SEGMENTATION PROCESSES IN VISUAL PERCEPTION:
A COOPERATIVE NEURAL MODEL[†]

Parvati Dev
Neurosciences Research Program
Boston, Massachusetts

COINS Technical Report 74C-5

(June 1974)

Abstract

Research on the system for recognition of object identity has concentrated on the neural detection of features such as edges, color and movement (Hubel and Wiesel, 1962; Hubel and Wiesel, 1968), with the assumption that subsequent recognition is based on the set of features detected (Barlow, 1972). Since the visual field, in general, contains many more than one object, at some stage prior to recognition of an object there must occur a separation into subsets of the set of features detected, such that each subset contains features corresponding to a single object or a meaningful group of objects. This process of segmentation has not

---

[†] This work was supported in part by NIH Grant No. 5 R01 NS09755-03 COM awarded to M. A. Arbib, Computer and Information Sciences, University of Massachusetts, Amherst, Mass., and by the Neurosciences Research Program, Jamaica Plain, Mass.

been studied neurophysiologically. Theoretically, it has been of importance in artificial intelligence systems where pattern recognition also involves segmenting out the optical data relevant to a single object from an image of a scene containing many objects (Guzman, 1967). However, the algorithmic and heuristic techniques developed for artificial intelligence systems do not explain how the activity of individual neurons in a neural network can generate the segmentation of a visual scene.

In this paper, a model is developed for the segmentation of a visual scene by excitatory and inhibitory interaction between feature-detecting neurons in a neural network. Analytic equations are developed for the neural mechanism of segmentation. The model is tested by computer simulation of the process of binocular depth perception in Julesz random-dot stereograms (Julesz, 1971), where the segmentation process generates the surfaces at different depths. Segmentation is performed on the basis of feature similarity, that is, a segment is defined as a region of the visual field where the same feature occurs at all points within that region. The network detects feature similarity by responding with high levels of neural activity in the corresponding region of the network.

SEGMENTATION PROCESSES IN VISUAL PERCEPTION:
A COOPERATIVE NEURAL MODEL

1. Introduction

Research on processes for recognition of object identity has concentrated on processes such as feature detection, pattern recognition and visual search. The process of visual field segmentation, an important prerequisite in visual information processing, is generally ignored. Neurophysiological data for more than a decade (Hubel and Wiesel, 1962) suggests that pattern analysis is based on the neural detection and synthesis of pattern elements, such as edges, angles and texture, known collectively as features. Since the visual field, in general, contains more than one object, at some stage prior to recognition of the object the set of features detected must be separated into subsets such that each subset of features corresponds to a single object or a meaningful group of objects. The process of generating appropriate feature subsets is termed here as 'segmentation'.

The process of segmentation has been of some theoretical interest in the field of artificial intelligence, where pattern recognition involves segmenting out the optical data relevant to a single object from an image of a scene containing many objects (Guzman, 1967). However, the algorithmic and heuristic techniques employed in artifical intelligence systems cannot explain how activity in networks of neurons can generate segmentation of a visual scene.

In the neurosciences, however, segmentation has elicited almost no interest, except for some early work by Gestalt psychologists on the related problem of figure-ground separation (Koffka, 1935). The current paradigm of neuronal action makes segmentation a particularly difficult problem to study. Neurophysiology, which is one of the main tools for the study of neural mechanisms of perception, is based on the assumption of localized connectivity. A neuron in one brain area projects only to neurons in a corresponding spatial location within another brain area. The result is that a neuron in the visual system, for example in the visual cortex, is assumed to receive input only from a localized region of the visual field, generally known as its receptive field. Segmentation, however, requires synthesis of information from many locations in the visual field so that a decision can be made as to which features in the input belong to a single subset.

Lateral interaction between neurons leads to modification of a neuron's receptive field (Blakemore and Tobin, 1972). This localized lateral interaction, however, also has a global effect, causing input at any location to eventually influence all neurons in that brain area. It is this global effect through local interaction that is explored here as a possible basis for the process of segmentation.

1.1. Definition of the Problem

The process of segmentation, that is, subdivision of the visual field, may be based on a number of criteria. A region of similar feature properties forms a 'primitive unity' (Hebb, 1949), with a high degree of probability that the region corresponds to a single object. For example, a brown area is eventually recognized as a table, or a rough-textured area defines the extent of a wall. Certain patterns of features, by enclosing an area, may define a segment, for example, four lines appropriately arranged may define a square-shaped segment.

Segmentation may also be a feedback process. For example, computation on a group of features can define what other features should be included in the group. This is seen in the perception of hidden figures. Once a part of the figure is perceived, the rest of the figure stands out from a meaningless background of lines and angles.

In artificial intelligence programs for object recognition, segmentation is carried out by iteration of a series of processes. Computation on a group of features causes generation of a hypothesis. This specifies other features that should belong to the group. If these features are found, the hypothesis is satisfied, and computation is carried out on the remaining features. Otherwise, a new hypothesis is generated, and a new attempt at segmentation occurs (Minsky and Papert, 1972).

Thus, segmentation criteria may vary from the very simple to the very complex, and, in actual perception, segmentation probably occurs repeatedly, at a number of different stages. Here we examine, and develop a model for segmentation on the basis of feature similarity, the most basic process of segmentation.

Sections 2, 3 and 4 develop and analyze the model which is based on neural networks operating through parallel, distributed computation. Section 5 is a specific application of the model to the segmentation of random-dot stereograms (Julesz, 1971), where each segment corresponds to a surface at a different depth from the observer. While other models have been developed for the process of stereopsis (one of the processes by which depth is perceived), the model presented here is neurally oriented in contrast to the more abstract approaches used elsewhere. It also provides a neural basis for the cooperative, dipole model of stereopsis presented by Julesz (1971).

2. The Biological Basis of the Model

The model of the segmentation process to be developed in this chapter uses feature similarity as the criterion for segmentation. The aim is to develop a neural network capable of detecting features and associating them appropriately so that the existence of regions of similar feature input can be perceived. In order to define the characteristics and connectivity of such network, available psychophysical and neurophysiological data are explored.

In a cluttered visual field, with objects of different sizes, shapes and colors, the search for a target object is reduced if cues are given. These cues structure (or segment) the visual field. For example, search for an orange object causes perception of a pattern of orange objects on a background of other colored objects (Williams, 1966). Cues such as color cause much more efficient segmentation than cues such as shape, suggesting that segmentation on the basis of feature similarity must use simple features, where color is considered a simpler feature than shape (for example, a square or triangle). Beck (1967) also comes to the conclusion that simple features, such as brightness or orientation, provide better "perceptual grouping" than more complex features such as shapes.

Through neurophysiological observations it is known that features that are detected more peripherally in visual processing are color, movement, texture (spatial frequencies), orientation and, perhaps, depth. Detection of specific shapes appears to be the result of further processing (Hubel and Wiesel, 1962). We will assume that the features that are detected peripherally are the simplest and that, when a cluster of similar features occurs in a region, this provides the basis for the formation of a segment.

In figure-ground situations, the threshold for discrimination of a feature is lower in the figure region than in the ground, and accurate discrimination occurs earlier in the figure area. This may be the effect of figural facilitation, ground inhibition, or a combination of both types of processes (Weitzman, 1963). Further, when a cluttered field is segmented on the basis of a color, a pattern in that color stands out from the background, also suggesting some figural facilitation or ground inhibition effects.

We suggest that the neuronal correlate of figural facilitation or ground inhibition may be, very simply, a higher level of neural activity in the figure area and a correspondingly lower level of activity in the ground.

In the corresponding case of segmentation based on feature similarity, we suggest that the occurrence of a cluster of similar features generates a high level of activity in the corresponding feature-detecting neurons. Random activation of feature detectors results in low activity levels in these neurons.

In the following section, a model of interacting arrays of feature detectors is described. It is shown qualitatively that certain simple patterns of excitatory and inhibitory connectivity between the neurons suffice to generate the desired high-activity regions corresponding to segments. The model is embedded into the overall pattern recognition process.

A mathematical analysis of the dynamics of these interacting arrays is developed in Sec. 4.

3. Formulation of the Model

In the previous section, it was suggested that segmentation on the basis of feature similarity is the earliest segmentation in the visual system. The hypothesis is made that the neural basis for such segmentation is simple -- cortical regions receiving similar feature input show increased levels of neural activity because of excitatory interaction between similar feature detectors. The region of increased activity defines the extent of the segment.

In section 3.1 we show in a qualitative manner how interacting neural arrays, through feature extraction, can lead to generation of high activity regions corresponding to segments. The model of the segmentation process is then embedded in the overall pattern recognition process (in Sec. 3.2), indicating how regions of high activity can restrict the pattern recognition process to the region defined as the segment.

3.1. Neural Representation of the Segmentation Stage

Figure 2.1 indicates the neural connectivity that causes regions of similar feature input to increase in activity. The process is one of positive feedback (or cooperative facilitation). The neurons of this stage correspond to the simple cells of Hubel and Wiesel (1962), that

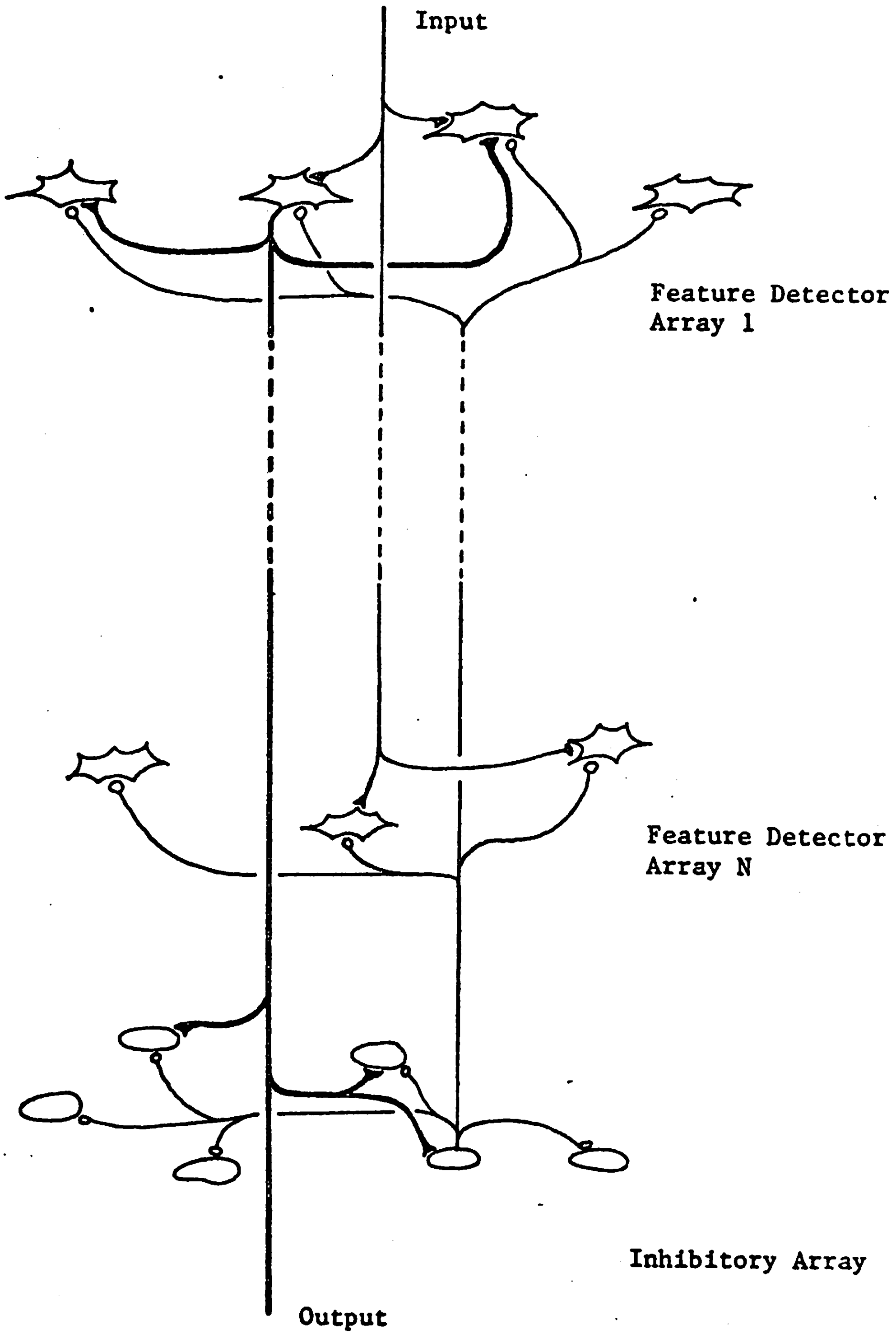


Fig. 1. Neural Connectivity of the Segmentation Stage

detect features such as edges, orientation, velocity and direction of movement. Neurons that detect similar features are assumed to excite one another. For example, all neurons that respond selectively to the color red interact by exciting each other.

Therefore, as a first approximation, this stage can be conceived of as a set of neural nets. Within a net, the neurons detect similar features and interact through excitation. The nets do not interact with each other because neurons detecting dissimilar features do not interact. In the mathematical analysis developed in Sec. 4, these non-interacting nets of feature detectors are modelled as separate arrays, with each array containing only neurons detecting similar features.

A population of inhibitory interneurons, with wide dendritic and axonal spread for each neuron, is interspersed among these feature detectors. (In figure 1, they are shown in a separate array for simplicity, to aid in showing the flow of information). These inhibitory interneurons receive input from the activated feature detectors and inhibit all the feature detectors in the corresponding location. The net effect is that activation of a set of feature detectors reduces the activity of all the other feature detectors in that region. However, because of the excitatory interaction between similar feature detectors (or positive feedback) a region of high activity tends to expand its boundaries. If the input to neighbouring similar feature detectors is subliminal (that is, input exists but is not sufficient to activate the neuron), the excitatory influence of the nearby high-activity region causes these neurons to become activated too.

Thus, the effect of the connectivity diagrammed in Fig. 1 is twofold:

a) a region of near-uniform high activity is developed, corresponding to the existence of a segment characterized by a specific feature, even if the input to this region is nonuniform.

b) spurious activation of other feature-detectors in that region is suppressed.

A detailed mathematical analysis and simulation is presented in Sec. 4.

3.2. Embedding the Segmentation Stage in a Pattern Recognition Process

The segmentation stage described in Sec. 3.1 can generate regions of high activity corresponding to the occurrence of a cluster of similar features in the visual input. However, in order to eventually recognize an object, this region of increased activity must be used by a later stage which carries out a pattern recognition procedure on the region of visual input defined by the increased activity. The actual process of object recognition is still not known but current theories are based on the assumption that the detection of complex combinations of features must underlie any such process (Barlow, 1972). In this section, we show how regions of increased neural activity can guide information flow so as to restrict pattern recognition procedures to definite regions of the visual input.

While the segment is defined by a specific feature, such as its color or texture, it generally contains other features as well, such as the lines defining its edges, its depth from the observer, and any details on its surface. All these features contribute to the process of object recognition.

For the process of segmentation to be completed, computation must be restricted to the features lying within the region of high neural activity, and features lying outside this region must be excluded. Hence the occurrence of a region of high activity must in some way separate the computation on the features lying within and outside this region.

Figure 2 outlines one possible scheme that achieves spatial separation of computation on these two regions. The segmentation stage projects to two other stages. Stage 1 has high threshold neurons and can only be activated in the region corresponding to high activity in the segmentation stage. Object recognition is carried out on the input to stage 1. The segmentation stage also projects to the lower threshold stage 2. The region corresponding to high activity in the segmentation stage receives inhibitory input from stage 1. Thus stage 2 computes only on input from outside the segment region.

Separation of computation may be achieved by other means. Temporal separation, for example, may occur if information from regions of high activity progresses through successive stages with shorter latencies than does information from regions of low activity.

Stage 2, which receives input from all areas outside the segment region, may also carry out a recognition procedure, and may be the location at which the perception of 'ground' occurs (while perception of the 'figure' occurs in stage 1). A characteristic of the 'ground' is that, even though it may be interrupted by several objects,

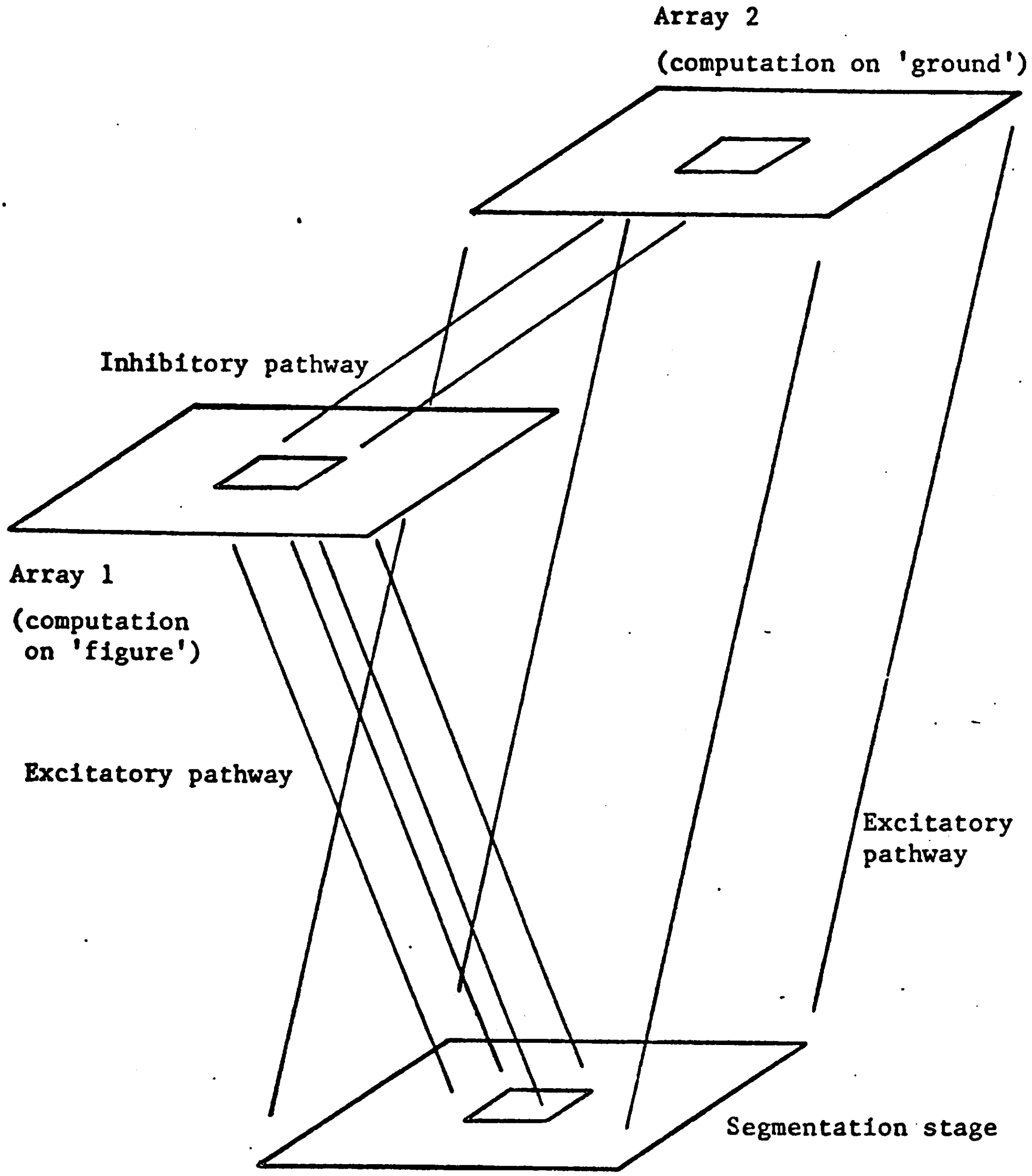


Fig. 2. Spatial Separation of Computation on Figure and Ground

it is still recognized as one entity. Input to stage 2 is interrupted in the region corresponding to the segment which projects to stage 1, and any recognition procedure must span this gap in the input. It is possible that a similar mechanism applies to the 'completion' that is seen across small scotoma. The blank in the visual field that should be produced by the scotoma is often not perceived, and the ground is perceived as continuous across that region. However, any details lying within that region are lost.

In general, there is more than one segment present in the visual field. These give rise to high-activity regions at different locations in the different feature-detector arrays. Again, only the region with the highest activity projects to stage 1 and is recognized. For the pattern recognition procedure to operate on the other segments as well, a feedback inhibitory process must exist which inhibits those regions in the segmentation stage which have projected to stage 1.

4. Mathematical Analysis of the Segmentation Stage

Figure 1 outlined the neural connectivity of the segmentation array. Neurons detecting similar features have excitatory interconnections. A population of inhibitory neurons is interspersed among the feature detectors, receiving input from all feature detectors and inhibiting all feature detectors non-specifically.

Input to the segmentation array is a spatial pattern of light intensity. Feature detectors in the segmentation array extract features from this intensity pattern. Corresponding to each small region of the visual field, a number of features are extracted by the array. Some of these features are common to a number of adjacent regions, and form the basis for segmentation.

The following section presents a mathematical analysis of the dynamics of the segmentation array, showing that:

a) when a number of different features are extracted from a region of the visual field, activated neurons detecting similar features tend to increase in activity because of their excitatory interconnections (positive feedback).

b) activity of the other feature detectors in that region is reduced because of the non-specific inhibition and the lack of other activated neurons detecting similar features.

Thus segmentation is achieved by a process of facilitory concensus (between similar feature detectors) and inhibitory suppression (against other, scattered feature detectors).

The interacting feature detectors are formally defined as N disjoint subpopulations of excitatory neurons and one subpopulation of inhibitory neurons (Fig. 3). The excitatory subpopulations do not interact with each other. However, within each of these subpopulations, the neuronal interaction is excitatory. The N excitatory subpopulations also excite the single inhibitory subpopulation. Each inhibitory neuron inhibits other neurons in its topographic area, including other inhibitory neurons.

Each excitatory subpopulation consists of neurons which detect identical features. Thus, there is excitatory interaction between neurons detecting identical features but no interaction between neurons

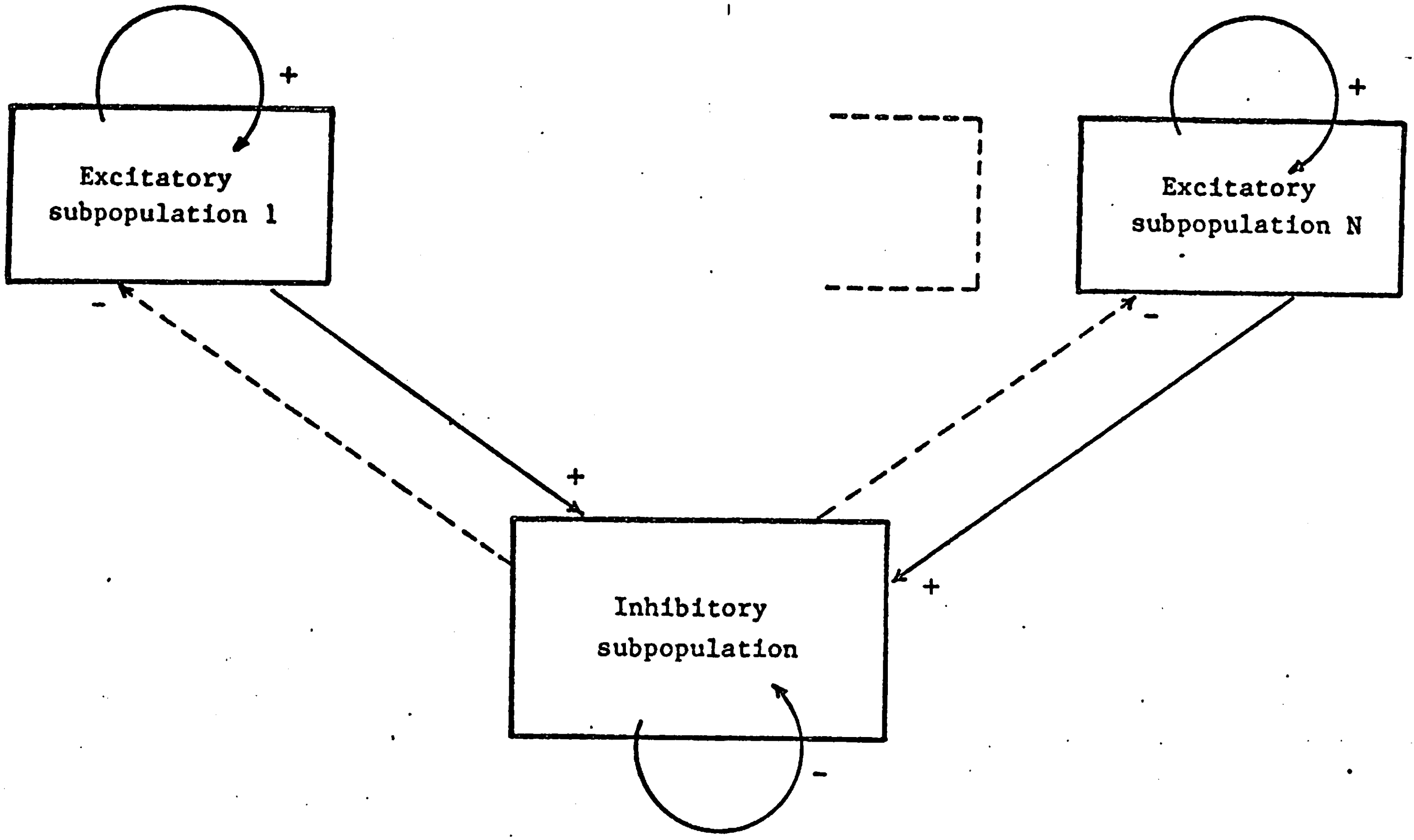


Fig. 3. Illustrating the Connectivity between N disjoint Subpopulations of Excitatory Neurons and one Subpopulation of Inhibitory Neurons

detecting different features, except through the population of inhibitory neurons.

4.1. Dynamics of Population Interaction

In this section, the interaction of N excitatory subpopulations and one inhibitory subpopulation is investigated. The parameter of spatial location is ignored. Hence, it is the behavior of ensembles that is being considered. In Sec. 4.2, we will return to the investigation of spatial arrays.

Each ensemble is composed of discrete units, namely neurons. The quantity of interest is the number of neurons activated in each ensemble. This is termed E_j, where $j=1,2,\ldots N$ is the index denoting the subpopulation or ensemble of interest. Equation 2.1 shows how this quantity, E_j, is determined by the number of activated neurons in each ensemble at the preceding instant. For the purpose of analysis, we also choose to examine the system at discrete points in time.

Just as E_j denotes the number of active neurons in the jth excitatory ensemble, I denotes the same quantity in the inhibitory subpopulation. Since the number of active neurons cannot fall below zero, E_j and I are never negative. P_j denotes the feature input to the jth ensemble, and Q denotes the input to the inhibitory ensemble. Q may be considered a bias input related to the level of arousal.

The degree of interaction between the ensembles is specified by the weighting factors a_j, b_j, c_j, d_j, g and h, as shown in eqn. 2.1. In Sec. 2.4.2, the definition of these factors is extended to make them spatial weighting factors.

Given the number of active neurons at time t, activation at time (t+1) is given by

$$E_j(t+1) = a_j.E_j(t) - b_j.I(t) + c_j.P_j(t) \tag{1}$$

where $j = 1,2,...N$, and

$$I(t+1) = \sum_{j=1}^{N} d_j.E_j(t) - g.I(t) + h.Q(t) \ . \tag{2}$$

At equilibrium, (that is, for $E_j(t+1) = E_j(t)$, $I(t+1) = I(t)$, for all t), for constant P_j and Q, a unique solution exists for this set of simultaneous equations. The exact solutions for the various E_j and I are functions of P_j and Q.

The behaviour of these interacting populations is examined in greater detail for the system with only two excitatory subpopulations, that is, for $N = 2$.

For $N = 2$, eqns. 1 and 2 reduce to the following set:

$$E_1(t+1) = a_1.E_1(t) + b_1.I(t) + c\ .P\ (t) \tag{3}$$

$$E_2(t+1) = a_2.E_2(t) + b_2.I(t) + c_2.P_2(t) \tag{4}$$

$$I\ (t+1) = d_1.E_1(t) + d_2.E_2(t) - g.I(t) + h.Q(t) \tag{5}$$

Assume now that

$$a_1 = a_2 \doteq a, \tag{6}$$

$$b_1 = b_2 \doteq b, \text{ and} \tag{7}$$

$$d_1 = d_2 \doteq d. \tag{8}$$

This implies that the excitatory subpopulations have identical interaction parameters. With these conditions, the situation at equilibrium is

$$E_1 = a.E_1 - b.I + c.P_1 \,, \tag{9}$$

$$E_2 = a.E_2 - b.I + c.P_2 \,, \text{ and} \tag{10}$$

$$I = d.(E_1 + E_2) - g.I + h.Q. \tag{11}$$

Solving for E_1, E_2 and I gives

$$E_1 = \{ \frac{c}{1 - a} - \frac{bdc}{\Delta} \}.P_1 - \frac{bdc}{\Delta}.P_2 - \frac{bh(1 - a)}{\Delta}.Q \,, \tag{12}$$

$$E_2 = \{ \frac{c}{1 - a} - \frac{bdc}{\Delta} \}.P_2 - \frac{bdc}{\Delta}.P_1 - \frac{bh(1 - a)}{\Delta}.Q \,, \tag{13}$$

and

$$I = \frac{cd(1 - a)}{\Delta}.(P_1 + P_2) + \frac{h(1 - a)^2}{\Delta}.Q \,, \tag{14}$$

$$\Delta \doteq (1 - a).\{(1 - a)(1 + g) + 2bd\} \,. \tag{15}$$

From these results, the following conclusions can be drawn:

a) If $P_1 = P_2$, then $E_1 = E_2$, because the populations have identical parameters.

b) Increase in P_1 causes an increase in E_1 and a decrease in E_2, when $\frac{c}{1 - a} > \frac{bdc}{\Delta}$. (It is easily shown that this condition is satisfied when all the coefficients are positive, and $a < 1$.)

c) The ordering of E_1 and E_2 corresponds to the ordering of P_1 and P_2, that is, if $P_1 > P_2$, then $E_1 > E_2$, for the same conditions as above.

d) For the isolated excitatory population, given by

$$E(t+1) = a.E(t) + c.P(t), \tag{16}$$

the equilibrium value of E is

$$E = \frac{c}{1 - a}.P. \tag{17}$$

Thus, interaction with other excitatory populations, via an inhibitory population, reduces the equilibrium value of E_1 and E_2.

e) A change in input, δP, causes the changes

$$\delta E = \{ \frac{c}{1 - a} - \frac{bdc}{\Delta} \}.\delta P \,, \text{ and} \tag{18}$$

$$\delta E_1 = -\frac{bdc}{\Delta}.\delta P_1 \ . \qquad (19)$$

Therefore, $\delta E_1 - \delta E_2 = \frac{c}{1-a}.\delta P_1$. (20)

For the isolated population, a change δP causes the change

$$\delta E = \frac{c}{1-a}.\delta P \ .$$

Thus, though interaction lowers the absolute value of activity, the amount of change in activity for change in input is not altered. Hence the relative amount of change in activity is increased.

However, we cannot say yet whether this equilibrium point is stable or not. The stability of the system can be investigated by determination of its various eigenvalues.

[Given a system characterised by the equation

$$\underline{x}(t+1) = \underline{A}.\underline{x}(t) + \underline{B}.\underline{u}(t) \qquad (21)$$

where $\underline{A}$ and $\underline{B}$ are matrices, $\underline{x}$ and $\underline{u}$ are vectors, the stability of the system can be expressed in terms of its eigenvalues. For the autonomous system, that is, $\underline{u}(t) = 0$ for all t,

$$\underline{x}(t+1) = \underline{A}.\underline{x}(t) \ . \qquad (22)$$

The system is stable if the eigenvalues, λ_i, of $\underline{A}$, satisfy the following condition:

$|\lambda_i| \leq 1$, for all of the eigenvalues.

For the system with n excitatory subpopulations,

$$\underline{A} = \begin{bmatrix} a_1 & & & & -b_1 \\ & a_2 & & & -b_2 \\ & & \ddots & & \vdots \\ & & & a_n & -b_n \\ d_1 & d_2 & \cdots & d_n & -g \end{bmatrix} \qquad (23)$$

The characteristic equation ($\underline{A} - \lambda\underline{I}$) = 0 can then be solved for its eigenvalues.]

In solving for the eigenvalues of the system, we first obtain the characteristic equation

$$\lambda^3 + \lambda^2(-a_1-a_2+g) + \lambda(-a_2g+b_2d_2+a_1a_2-a_1g-b_1d_1)$$

$$+ (a_1a_2g-a_1b_2d_2+a_2b_1d_1) = 0 \qquad (24)$$

For

$$a_1 = a_2 \doteq a \ , \qquad (6)$$

$$b_1 = b_2 \doteq b \ , \qquad (7)$$

$$d_1 = d_2 \doteq d \ , \qquad (8)$$

the eigenvalues are found to be simply

a, a, and -g.

Both a and g are positive as assumed before. Hence the system is stable as long as

$$a \leq 1 \ . \qquad (25)$$

This implies that, if the jth population receives no external inputs from other neural populations, then $E_j(t+1) \leq E(t)$ for all t. If $a < 1$, the activity of the population decreases to zero.

It is interesting that, in spite of all the interaction between the populations, stability depends only on the internal dynamics of each population.

A computer simulation of the system with two excitatory populations was carried out. Figures 4 and 5 show the dynamics for a step increase in input P_1. In Fig. 4, these populations are isolated and do not interact with each other, Increase of the

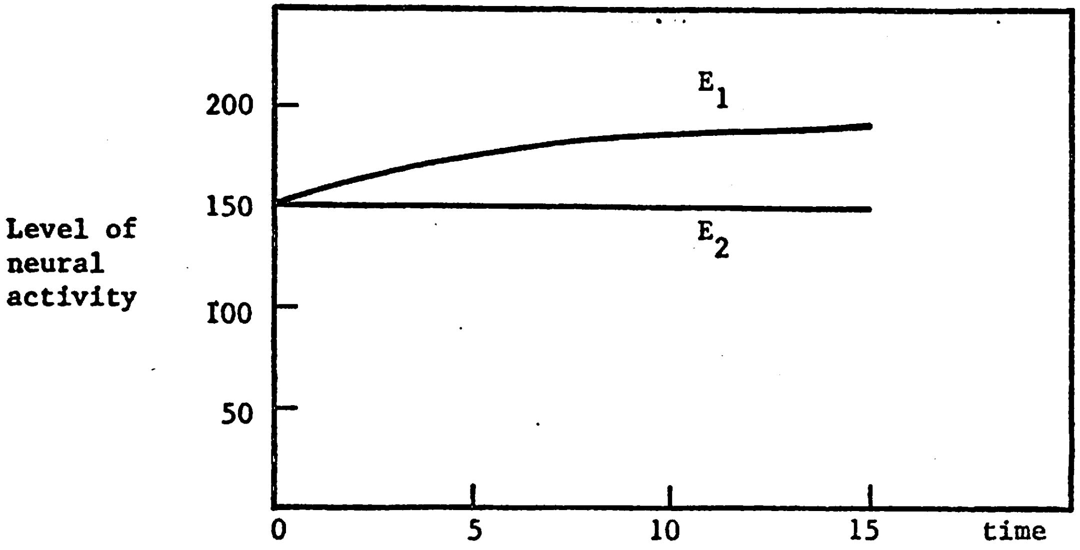


Fig. 4. Response of System without Inhibitory Subpopulation to a Step Increase in P_1

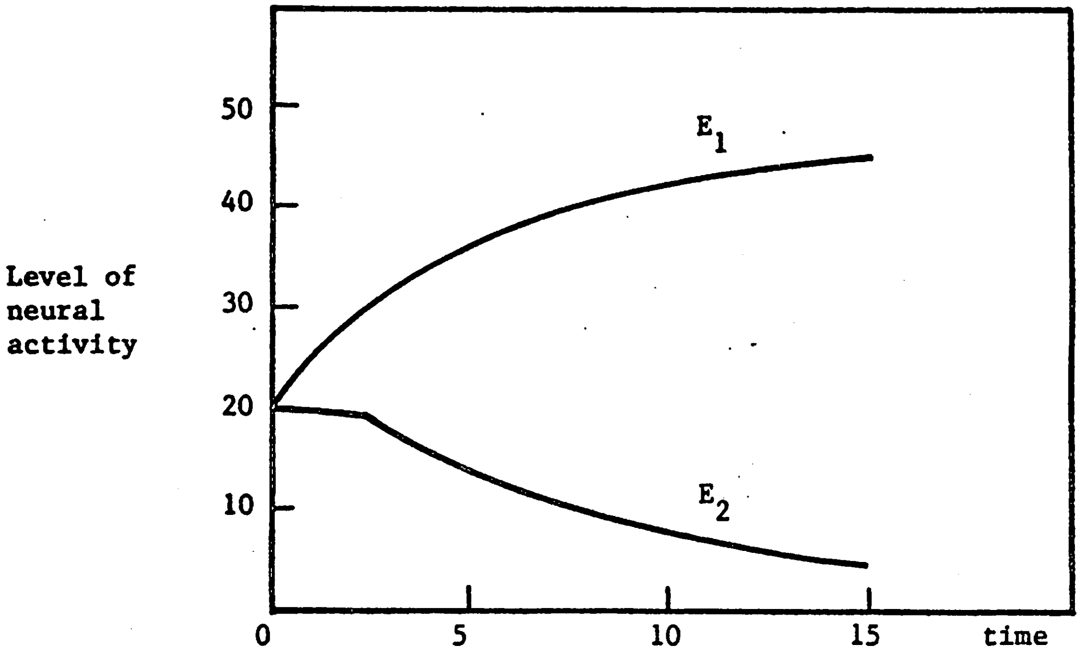


Fig. 5. Response of System with Inhibitory Subpopulation to a Step Increase in P_1

input P_1 causes an increase in the corresponding activity level E_1, but naturally has no effect on E_2.

Introduction of a common inhibitory population (Fig. 5), through which the excitatory populations interact with each other, dramatically alters the neural dynamics. Now, an increase in P_1, with P_2 fixed causes a corresponding increase in E_1, and a decrease in E_2.

It is this suppressive effect that we will utilise in our model, to show how spurious feature detection activity in one subpopulaton of neurons can be suppressed by the existence of greater activity in some other subpopulation.

By moving into the spatial domain, we observe not only the suppression of spurious input, but also the ability of large clumps of high activity to fill in the gaps between these clumps by exciting neurons in these gaps, and thus building up and filling out the shape of the segment.

4.2. Interaction of Spatial Arrays

The concept of interacting populations can be generalised to that of interacting spatial arrays. Location of a neuron becomes important, and we assume that interaction is greater between near neighbors than across a large distance.

This is represented by making the E_j and I functions of space as well as time. a, b, d, and g then become spatial parameters. Therefore,

$$E_j(x,y;t+1) = \iint_{\xi,\zeta} a(\xi,\zeta).E_j(x-\xi,y-\zeta;t)\, d\xi d\zeta - \iint_{\xi,\zeta} b(\xi,\zeta).I(x-\xi,y-\zeta;t)\, d\xi d\zeta + c.P_j(x,y;t)\ , \qquad (26)$$

where $j=1,2,\ldots N$, and j represents the jth layer of excitatory neurons. Note that the N excitatory populations have identical interaction parameters. Also,

$$I(x,y;t+1) = \iint_{\xi,\zeta} d(\xi,\zeta).\sum_{j=1}^{N} E_j(x-\xi,y-\zeta;t)\, d\xi d\zeta - \iint_{\xi,\zeta} g(\xi,\zeta).I(x-\xi,y-\zeta;t)\, d\xi d\zeta + h.Q(x,y;t). \qquad (27)$$

The limits for ξ and ζ are defined by the individual functions a, b, d and g; c and h are retained as constants because any actual spatial divergence of input can be rewritten so as to obtain the total input to each point (x,y).

It is no longer possible to solve directly for the E_j and I when P_j and Q are spatially varying inputs and there is lateral interaction in the arrays.

The stability criterion for the array is very similar to that obtained for the neural population, and can be explicitly stated when $P_j(x,y;t)$ and $Q(x,y;t)$ are constant and uniform over the array, thus causing $E_j(x,y;t)$ and $I(x,y;t)$ to be independent of spatial coordinates. Thus

$$P_j(x,y;t) = P_j(t) = P_j,\ \text{constant} \qquad (28)$$

$$Q(x,y;t) = Q(t) = Q, \text{ constant} \qquad (29)$$

$$E_j(x,y;t) = E_j(t) , \quad \text{and} \qquad (30)$$

$$I(x,y;t) = I(t) . \qquad (31)$$

Substituting in the previous equation,

$$E_j(t+1) = E_j(t).\iint_{\xi,\zeta} a(\xi,\zeta)\, d\xi d\zeta - I(t).\iint_{\xi,\zeta} b(\xi,\zeta)\, d\xi d\zeta + c.P_j , \text{ and} \qquad (32)$$

$$I(t+1) = \sum_{j=1}^{N} E_j(t).\iint_{\xi,\zeta} d(\xi,\zeta)\, d\xi d\zeta - I(t).\iint_{\xi,\zeta} g(\xi,\zeta)\, d\xi d\zeta + h.Q . \qquad (33)$$

Solving for eigenvalues as before, one obtains the conditon for stability:

$$\iint_{\xi,\zeta} a(\xi,\zeta)\, d\xi d\zeta \le 1 , \text{ and} \qquad (34)$$

$$-\iint_{\xi,\zeta} g(\xi,\zeta)\, d\xi d\zeta \le 1 , \qquad (35)$$

that is, when an array receives no input from other arrays or from outside, activity in the array should die down (<1, condition) or should remain constant (=1, condition), but should not increase.

Note that, though this criterion for stability is obtained for the case of an array with uniform activity, it is valid even when the activity in an array is non-uniform, as is shown below.

Consider the isolated array

$$E_j(x,y;t+1) = \iint_{\xi,\zeta} a(\xi,\zeta).E_j(x-\xi,y-\zeta;t)\, d\xi d\zeta . \qquad (36)$$

Activity at a point (x,y) increases, that is,

$$E_j(x,y;t+1) > E_j(x,y;t) \qquad (37)$$

only if

$$\iint_{\xi,\zeta} a(\xi,\zeta).E_j(x-\xi,y-\zeta;t)\ d\xi d\zeta > E_j(x,y;t)\ . \qquad (38)$$

For those neurons surrounded by more active neighbours, this condition may hold. However, there is at least one neuron, the most active one, for which this does not hold. Say it is located at (x',y'). For this neuron,

$$E_j(x',y';t+1) = \iint_{\xi,\zeta} a(\xi,\zeta).E(x'-\xi,y'-\zeta;t)\ d\xi d\zeta\ . \qquad (39)$$

But, the right hand side $\le E(x',y';t).\iint_{\xi,\zeta} a(\xi,\zeta)\ d\xi d\zeta$ (40)

since E(x',y';t) is the highest level of activity. Also

$$\iint_{\xi,\zeta} a(\xi,\zeta)\ d\xi d\zeta \le 1 \qquad \text{(stability criterion)} \qquad (41)$$

Therefore, the right hand side of eqn. 2.39 is less than or equal to E(x',y';t).

Hence, $E_j(x',y';t+1) \le E(x',y';t)$, (42)

and the activity of this neuron does not increase.

Even if the activity of other neurons reach this value, the inequality of eqn. 2.40 becomes an equality, that is,

$$E_j(x',y';t+1) = E(x',y';t).\iint_{\xi,\zeta} a(\xi,\zeta)\ d\xi d\zeta \qquad (43)$$

and the stability criterion is still satisfied.

Basically, the preceding analysis states that the activity of any neuron in an array cannot increase above the activity level of the most active neuron in that array when there is no external input.

4.3. Simulation of Interacting Spatial Arrays

A system with two identical excitatory arrays and one inhibitory array was simulated. The arrays are one-dimensional, and composed of discrete units rather than being a continuous sheet. For the present, each such unit is considered a neuron, though this is not necessary.

The spatial functions a, b, d, and g may be of any form such that interaction falls off with distance. For computational ease, triangular shapes were chosen (Fig. 6). The inhibitory parameters (b and g) extend over a larger distance than the excitatory parameters (a and d), that is, the inhibitory influence of a neuron extends further than its excitatory influence. If the range of excitatory influence is greater, travelling waves are set up, with the excitatory wavefront always advancing ahead of the inhibitory one.

Because of excitatory lateral interaction between neurons in an array, it is expected that slightly separated regions of high activity will tend to spread and coalesce with each other until the total extent of the high activity region approximates the extent of the segment. Further, because of the indirect reciprocal inhibition between arrays, spurious input should be suppressed.

Figures 7 and 8 show the spatio-temporal dynamics of such arrays. In Fig. 7, input P_1 is higher than input P_2 at all locations except one. Here we have spurious activation of a neuron in the incorrect layer, while the input to the desired layer is low.

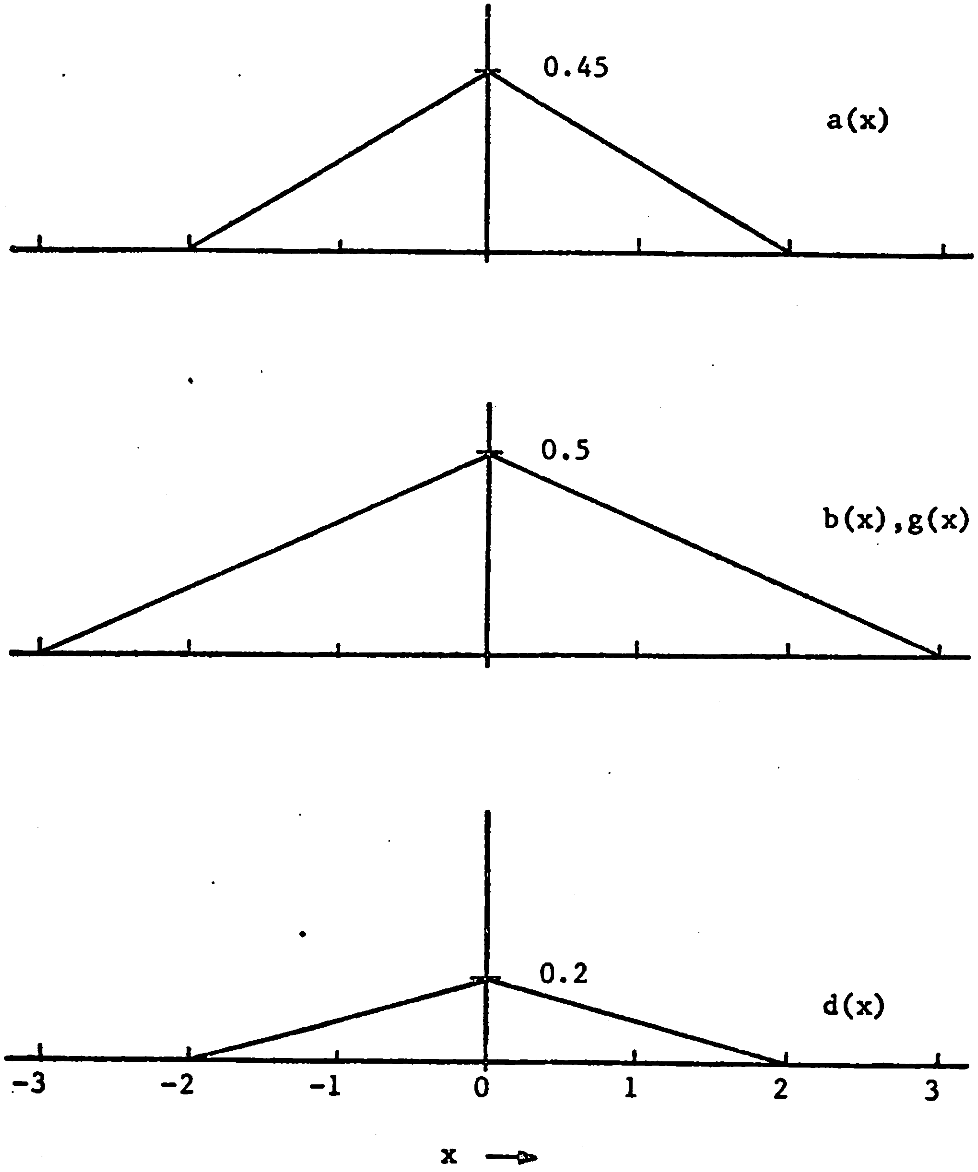
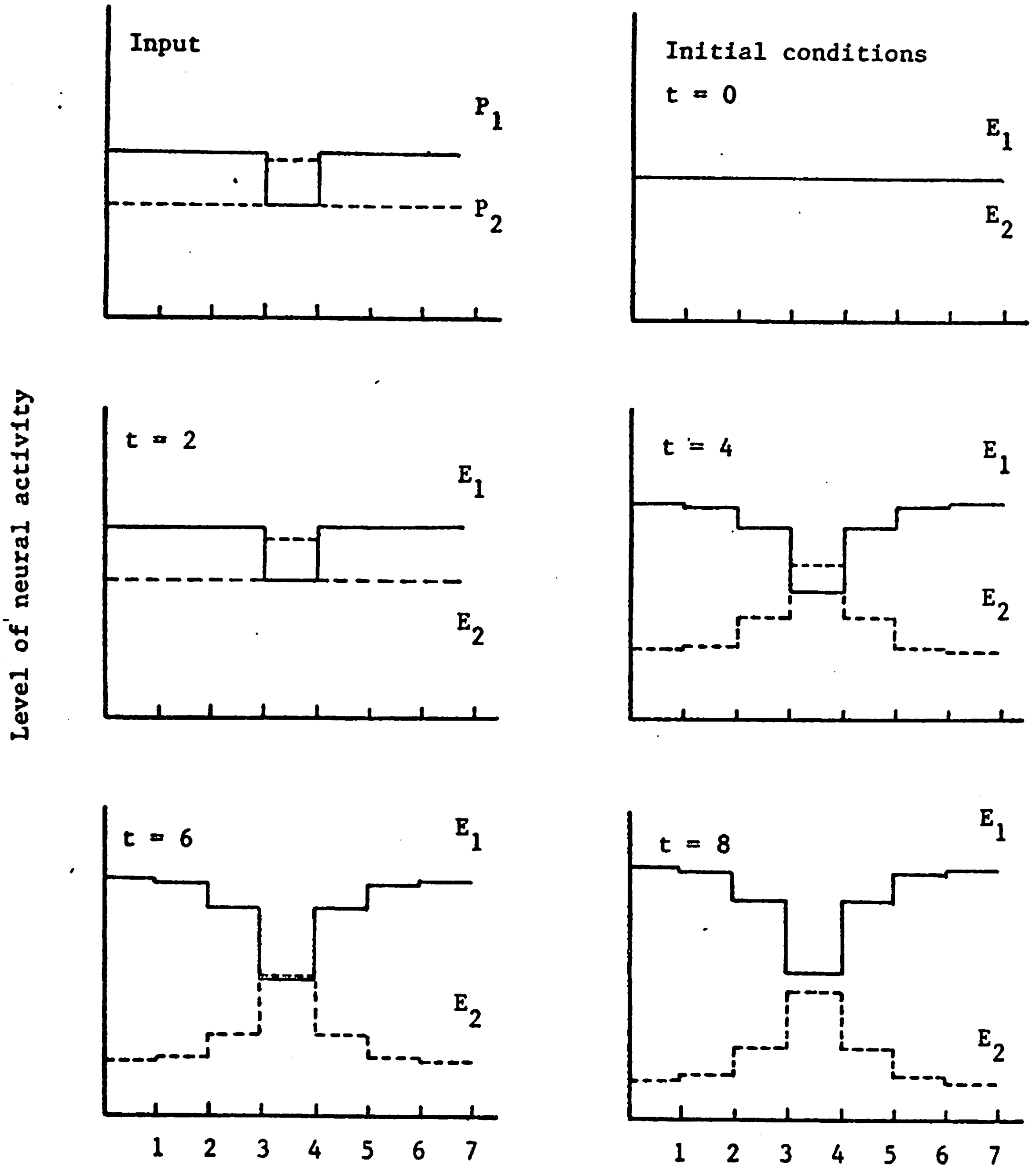


Fig. 6. Spatial Weighting Functions for the Simulation of Interacting Spatial Arrays



Distance along spatial array

Fig. 7. Dynamics of Activity of Excitatory Arrays in Response to a Notch Input

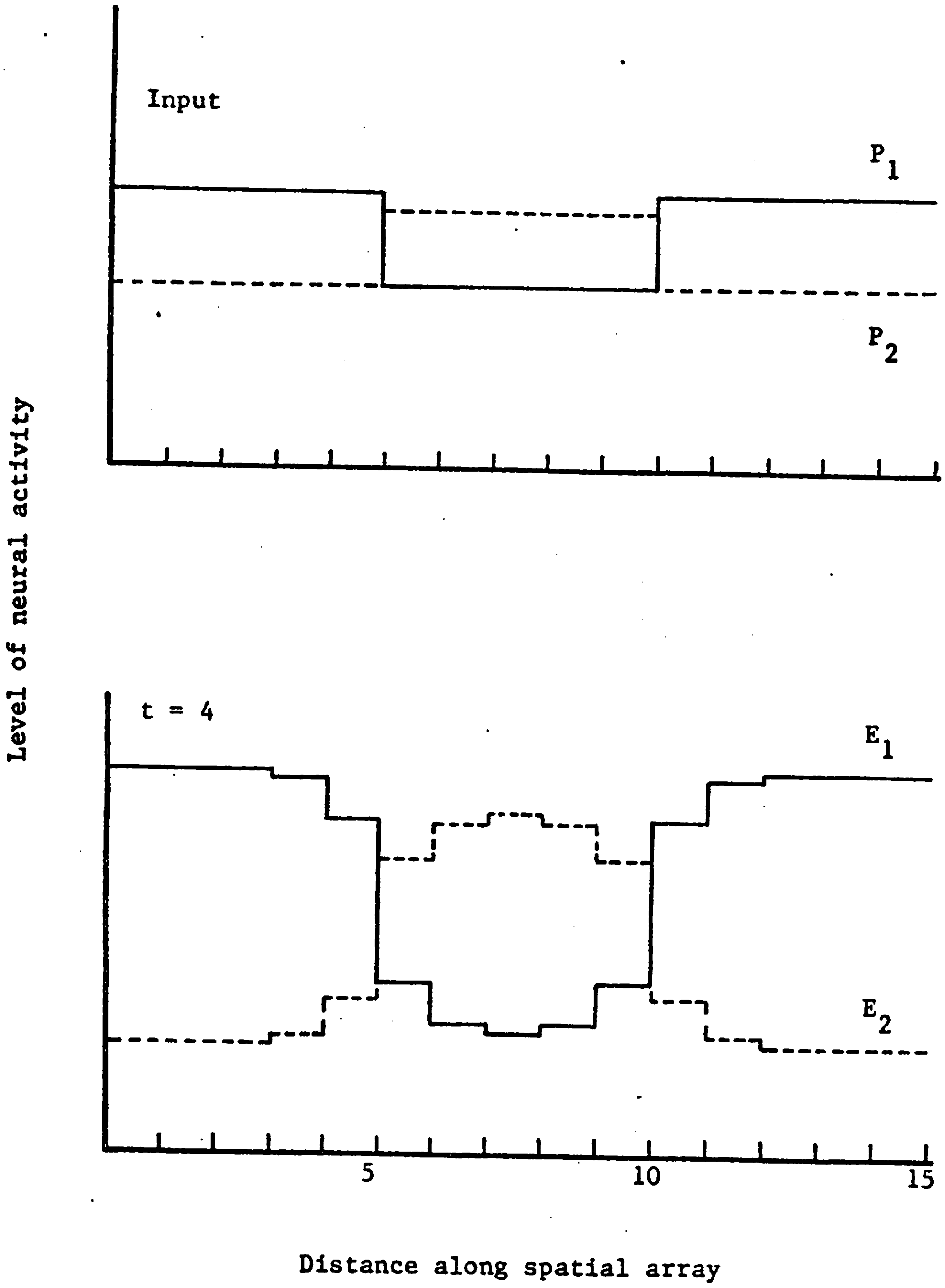


Fig. 8 Typical Activity of Excitatory Arrays in Response to a Rectangular Input

The sequence of diagrams shows how the spurious input is suppressed, and how the gap between two regions of high activity, in the E_1 layer, is closed. Only the first eight iterations are shown, but the effect is strong enough to reverse the normal order of activation levels at that location, that is, though $P_2 > P_1$, at that location, $E_2 > E_1$.

Suppression of the spurious input is assisted by the fact that activity levels in neighboring neurons decrease over time. However, there is some retardation of suppression because P_1 (and hence E_1), at that location is lower than over the rest of the layer. Similarly, filling-in of the gap in activity in the E_1 layer is dependent on the high activity levels in surrounding neurons. The fact that P_2, and E_2, are high at that location retards and reduces the amount of filling-in achieved.

However, if the gap between the regions of high activity is too big (Fig. 8 and 9), no such reversal of order of E_1 and E_2 occurs, that is, where P_2 is greater than P_1, E_2 is greater than E_1 (neglecting edge effects). The regions of high activity remain separate.

In Fig. 8, P_1 is high except in the region of the gap. However, in this gap area, there is an increase of P_2. E_2 shows a corresponding increase, resulting in a localized region of high activity in the E_2 layer. Thus, a gap in an input to one layer, together with higher input to a corresponding location in another layer, causes the development of a new region of high activity, a new segment, as long as the gap is wide enough.

In Fig. 9, the gap in the P_1 input is not filled by an increase in any other input, so that no new region of high activity develops. However, the space is too wide to be bridged by the excitatory influence of the two regions of high activity in the E_1 layer. Thus, without a

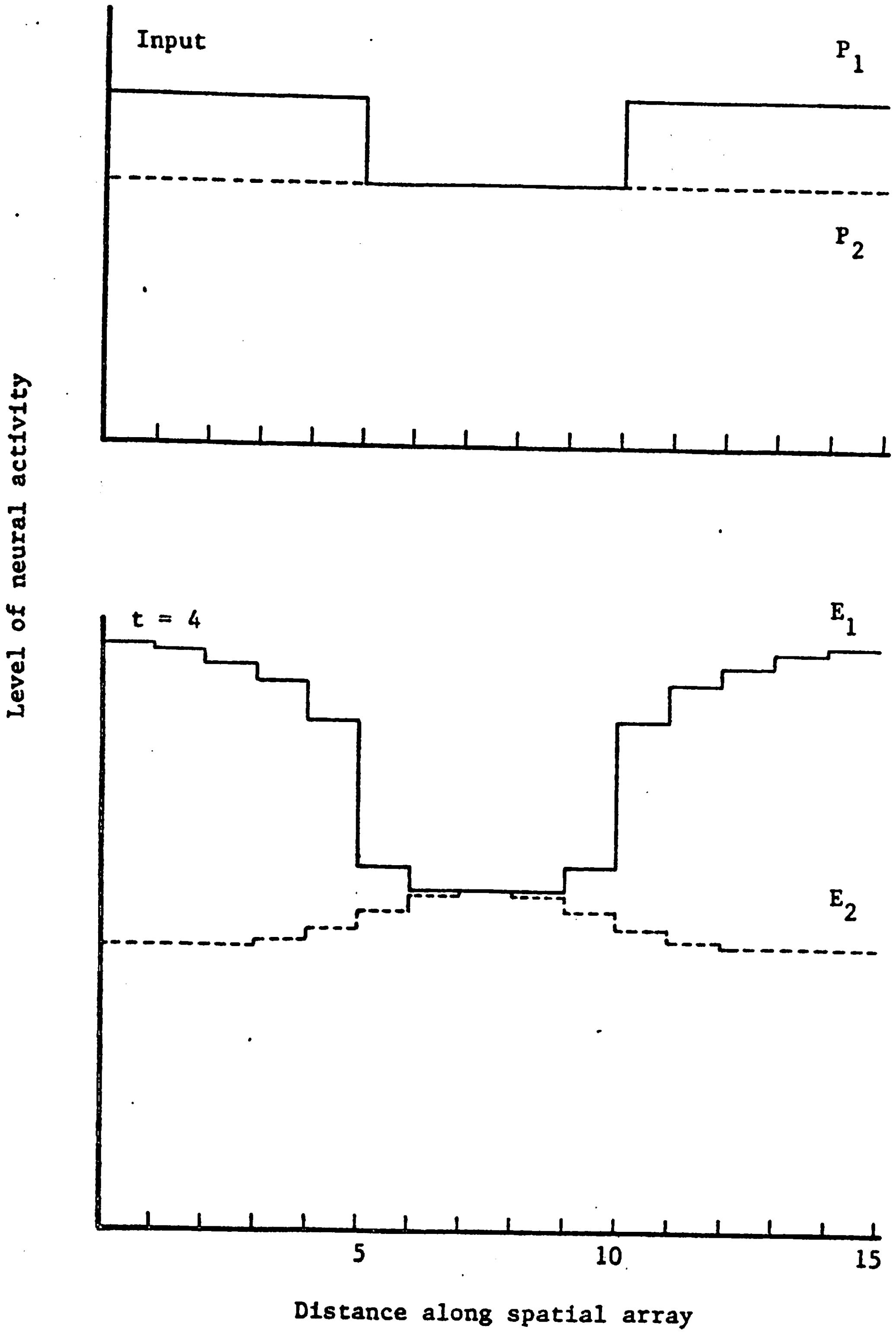


Fig. 9. Typical Activity of Excitatory Arrays in Response to a Single Rectangular Input

sufficient substrate of input, activity cannot spread far outwards from nuclei of high activity. (This is true in a linear system, and may be different in a nonlinear one.) Therefore, in order to form the shape of the segment, besides localized regions of high input, scattered input should be available in order to bridge the gaps between these regions.

5. Application of the Model: Perception of Depth Surfaces in Julesz Random-Dot Stereograms

Random-dot stereograms consist of two slightly different, two-dimensional random-dot pictures, one presented to each eye, whose synthesis, at some location in the brain, causes perception of a three-dimensional scene (Julesz, 1971).

Stereopsis, or the perception of the third dimension, is based on the existence of retinal disparity between the input to the two eyes (Fig. 10). All points on the fixation surface project to homotopic points on the two retinas. However, a point not on the fixation surface projects to two non-homotopic retinal locations. The retinal disparity between these locations is the retinal distance by which they fail to be homotopic. The amount of retinal disparity is a measure of how far the point lies before or behind the fixation surface. (See legend of Fig. 10 for a more detailed exposition.)

The disparity can be detected by the visual system since some neurons in visual cortex fire maximally when an optical stimulus (e.g., a bright dot or bar) illuminates non-homotopic points on the two retina. For different neurons, the firing is maximal for different disparities

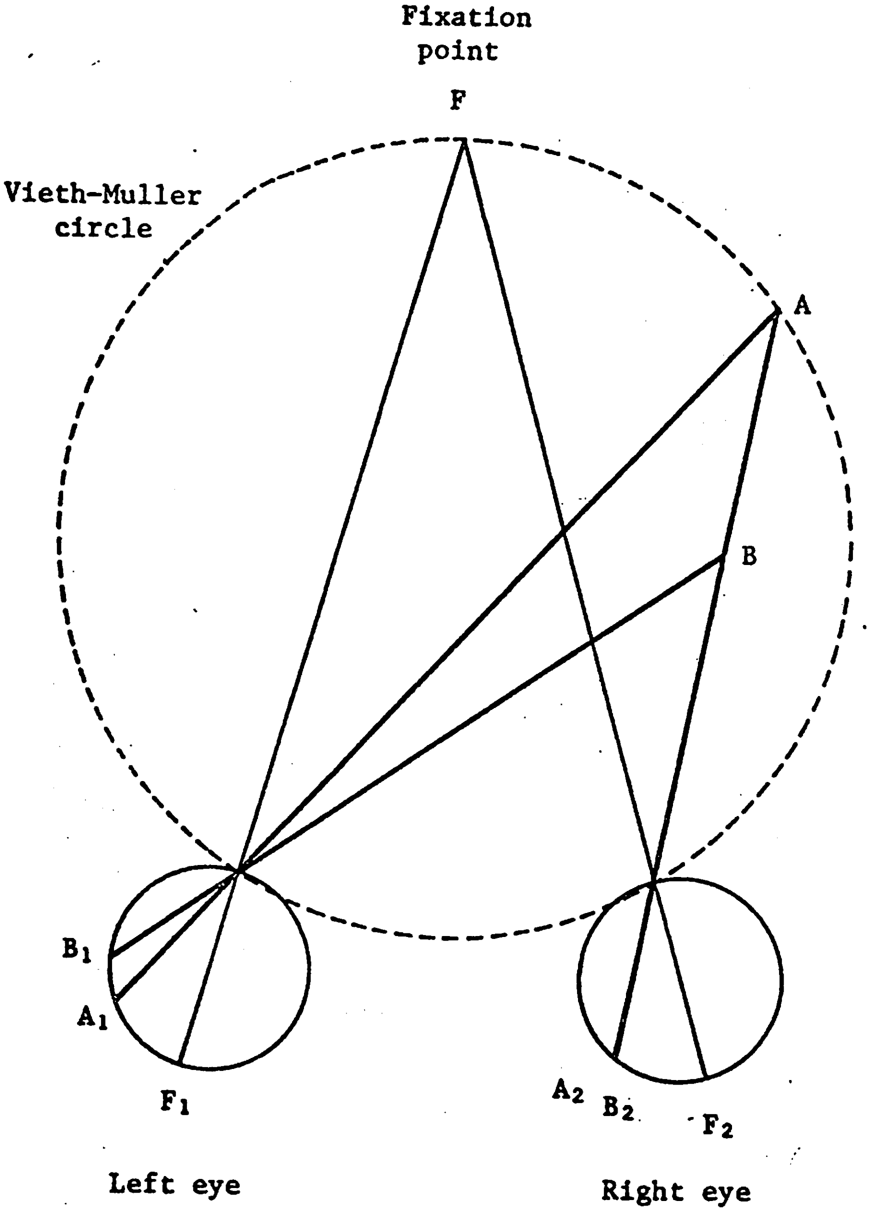


Fig. 10. The Vieth-Muller Circle defines the locus of points in visual space whose images fall on homotopic points in the two retinas, such as A_1 and A_2. Point B, which is not on the Vieth-Muller Circle, projects to the non-homotopic points B_1 and B_2. The distance A_1B_1 defines the retinal disparity of B.

(Barlow, Blakemore, and Pettigrew, 1967).

Consider the simplest random-dot stereogram -- a square surface 'floating' above the plane of the background. The scene has surfaces at two depths, and generates in general two different retinal disparities. If the eyes fixate on one of the surfaces, the corresponding disparity becomes zero.

The stereogram is produced in the following manner. A random two-dimensional pattern of dots is generated. This pattern provides the input to one eye. If the other eye receives the same input, the two patterns can be exactly matched. However, for the stereogram, the input for the second eye is obtained by using a pattern identical to the original pattern but with a square region of dots, in the center, displaced laterally by a fixed amount. Now the inputs to the two eyes no longer match in the center, that is, the central square generates a retinal disparity which is perceived as a displacement of the square out of the background plane.

Perception of retinal disparity is based on the matching of pairs of dots, one from each eye. Since every dot is the same as every other dot, tremendous ambiguity in matching occurs. Each dot may participate in a number of such pairs, and each such pair corresponds to a specific retinal disparity.

Perception of a flat surface implies that the same retinal disparity is perceived at all points on the surface. However, ambiguity in pairing causes production of a variety of retinal disparities, including the correct one. Therefore, perception of the raised central square depends on extraction of the disparity that is common over this region, as well as suppression of all spurious disparities. This is analogous to the problem of segmentation, where extraction of a segment is based on the search for a region where all the locations possess some common feature.

5.1. Computer Simulation

In this section, the segmentation model is tested through its application to the perception of surfaces of different depths in random-dot stereograms. A surface, at a specific depth from an observer, presents the same retinal disparity over its entire extent. Thus, a segmentation process detecting similarity of feature would segment out regions of the visual field such that the disparity within each region is constant.

The test is carried out by means of a computer simulation. It is predicted that, if the segmentation process is effective, the level of activity in the disparity-detecting neurons will reflect the existence of surfaces at different depths from the observer, by the development of corresponding regions of high activity in the appropriate disparity detectors. Disparity detectors not corresponding to an existing surface, but activated by a match between two non-corresponding dots (one from each eye), will have their activity suppressed, so that incorrect perception of disparity does not occur.

The system is simulated on the computer, as shown in Fig. 11. For simplicity of computation, the neuronal arrays are assumed to be one-dimensional.

The system consists of two retinas, whose output is fed to the

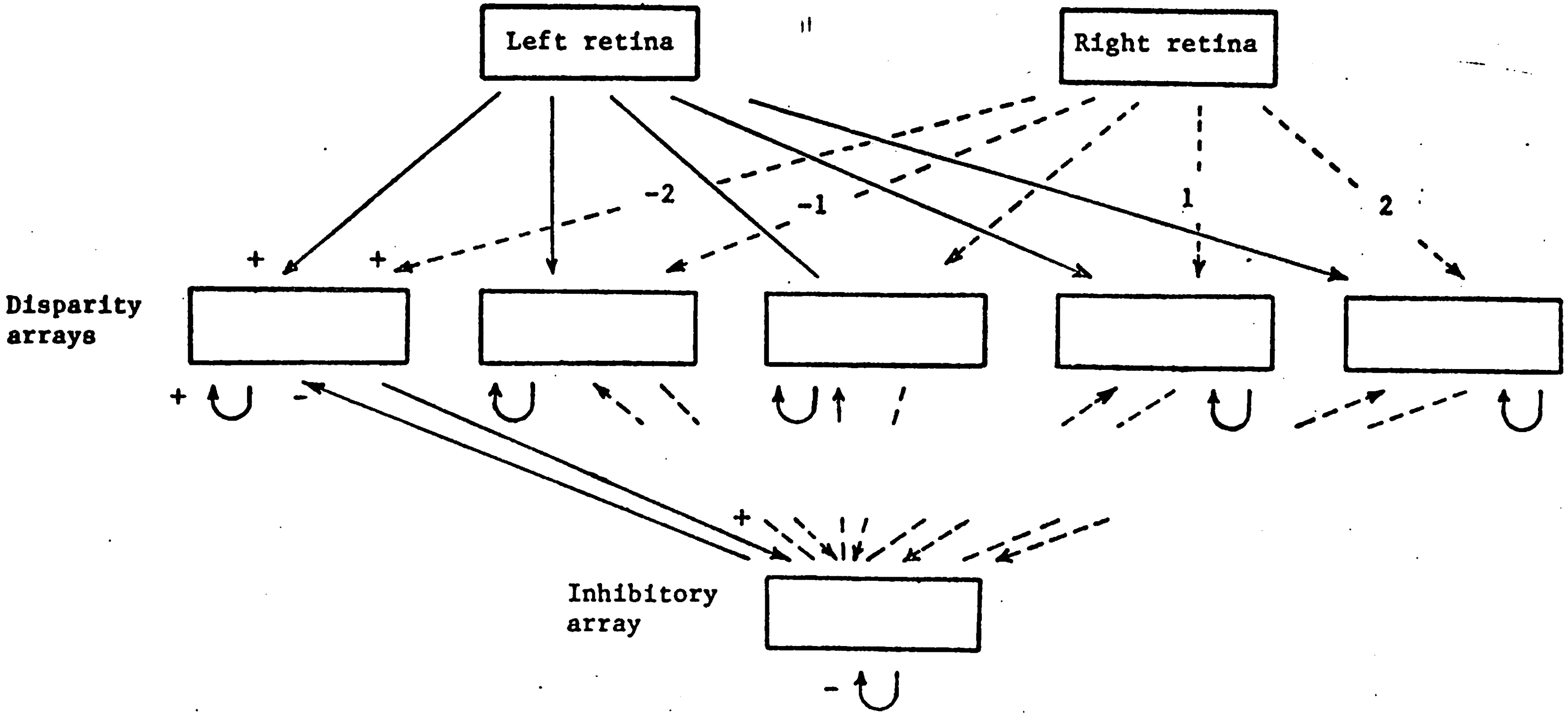


Fig. 11. Application of the Segmentation Model to the Detection of Surfaces at Different Depths from the Observer

feature detectors. In this case, the feature detectors are disparity detectors. The five arrays of disparity detectors interact with the single inhibitory array to generate the segmentation process. All disparity detectors within an array detect the same disparity.

Each retina is modelled as a one-dimensional array of 100 receptors, receiving a binary pattern of light intensity. The input to the retinas is a one-dimensional random-dot stereogram, analogous to the two-dimensional case of a square surface raised above a background field. It consists of a line with a central raised region. The left and right retinas receive the left and right patterns of the random-dot stereogram, respectively.

A random one-dimensional binary pattern of length 100 is generated as the input to the left retina. The right retina receives the same input as the left retina does except in the middle region, where the central 20 units of the pattern have been shifted to the left by 2 units. Thus, when the input patterns to the left and right retinas are compared, the central region shows a disparity of -2 units (because one pattern has to be slid 2 units to the left to match with the other) while the regions on either side have a disparity of 0 units, giving rise to the percept of a line, with the central region raised above a background.

The five disparity-detecting arrays, each of length 100, detect disparities from -2 to +2 units. Each disparity-detecting neuron receives two inputs, one from each retina. A neuron, at a location j, and detecting a disparity of n units, receives an input from the j-th location in the left retina and from the (j-n)th location in the right retina. The two inputs are compared, and if they are the same, the disparity detector is activated. Therefore, a neuron detecting a disparity of n units compares an input from the left retina with an input from the right retina which has been slid to the right by n units, and extracts the modulo-2 sum as shown below:

$$\text{Disparity}\,(j) = \text{Eye}_{\text{left}}\,(j) \oplus \text{Eye}_{\text{right}}\,(i+j) \qquad (44)$$

for $i = -2$ to $+2$, indexes the disparity arrays corresponding to disparities from -2 units to +2 units

$j = 1$ to 100, indexes the location in the arrays.

The output of interest is the final state of activity of the neurons in the disparity-detecting arrays. The theory predicts that the array of neurons detecting a disparity of 0 units will show high activity in the regions corresponding to the background line (i.e., from locations 1 to 38 and 59 to 100), while the array detecting a disparity of -2 units will show maximum activity in the center (from locations 39 to 58), corresponding to the raised central region of the line. Activity of any other disparity detector is undesired and will be suppressed or reduced.

Figure 12a shows the inputs to the two retinas. The initial state of activity of the neurons in the disparity-detecting arrays is shown in Fig. 12b. Each neuron is in one of two states, indicating whether or not it is receiving disparity input. As can be seen, it is difficult to tell which disparity inputs correspond to an actual surface and which to spurious activation of disparity detectors by the matching of non-corresponding dots.

Figure 13 is the activity of the disparity-detectors in the steady state. Three regions of high activity are clearly visible. Array E_1, detecting a disparity of -2 units, shows greatest activity in the

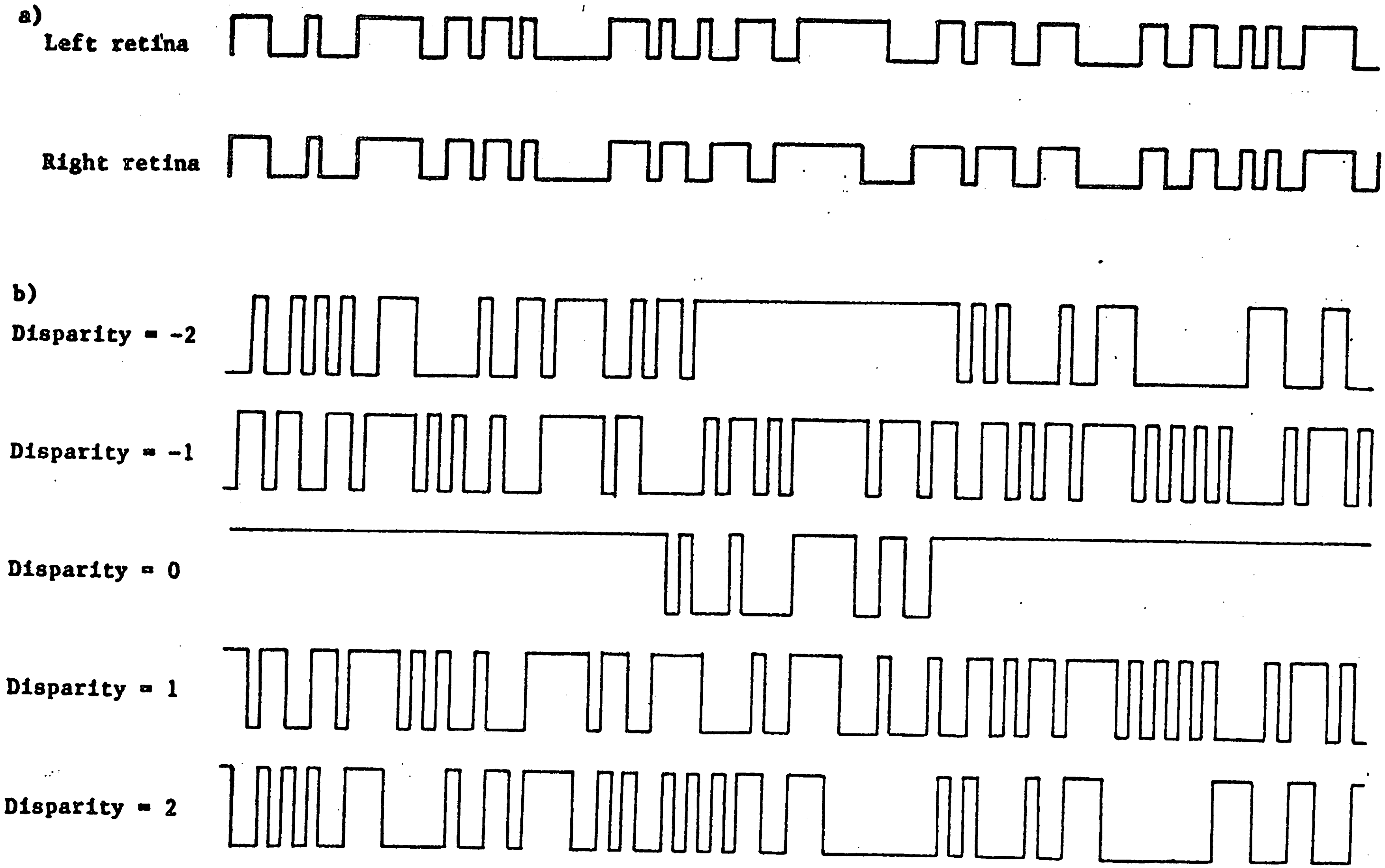


Fig. 12. Random-dot Stereogram Inputs to the Retinas and the Response of the Disparity Arrays

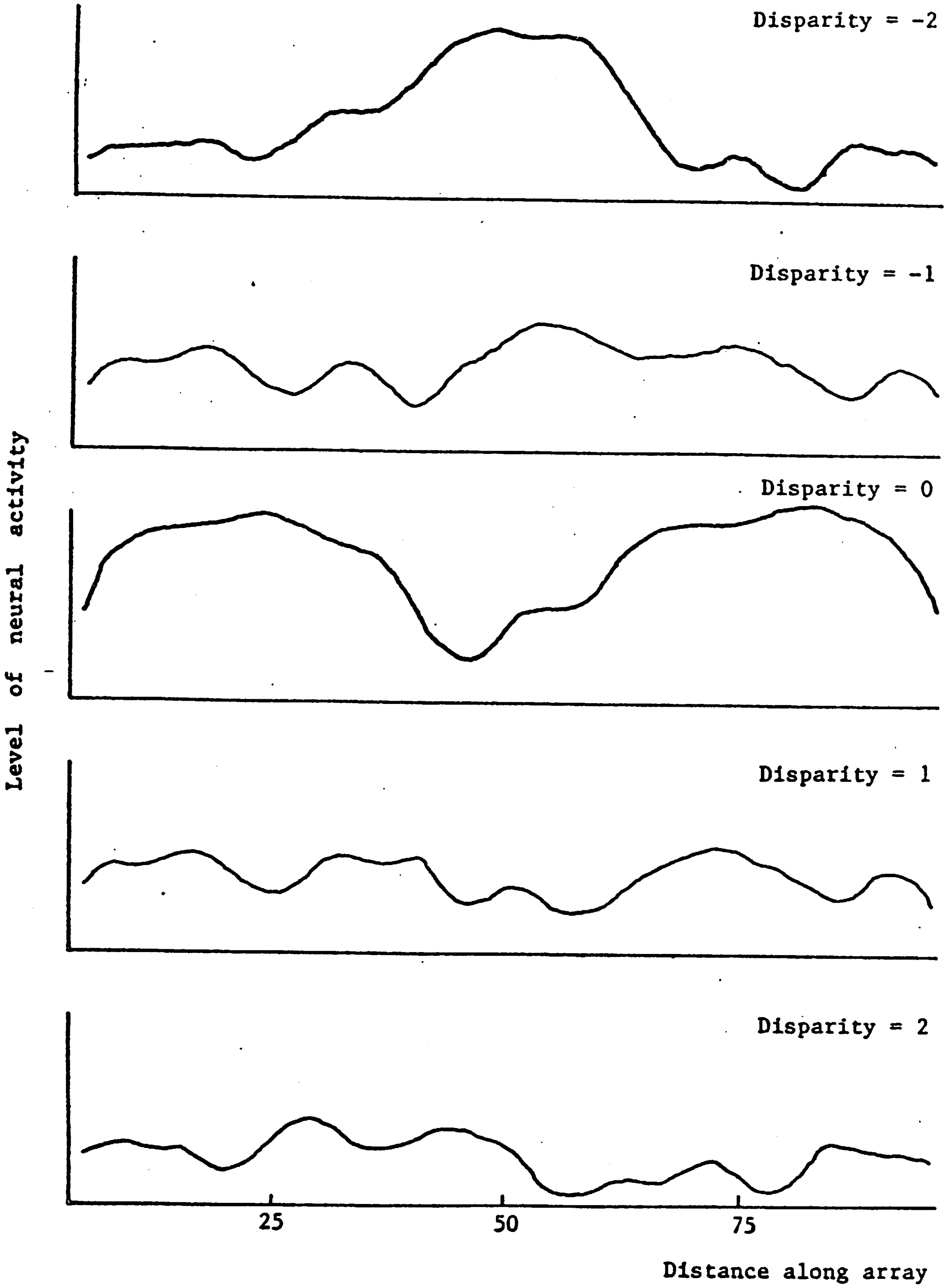


Fig. 13. Activity in Disparity Arrays in Response to Random-dot Stereogram Inputs. Here the Response of each Neuron is Linear, as described in the text.

center, indicating a central raised surface (or line) in the input pattern. Array E_3, on the other hand, has two regions of high activity, one on each side of the central peak. This array detects disparities of 0 units, thus indicating two flanking surfaces (or lines) lying behind the central raised region.

It should be noted that some spurious activation of disparity detectors in other arrays does occur. However, the level of activation is far less than the activity levels developed in the relevant regions of arrays E_1 and E_3, and is at or below the level of activity that would be developed with noise input. This occurs both because of the suppression exerted by the high levels of activity in E_1 and E_3 and also because of a lack of the enhancement of activity which occurs when adjacent disparity detectors are activated.

So far, each neuron has been modelled as a linear threshold element. In a subsequent simulation, the effect of nonlinearity was investigated. The nonlinearity introduced is depicted in Fig. 14.

Figure 15 shows the effect of the nonlinearity on the final levels of activity in the disparity detector arrays. The regions of high activity, in arrays E_1 and E_3, are even more clearly visible. Spurious activity is further reduced. However, the basic characteristics of the activity pattern are similar to that of the linear case.

The introduction of nonlinearities in the response of both the excitatory and the inhibitory arrays leads, under certain conditions, to the occurrence of bistable disparity detectors, or disparity detectors with hysteresis. (Wilson and Cowan (1971) have developed the analysis of such a system in the case of a single excitatory and a single inhibitory array.) In such a system, the regions of high and low activity are

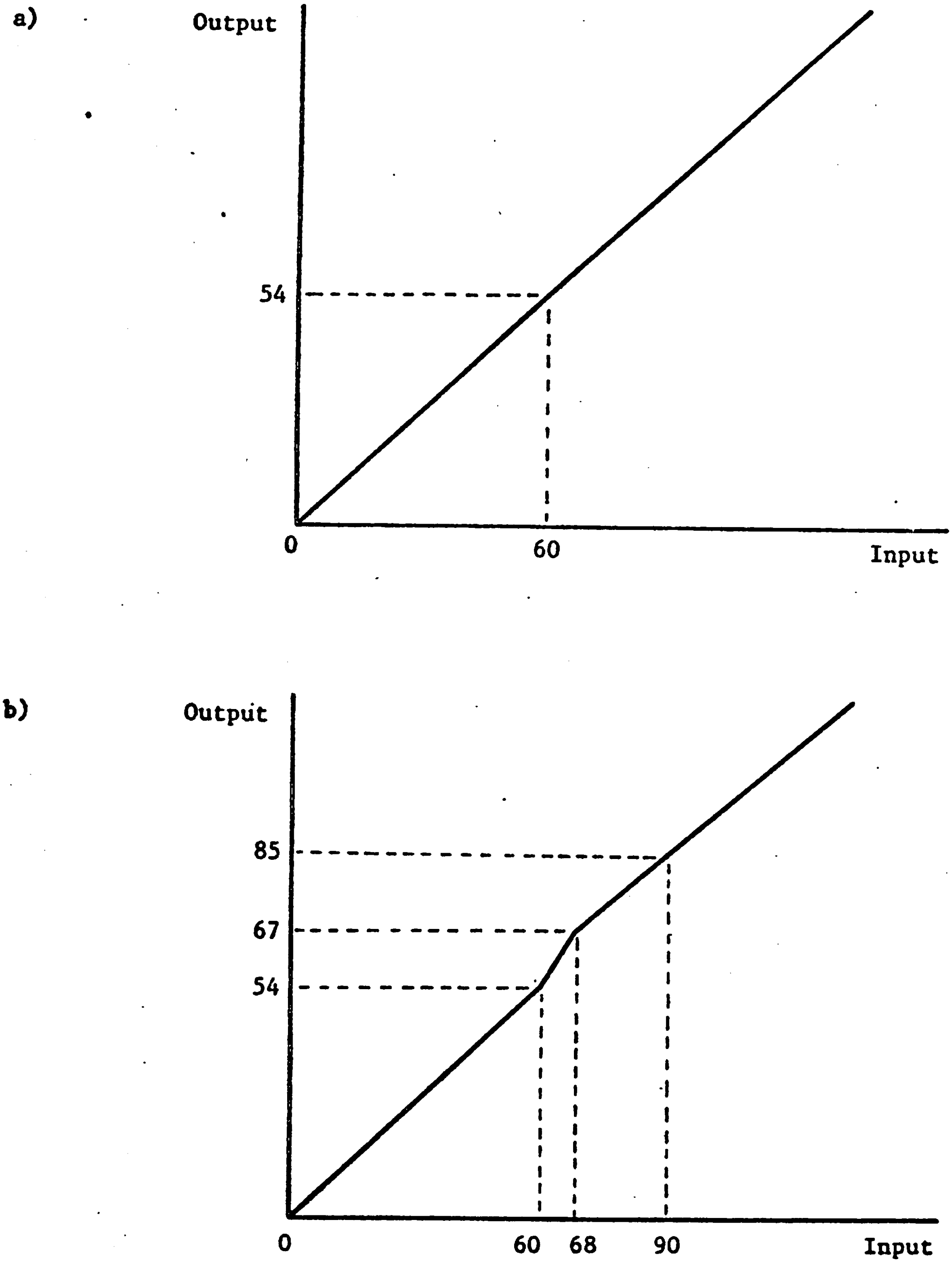


Fig. 14. Response Characteristics of an Excitatory Subpopulation a) Linear System Response, b) Nonlinear System Response

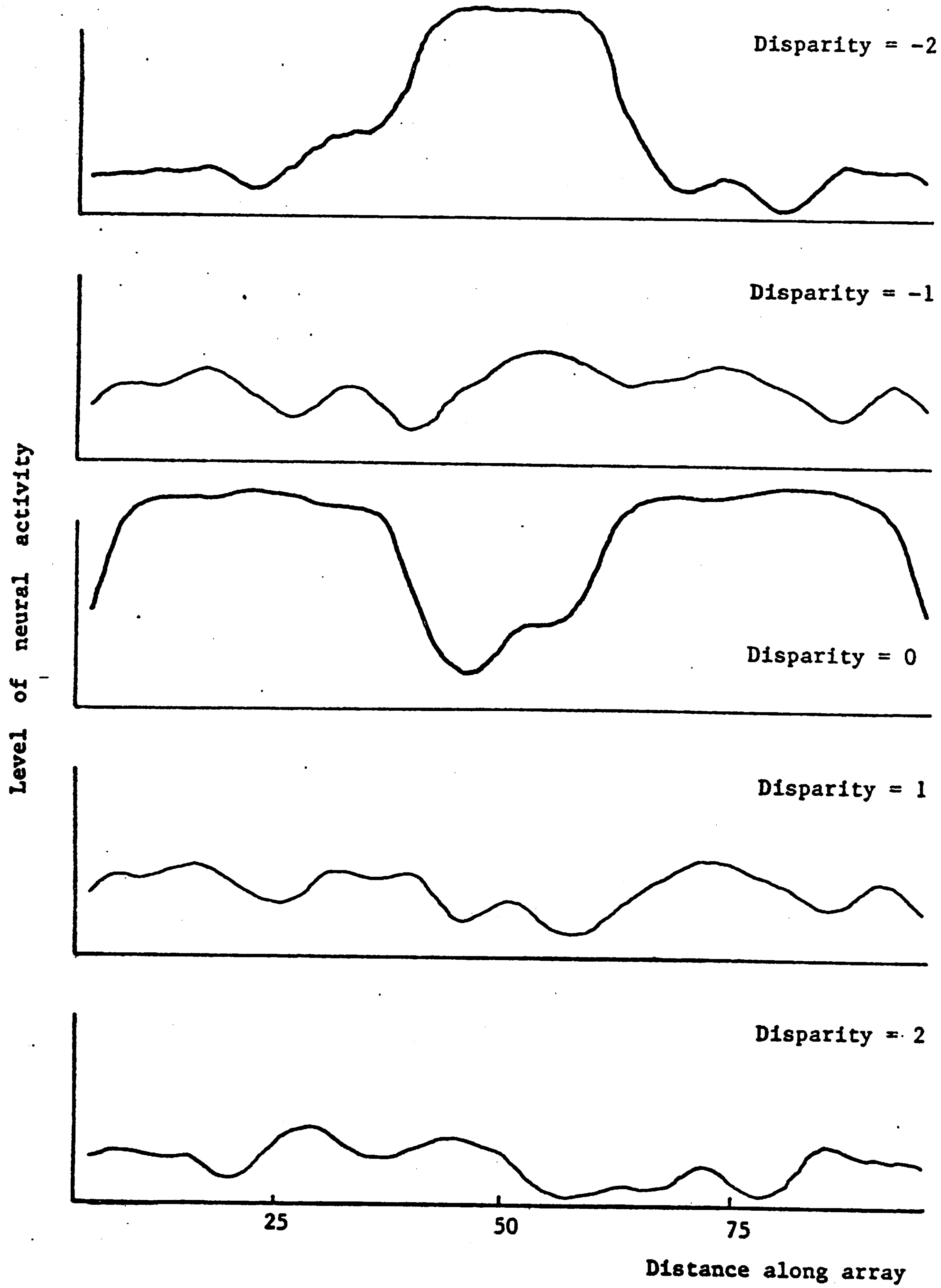


Fig. 15. Activity in Disparity Arrays in Response to Random-dot Stereogram Inputs. Here the Response of each Neuron is Nonlinear, as described in the text.

even more clearly defined, and the extraction of segments occurs very suddenly. Such a system is a 'cooperative' one, because although hysteresis retards a change of state in the disparity detectors, when some of the detectors switch to a state of high activity, they facilitate a change of state in their neighbors. Segmentation, in such a system, is resistant to adaptation of the input, because of the characteristics of hysteresis.

While nonlinearities can improve the performance of the segmentation model, the simulation shows that the basic requirement of segmentation, that is, the development of appropriate regions of high activity in the arrays of feature detectors, can be satisfied by the kind of neural interaction process between linear, threshold elements, proposed in the earlier sections.

5.2. Limitations and Extensions of the Model

We will consider two more complex Julesz figures; namely, the step pyramid, because it possesses a number of flat surfaces at different depths, and the sinusoidal trough, because of its smooth change in disparity.

Why is it that once one of the stages of the step pyramid is perceived, the others become apparent so much more rapidly? Why is it that once the slope at one point of the trough has been determined, the remainder of the changing slope is easily seen? Can a theory of segmentation based on similarity, as suggested above, explain these phenomena?

First we try to work within the framework of the theory with disjoint subpopulations of neurons, each subpopulation tuned to a different feature or class of features.

In the case of the <u>sinusoidal trough</u>, the change in disparity from point to point may be small. Detection of a specific disparity excites neurons in the corresponding subpopulation. If there is some distribution in the disparities detected by this subpopulation, detection of one disparity makes detection of other disparities, within that distribution, easier. However, if the feature distribution for the different subpopulations are disjoint, detection of some region of the trough does not facilitate detection of some other region.

Thus, one possible extension to the model is to permit each subpopulation to have a broadly-tuned disparity detector function, overlapping with the disparities detected by other subpopulations. Consequently, the subpopulations will no longer be disjoint. This may be formalized as follows:

$$E(f;x,y;t+1) = \iiint_{\eta\xi\zeta} a(\eta;\xi,\zeta).E(f-\eta;x-\xi,y-\zeta;t).d\eta.d\xi.d\zeta$$

$$-1(f) \iint_{\xi\zeta} b(\xi,\zeta).I(x-\xi,y-\zeta;t).d\xi.d\zeta$$

$$+ c.P(f;x,y;t) \qquad (\ 45\ )$$

where f is the feature dimension,
1(f) is defined as unity for all values of f,
and a(η;ξ,ζ) is a unimodal function, peaking at (f;x,y) and indicates that a neuron receives most input from those neurons that are nearest to it, spatially, and along the feature dimension.

Mathematically, the variable f is similar to the variables x and y, and the segmentation model is simply extended to a higher dimension.

In the case of the step pyramid, a segmentation model with disjoint subpopulations, detecting different features, does not allow detection of one depth surface to facilitate detection of the adjacent step, unless the steps are very shallow. The extended model, with the subpopulations no longer disjoint, does permit such facilitation. The facilitation decreases with increasing height of the steps.

This extension of the model is a testable one. The time taken to detect the step pyramid can be measured, as a function of the height of the step. Figure 16 indicates the expected monotonic relationship between detection time and the height of the step.

Prior knowledge of what the figure is contributes to the speed with which a Julesz figure is detected. Julesz (1971) suggests that this is based on learning the appropriate sequence of vergence movements. Each vergence movement brings a new surface into the fixation plane so that it can be perceived. An alternative approach can be based on the expectation-tuning of the function a(f;x,y).

The function a(f;x,y) is a unimodal function indicating that a neuron receives most input from those neurons that are nearest to it spatially and along the feature dimension.

For the detection of flat surfaces, at different depths, a(f;x,y) should be sharply tuned along the feature dimension, that is, for a particular disparity but broadly tuned in the x-y plane. Such a

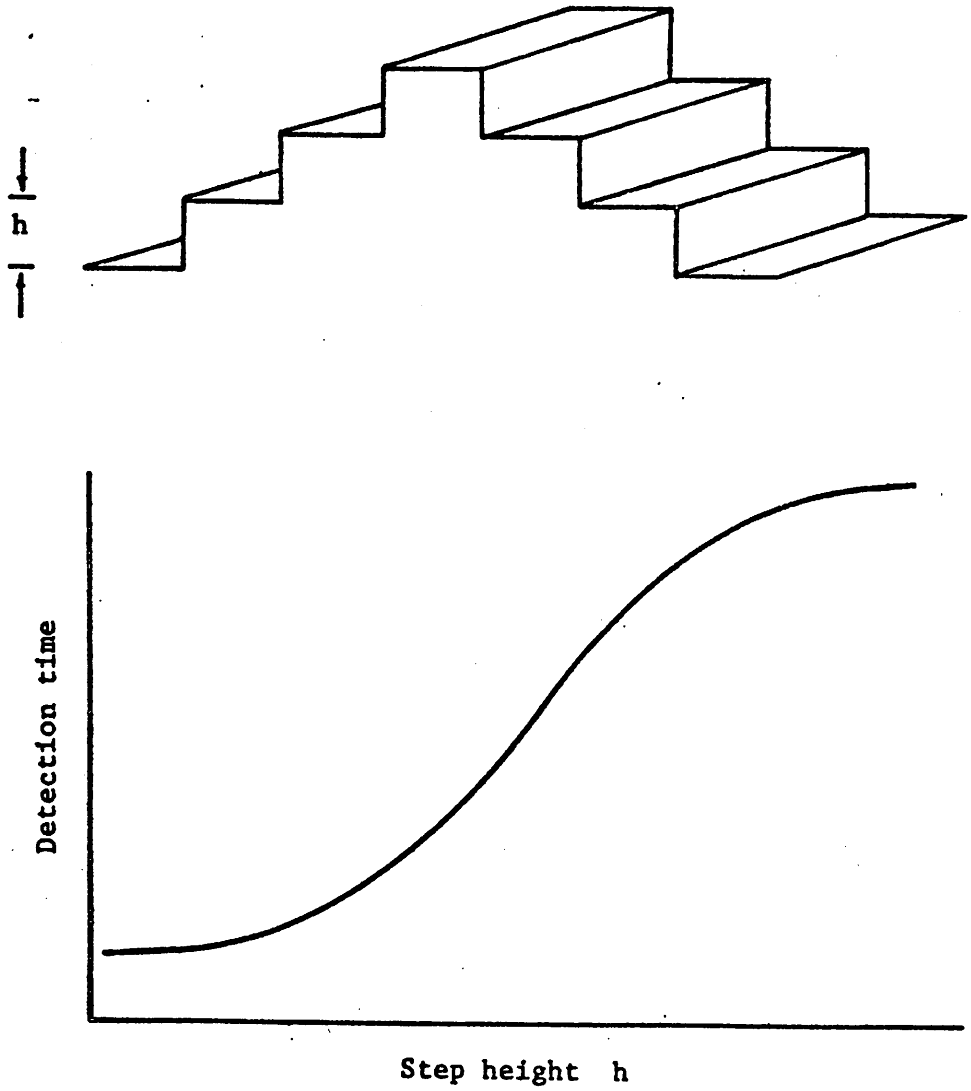


Fig. 17. Illustrating the Monotonic Relationship between the Height of the Step in a Random-dot Stereogram of a Step Pyramid and the Time taken to detect the Step

system searches for large surfaces which show very little change in disparity.

On the other hand, when the surface slopes away, the disparity changes smoothly with location. One possibility is an a(f;x,y) more broadly tuned along the feature dimension. Another alternative is that the tuning along the feature dimension remains sharp, but is a function of (x,y). For example, a disparity detector receives from neurons detecting small disparities on one side, but, from the other side, it receives from those detecting large disparities. A third, and interesting alternative is the possible existence of neurons detecting <u>changes</u> in features, analogous to neurons in the auditory system sensitive to a sweep in frequency, or to neurons in the visual system sensitive to movement. Thus a(f;x,y) can be extended to $a(f,\dot{f};x,y)$, and tuning for a slope consists of sharpening the tuning for $\dot{f}$.

This tuning of $a(f,\dot{f};x,y)$ may be self-organizing, besides being actuated from higher levels. Detection of a disparity, or a disparity change, at some location, causes tuning of the feature detectors at that, and nearby, locations thus facilitating further detection. As the region of accurate detection spreads, the tuning effect spreads in advance of it.

5.3. Comparison with Other Models of Stereopsis

This segmentation model applied to stereopsis differs from others in that it is not formulated expressly for stereopsis. In the general case of segmentation, it proposes a neural basis for partitioning or segmenting the visual field such that each segment is defined by a single feature, for example, a color or a texture. In the particular case of stereopsis, the retinal disparity is the feature defining the depth of different surfaces from the observer. From experiments with random-dot stereograms, Julesz (1971) has shown that a cluster of similar disparities in some location of the visual field leads to perception of a surface of the corresponding depth even if there are no other cues indicating the presence of the surface.

Models of the phenomenon of stereopsis are better developed than those for segmentation. The most detailed model of stereopsis is the spring-coupled dipole model of Julesz (1971). Others are a model of binocular fusion by Dodwell and Engel (1963), Sperling's hypothesis of 'stereopsis energy' (1970), and a computer model, AUTOMAP-1, by Julesz (1962).

The binocular fusion model presented by Dodwell and Engel (1963) is based on parallel processing. They postulate the simultaneous processing of many depth planes, with the plane containing the most activity being the one that is perceived. The concept is similar to the one proposed here for segmentation, but the model does not discuss possible neural mechanisms that explain the development of high neural activity for a particular depth plane.

The hypothesis of 'stereopsis energy' (Sperling, 1970) and the modulation of stereopsis energy by image disparity is of conceptual importance. However, it is a global model, not concerned with the effect of local variations in image disparity. Further, it does not propose a neural mechanism for the determination of stereopsis energy or image disparity.

The computer model, AUTOMAP-1 (Julesz, 1962) operates on the basis of a difference-field. The left and right retinal images are superposed, and subtracted point-by-point from each other to give a difference field. The images are then shifted laterally with respect to each other and the process is repeated for many such shifts, each shift corresponding to a different disparity. This method detects clusters of similar disparity, and thus detects surfaces at different depths. Note that there is no further interaction between the difference fields. In the segmentation model proposed here, AUTOMAP-1 corresponds to the first step of initial activation of the disparity arrays.

The spring-coupled dipole model (Julesz, 1971) is of much greater richness and can explain many of the phenomena associated with stereopsis. The magnetic and spring coupling between the dipoles permits interaction that is not permitted in the AUTOMAP-1 model, thus generating properties such as a maintained perception of depth surface under transient occlusion of the visual input, or the alternate perception of ambiguous surfaces. One drawback of the model is that a serial search must be made through the various disparities to locate clusters of similar disparity. Further, the elements of the model, springs and dipoles, are not explained in neurophysiological terms.

The application of our segmentation model to stereopsis generates a system with characteristics similar to those of the spring-coupled dipole model, but with the added advantage that the elements used are neural and that all processing is done in parallel. Facilitatory interaction between similar feature detectors and inhibitory interaction between detection of dissimilar features produce system characteristics analogous to the facilitatory (attractive) spring coupling and the inhibitory (repulsive) magnetic coupling. Further, the versatility of the excitatory feedback factor a(f;x,y), in the extension of the mathematical model

(Sec. 2.5.2), allows investigation of complex forms of interaction that may be difficult to model with the spring-coupled dipole system. Further mathematical and computer analysis of the system is required.

6. Conclusion

This chapter deals with a process in the analysis of visual information, that is generally neglected in the neurosciences. Current theories about the neural basis of visual information processing invoke a hierarchical form of processing involving ever more complex combinations of features. The neural basis of processes involving feedback and interaction between elements of any one stage of this hierarchy are not investigated because of the complexity of the resulting response. The model developed here investigates, mathematically and through computer simulation, one such complex process, the process of segmentation of the visual field. The elements used in the model are neurons, with simple patterns of connectivity to near neighbours. The model shows how the assumption of certain constraints on this connectivity results in the segmentation of the visual field into separate regions, each region being characterised by a cluster of similar features. The hypothesis is made that the simplest form of segmentation is the extraction of regions of similar features, and each region so defined is operated on by a recognition process which computes on the other features present within that region, leading to the perception of an object or part of an object.

The segmentation model is applied to the detection of surfaces, at different depths from the observer, in a random-dot stereogram. A computer simulation shows how the model compares the two stereogram dot-patterns, detecting and enhancing regions with clusters of similar retinal disparities (the cue for depth perception). This specific application of the segmentation model is used to investigate the limitations of the model, and possible extensions.

In conclusion, we note that the model of segmentation developed here represents only the simplest from of the segmentation process. As outlined in Sec. 1.1 segmentation may occur through many different processes, and their investigation may require different approaches. The approach used here, of neural networks with excitatory and inhibitory interactions, was considered the most appropriate for a mechanism of segmentation that utilised detection of feature similarity.

REFERENCES

Barlow, H.B. (1972): Single units and sensation: a neuron doctrine for perceptual psychology?
Perception. 1, 371-394.

Barlow, H.B., Blakemore, C., and Pettigrew, J.D. (1967): The neural mechanism of binocular depth discrimination.
J. Physiol. 193, 327-342.

Beck, J. (1967): Perceptual grouping produced by line figures.
Percept. and Psychophys. 2, 491-495.

Blakemore, C. and Tobin, E.A. (1972): Lateral inhibition between orientation detectors in the cat's visual cortex.
Exp. Brain Res. 15, 439-440.

Dodwell, P.C., and Engel, G.R. (1963): The theory of binocular fusion.
Nature 198, 39-40, 73-74.

Gusman, A. (1967): Some aspects of pattern recognition by computer.
MAC-TR-37 (Thesis). Massachusetts Institute of Technology.

Hubel, D.H., and Wiesel, T.N. (1962): Receptive fields, binocular interaction and functional architecture in the cat's visual cortex.
J. Physiol. 160, 106-154.

Julesz, B. (1962): Towards the automation of binocular depth perception (AUTOMAP-1).
In the Proceedings of the IFIPS Congress, Munich, C.M. Popplewell (ed.).
Amsterdam: North-Holland.

Julesz, B. (1971): Foundations of cyclopean perception.
Chicago: University of Chicago Press.

Koffka, K. (1935): Principles of Gestalt psychology.
New York: Harcourt, Brace and Co..

Sperling, G. (1970): Binocular vision: A physical and neural theory.
J. Am. Psychol. 83, 461-534.

Weitzman, B. (1963): A threshold difference produced by a figure-ground dichotomy.
J. Exp. Psychol. 63, 201-205.

Williams, L.G. (1966): The effect of target specification on objects fixated during visual search.
Percept. and Psychophys. 1, 315-318.

Wilson, H.R., and Cowan, J.D. (1972): Excitatory and inhibitory interactions in localised populations of model neurons.
Biophys. J. 12, 1-23.