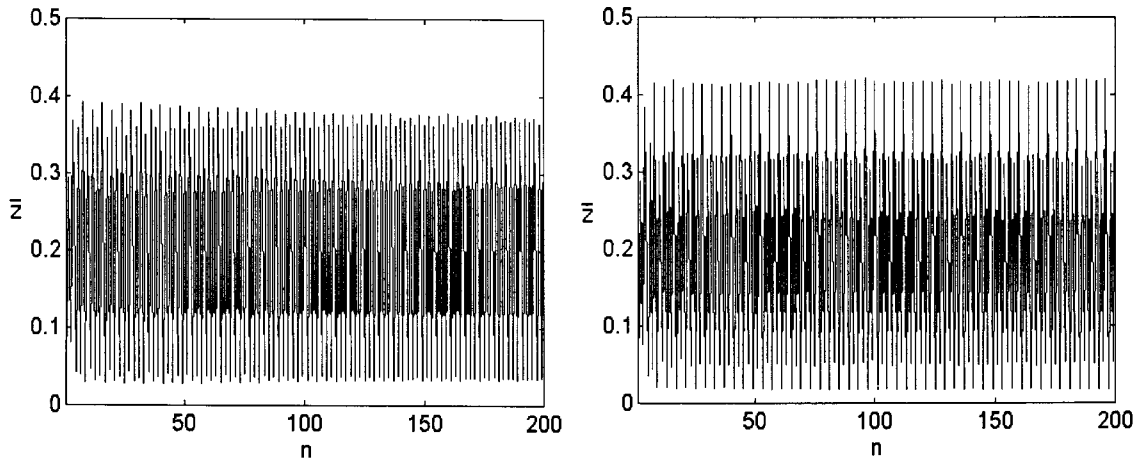


Figure 6.3: Average activity ($\overline{z}$) of a network of 100×100 elements, arranged in a two-dimensional plane, with coupling over a local neighborhood: (a) $r_{ex} = 1$, $r_{in} = 2$ and (b) $r_{ex} = 2$, $r_{in} = 2$.

oscillation should be constant in time. This is so for $r_{ex} < r_{in}$, as seen in Fig. 6.3 (a).

6.2.3 Simulation results

The network described above has been used to segment images, both synthetic and "real-life". The synthetic image chosen is that of a square of intensity I_2 (the "object") against a background of intensity I_1 ($I_1 < I_2$). Uniform noise of intensity ϵ is added to this image. The signal-to-noise ratio (SNR) is defined as the ratio of the range of gray levels in the original image to the range of noise added (given by ϵ). The image is then presented to the network, which is made to undergo 200 - 300 iterations. Afterwards, the elements which remain unchanged over successive iterations

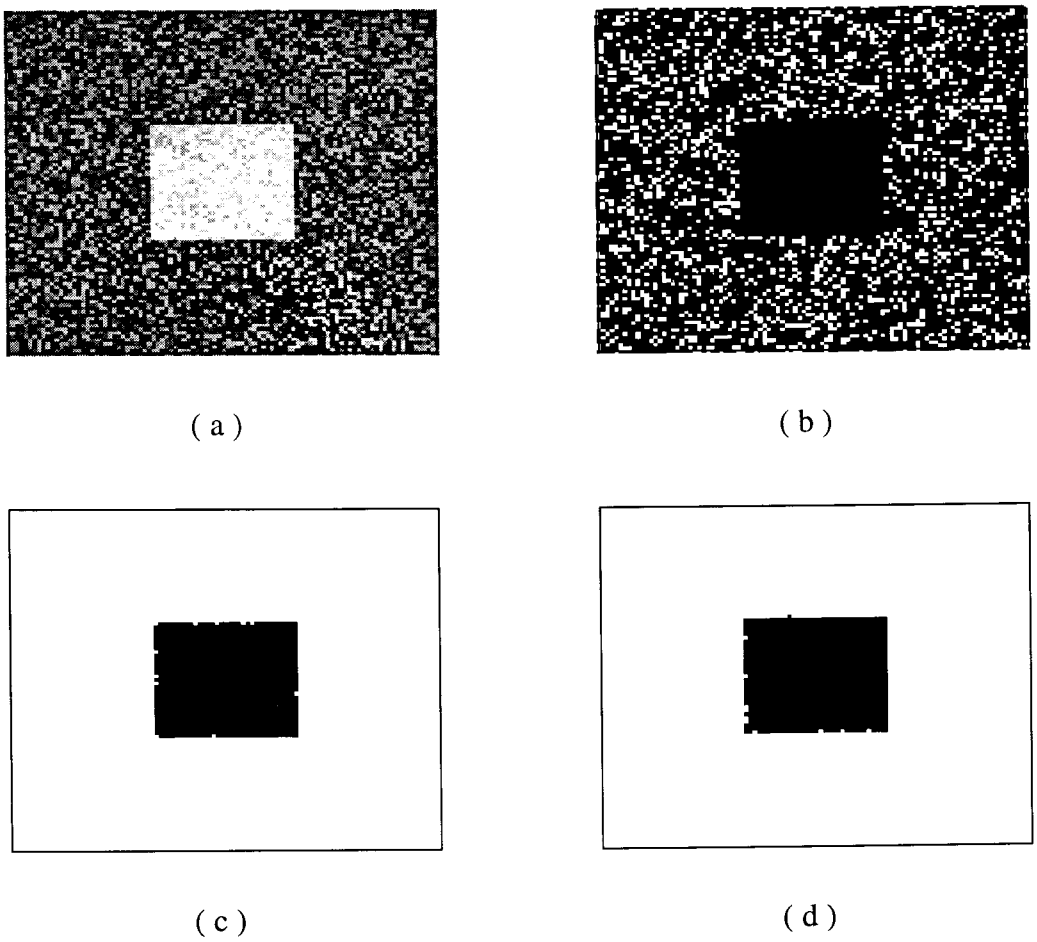


Figure 6.4: Results of implementing the proposed segmentation method on noisy synthetic image: (a) original image, (b) output by the uncoupled network, (c) output by the coupled network ($r_{ex} = 1, r_{in} = 2$), and (d) output by the coupled network ($r_{ex} = r_{in} = 2$), after 200 iterations (a=20, b/a=0.25 and threshold th=0.02).

(within a tolerance value, th) are labeled as the "object", the remaining being labeled the "background". For SNR = 1, the results of segmentation are shown in Fig. 6.4. Results for $r_{ex} = 1, r_{in} = 2$ and $r_{ex} = r_{in} = 2$ are shown. The two architectures show very similar segmentation results, at least upto the iterations considered here, although the latter is unstable (as discussed in the previous section). Excepting for the boundary of the "object", which is somewhat broken, the rest of the image has been assigned to the two different classes quite accurately.

We have also considered the 5-bit gray level "Lincoln" image as an example of a "real-life" picture. A suitable I_c has been estimated by looking at the histogram of the gray-level values, and taking the trough between two dominating peaks as the required value. Following the same procedure, as in the synthetic image, we have segmented the image. The results are shown in Fig. 6.5. Most of the image has been labeled accurately, except for a few regions (e.g., near the neck).

Note that, we have considered a single value of a (and hence I_c) for the entire image. This is akin to "global thresholding". By implementing local thresholding and choice of a on the basis of local neighborhood information, the performance of the network can be improved.

6.3 The retinal processing model

We next proceed to investigate image enhancement through adaptive smoothing in an excitatory-inhibitory network, followed by edge extraction from the processed image. The model used for this purpose has been inspired by the neural architecture of the retina.

6.3.1 Structure of the retina

The retina is where the physical image of a visual scene is converted to a neural representation, on being projected onto a receptor array. It is a thin sheet of neural tissue lining the rear hemisphere of the eyeball, being a projection of the brain itself. The basic design of a retina involves three interconnected layers and five broad classes of neurons: photoreceptor, bipolar, horizontal, amacrine and ganglion cells (Fig. 6.6). In all vertebrate retinas, the transformation from optical to neural image involves three stages [185]:

- phototransduction by a layer of receptor neurons,
- transmission of their signals by chemical synapses to a layer of bipolar neurons, and,
- transmission of these signals by chemical synapses to output neurons (ganglion cells).

(a)

(b)

(c)

(d)

Figure 6.5: Results of implementing the proposed segmentation method on "Lincoln" image: (a) original image, (b) output by the uncoupled network, (c) output by the coupled network ($r_{ex} = 1, r_{in} = 2$), and (d) output by the coupled network ($r_{ex} = r_{in} = 2$), after 300 iterations (a=30, b/a=0.25 and threshold th=0.02).

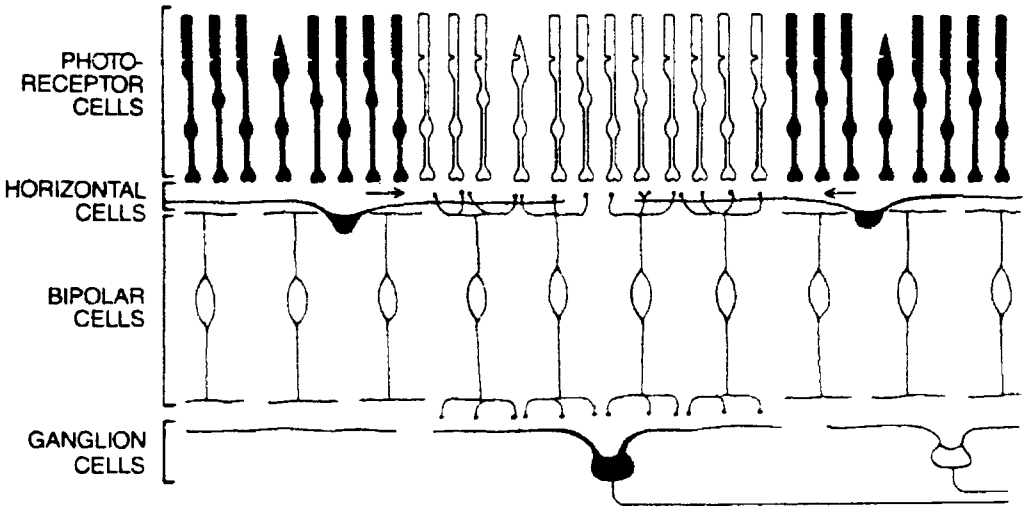


Figure 6.6: Structural organization of the retina. (Adapted from R. H. Masland, "The functional architecture of the retina", *Sci. Am.* **255**(6), 102-111 (1986).)

Synaptic connections in the retina are confined to two distinct zones: the outer plexiform layer (OPL) and the inner plexiform layer (IPL). The OPL contains the axon terminals of rods and cones, the dendritic processes of bipolar cells and the processes of horizontal cells. The IPL comprises the axon terminals of bipolar cells, the processes of amacrine cells and the dendrites of ganglion cells.

The two structurally similar systems of lateral interactions, at the OPL and the IPL, mediated by the horizontal and amacrine cells, respectively, modulate the signal transmitted along the input-output pathway. These interactions, which are inhibitory in nature [203], play an important role in retinal response, as they implement the process of *lateral inhibition.* This process is observed in a large variety of neural systems, whereby the stimulation of a neuron leads to the inhibition of other neurons lying in the surrounding area. Lateral inhibition in sensory systems has important functional significance, since, it provides for an enhancement of the acuity by increasing the contrasts of the pattern of the message carried to the brain [76, 137]. The modulation of the local input-output pathways due to the lateral interneurons (on the basis of an average signal received over a broad area) is used by the retina to adjust its operating characteristics to prevailing light conditions [200, 201, 178]. This ensures the generation of a high-contrast visual image. In this work, we have focussed on the interactions at the OPL, responsible for producing a high contrast response to visual stimuli.

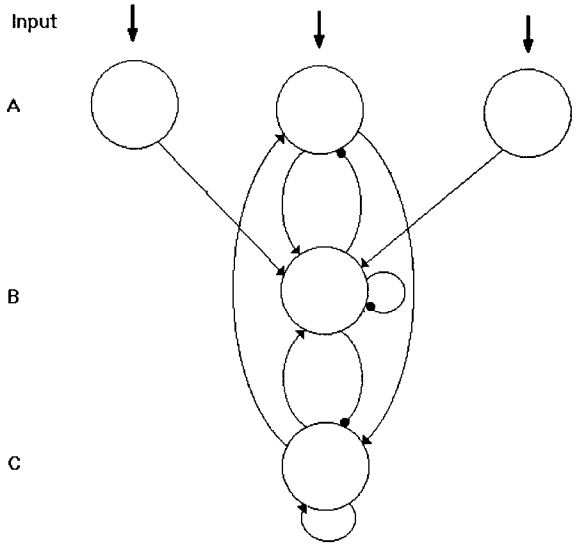


Figure 6.7: An excitatory-inhibitory neural network model for retinal information processing. Layers A, B and C correspond to the photoreceptor, horizontal cell and bipolar cell layers of the retina. Arrows indicate excitatory connections, while circles represent inhibitory synaptic couplings.

6.3.2 The model

A model comprising three layers of excitatory, inhibitory and excitatory elements (respectively), with anti-symmetric, sigmoid activation function, has been used to investigate the visual processing done in the outer plexiform layer of the retina. Figure 6.7 shows the arrangement of the network. Layers A and C consist of excitatory elements, while, layer B is composed of inhibitory elements. They correspond to the photoreceptor, horizontal cell and bipolar cell layers of the retina, respectively. The dynamics of the network is governed by the following set of equations:

$$x_{n+1} = F_b(I), \tag{6.18}$$

$$y_{n+1} = F_b(x_n + z_n - y_n - \lambda(x_n + z_n - \Sigma_{j\in R} y_n^j)), \tag{6.19}$$

$$z_{n+1} = F_a(x_n + z_n - y_n - \lambda(x_n + z_n - \Sigma_{j\in R} y_n^j)). \tag{6.20}$$

Here R indicates the local neighborhood in layer B, over which lateral connections are considered. In our simulations for a two-dimensional network, R is the 3×3 neighborhood around a given element. The external input to the network is denoted by I, while, the parameter, λ, controls the relative weightage of the lateral interactions. The activation functions, F, are given by Eqn. (2.10).

Initially (i.e. at $n = 0$), the input I is taken to be the input image, while x, y, z are all set equal to zero. The input image is then withdrawn, so that for all subsequent iterations, I, and hence, x, is zero. Therefore, for $n > 0$, if we neglect the lateral

interactions (i.e., $\lambda = 0$), the dynamics of the network reduces to that of the one dimensional map:

$$\xi_{n+1} = F_a(\xi_n) - F_b(\xi_n),$$

where, $\xi = z - y$. This is the map representing the dynamics of a single excitatory-inhibitory neural pair, that has been analyzed in Chapter 2. The additional features of the network that shall be reported here are, therefore, purely the result of lateral interactions.

The effect of the lateral interaction term is to enable computation of the local gradient of the intensity values in an image. This is used to generate an output, which is then added to the existing image. The result is *adaptive smoothing* [1] of the image. Note that, λ necessarily has to be small, as otherwise, the network becomes unstable, and the activity becomes spatially uniform.

The mechanism by which adaptive smoothing is implemented in the present model, can be easily understood by looking at the map governing the dynamics of an excitatory-inhibitory pair with sigmoid activation functions (Fig. 2.11). In the present context, the input is the gradient of the image. For edge regions, where the gradient is high, the resultant output is low, implying that there will be little or no change in these regions. However, isolated discontinuities, which are not part of any edge, also give high gradients. As a result of the selective smoothing process, they will be maintained almost unchanged, degrading the performance of the method. This can be avoided by decreasing b (keeping a fixed), which increases the output value of the map for high values of the gradient. Therefore, there will be smoothing of points having a very high gradient value.

Let us now consider the smooth regions of the image, for which the gradient is low. The neural map, considered above, gives a very high output value for these regions. Any small inhomogeneity present in these regions is, therefore, rapidly smoothed out. As the peak value of the map is a function of the parameter, a (with b fixed), the extent of smoothing is governed by the magnitude of a - i.e., a desired degree of smoothing can be obtained by a suitable choice of a.

6.3.3 Simulation results

The proposed model has been used to implement adaptive smoothing in several gray-level images, an example being shown in Fig. 6.8 (a). The results of adaptive smoothing is shown for various parameter values in Fig. 6.8 (b – d). The process was carried out for 30 iterations for all the parameter values considered. As is clear from the results, increasing the parameter a results in enhanced smoothing of the image, while, increasing b has the effect of increasing the gradient of isolated discontinuities of the image. This is in accordance with the qualitative understanding of the network behavior outlined previously.

[1] Adaptive smoothing is a nonlinear filtering mechanism that can achieve edge-preserving smoothing of an image [151]

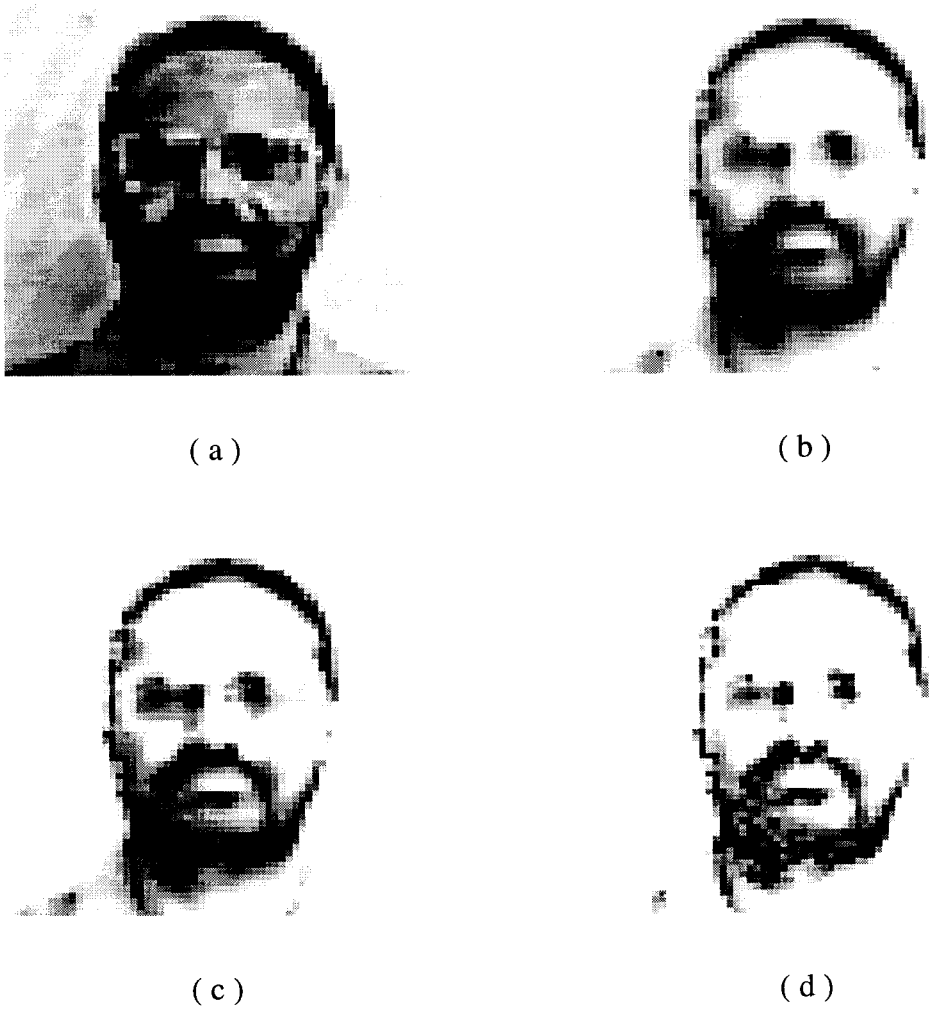
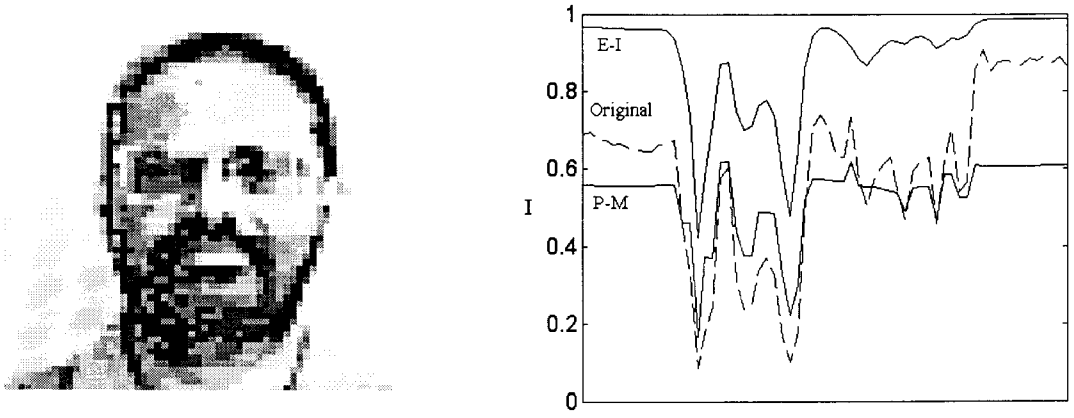


Figure 6.8: Results of implementing the proposed adaptive smoothing method on "Uma" image: (a) original image, (b) output by the model with a=20, b=5, (c) output by the model with a=10, b=5, and (d) output by the model with a=10, b=8, after 30 iterations (λ=0.01). It is evident that increasing a produces greater blurring of the final image, while increasing b leads to enhancement of isolated discontinuities.



(a)

(b)

Figure 6.9: Comparison with Perona-Malik method: (a) output by the Perona-Malik variable conductance diffusion method, and (b) the relative smoothing of a one-dimensional signal extracted from "Uma" image by considering the 32nd row. The broken line shows the original signal, while outputs by the Perona-Malik method (P-M) and the proposed model (E-I) ($a = 10$, $b = 5$, $\lambda = 0.01$) are shown with solid curves (the intensity values have been scaled to lie in (0,1)).

We have compared the proposed method with the variable conductance diffusion (VCD) method proposed by Perona and Malik [144] in Fig. 6.9. As can be seen clearly, isolated points having a high local gradient have been enhanced by the VCD method, leading to a speckled image. By extracting a one-dimensional signal from the image (viz. the 32nd row) we demonstrate the relative performance of the VCD method and the proposed method. While the former seems to enhance small local discontinuities (which may be only due to noise), the latter retains the broad features of the original signal, while averaging out smaller discontinuities, thus giving a smoother image.

The output of the adaptive smoothing process has been used for extraction of edges by using a conventional gradient thresholding technique. Note that, alternatively, an extra pair of excitatory-inhibitory layers can also be used for this purpose. This is because, as mentioned earlier, the interaction between excitatory and inhibitory neurons results in the gradient of the intensity values of an image to be obtained. In Fig 6.10, the edges obtained from the output image of the proposed method is compared with the results of applying the Canny operator method [29] on the original image. No post-processing technique, such as thinning or linking, has been applied. The performance of the two methods appear to be comparable, although the former method leads to a few broken edges.

(a) (b)

Figure 6.10: Performance in edge extraction: edges obtained from the "Uma" image on implementing (a) Canny's method and (b) the proposed model ($a = 10$, $b = 5$).

6.4 Discussion

The segmentation of images, through dynamical activity in neural network models, have been investigated previously by a number of researchers [68, 181, 183, 70, 194, 156, 195, 28, 73]. However, all these efforts had been confined to using synchronization and desynchronization of oscillatory activity among neural assemblies. We have instead, concentrated on using stimulus induced transitions from periodic to fixed point behavior, in order to segment images. As Malsburg [194] has indicated, the reason oscillatory synchronization has been studied so far, as a mean of segmenting sensory stimuli, is its relative ease of analysis. However, with the developments in nonlinear dynamics and chaos theory, we can approach the problem of sensory segmentation using more general dynamical behavior.

The adaptive smoothing method proposed in this chapter is similar to the variable conductance diffusion method discussed in the image processing literature [144, 131, 21]. Although the continuous space version of this method, expressed in terms of partial differential equations, is ill-posed [32, 202], the discretised version appears to be quite robust [199]. As the proposed model is based on a discrete lattice of elements, a relation with discrete scale space methods [111] may not be far-fetched. However, this is an area for future investigation.