The Holographic Hypothesis of Memory Structure in Brain Function and Perception

Karl II. Pribram Marc Nuwer

Robert J. Baron

WHY A HOLOGRAPHIC HYPOTHESIS?

Introduction

Recently a growing number of theorists have invoked the principles of holography to explain one or another espect of brain function. Historically the ideas can be traced to problems posed during neurogenesis when the activity of relatively remote circuits of the developing nervous system must become integrated to account for such simple behaviors as swimming. Among others, the principle of chemical 'resonances' that 'tune' these circuits has had a long and influential life (see, e.g., Loeb. 1907; Weiss, 1939). More specifically, however, Goldscheider (1906) and Horton (1925) proposed that the establishment of tuned resonances in the form of interference patterns in the adult brain could account for a variety of perceptual phenomena. More recently, Lashley (1942) spelled out a mechanism of neural interference patterns to explain stimulus equivalence and bearle (1956) developed a mathematically rigorous formulation of the origin of such patterns of plane wave interferences in neural tissue. But it was not until the advent of holography

From Contemporary Developments in Mathematical Psychology, Vol. II, Friited by R. C. Atkinson, D. H. Krantz, R. C. Luce and P. Suppes. W. H. Freeman & Company, Copyright 1974.

0

with its powerful damage-resistant image storage and reconstructive capabilities that the promise of an interference pattern mechanism of brain function became fully appreciated. As the properties of physical holograms became known (see Stroke, 1966; Chadman, 1966; Collier, Burckhardt, & Lin, 1971), a number of physical and computer scientists saw the relevance of holography to the problems of brain function, memory, and perception (e.g., von Heeriten, 1963; Inless & Pennington, 1965; Westlake, 1968; Daron, 1970; Cavanagh, 1972).

The purpose of the present chapter is threefold: (a) to summarize the neurological evidence that makes holographic processing, storage (temporary or permanent), and image reconstruction a plausible analogy to students of brain function; (b) to present a mathematical network model for the holographic process as a demonstration that optical systems are not necessary for its realization; and (c) to examine the evidence sustaining or negating the neurological assumptions involved in this and alternate realizations.

The Anatomical l'roblem

One of the best established, yet puzzling, facts about brain mechanisms and memory is that large destructions within a neural system do not seriously impair its function. Various controlled experiments have been performed to investigate this puzzle. Lashley (1950) showed that 80 percent or more of the visual cortex of a rat could be damaged without loss of the ability to correctly respond to patterns; Galumbos, Norton, and Frommer (1967) have severed as much as 98 percent of the optic tracts of cats with similar negative results; and Chow (1968) has combined the two experiments into one simultaneous assault, again with little effect on visual recognition behavior. In man, of course, hemianopia and other large scotomata also fail to impair the recognition mechanism. Even small punctal irritative lesions peppered throughout the cortical mantle of monkeys and shown to disrupt its electrical activity leave response to visual patterns intact (Kraft, Obrist, & Pribram, 1960).

These findings have been interpreted by everyone to indicate that the neural elements necessary to the recognition and recall processes must be distributed throughout the brain systems involved. The questions that arise are (a) how is the distribution effected, (b) how does recognition occur, and (c) how are associated events seconded by the network?

An answer that is often given is to consider the luput systems of the neuralisto be composed of large numbers of randomly connected neural elements (Rosenblatt, 1962) and to show either by computer simulation or by mathematical analysis that in a random network of neurons, replication and distribution of signals can occur. Unfortunately for this explanation, the anatomical facts are largely otherwise. In the visual system, for instance, the

reling and cortex are connected by a system of fibers that sun to a great, extent in parallel. Only two modifications of this parallelity occur.

- 1. The optic tracts and radiations that carry signals between the retina and cortex constitute a cheal which which the retinal events converge to some entent onto the interal geniculate nucleus of the thalanus from where they diverge to the cortex. The final effect of this parallel network is that each fiber in the system connects ten retinal outputs to about LCMI control receiving cells.
- 2. In the process of termination of the Abers at various locations in the pathway, on effective overlap develops (to about 5° of visual angle) between neighboring branches of the conducting fibers.

Equally striking and perhaps more important than these exceptions, however, is the interpolation at every cell station of a sheet of horizontally connected neurons in a plans perpendicular to the parollel fiber system. These horizontal cells are characterized by short or absent arons but spreading dendrites. It has been shown in the retina (Werblin & Dowling, 1969) and to some extent also in the cortex (Creutzfeldt, 1961), that such spreading dendrite networks may not generate nerve impulses; in fact, they usually may not even depolarize. Their activity is characterized by hyperpolarization that tends to organize the functions of the system by inhibitory rather than excitatory processes. In the retina, for instance, no nerve inquises are generated prior to the ganglion cells from which the optic nerve fibers originate. Thus, practically all of the complexity manifest in the optic nerve is a reflection of the organizing properties of depolarizing and hyperpolarizing events, not of interactions among nerve impulses.

Some Neurophyolological Considerations

Two mechanisms are therefore available to account for the distribution of signals within the neural system. One relies on the convergence and divergence of nerve impulses onto and from a neuronal pool. The other relies on the prescue of interal (mostly inhibitory) interactions taking place in sheets of horizontal dendritic networks situated at every cell station perpendicular to the assentially parallel system of input fibers. Let us explore the possible role of both of these nechanisms in explaining the results of the lesion studies.

Evidence is supplied by experiments in which conditions of anesthesia are used that suppress the functions of small nerve fibers thus leaving intact and clearly discernible the connectivity by way of major nerve impulse pathways. These experiments have shown that localized retinal stimulation evokes a receptive field at the cortex over an area no greater than a few degrees in diameter (e.g., Talbot & Marshall, 1941). Yet, the data that must be explained

indicate that some 80 percent or more of the visual cortex including the forcal region can be extirpated without marked impairment of the recognition of a previously learned visual pattern. Thus, whatever the mechanisms, distribution of input cannot be due to the major pathways but must involve the fine fibered connectivity in the visual system, either via the divergence of nerve impulses and/or via the interactions taking place in the horizontal cell dendritic networks.

Both are probably to some extent responsible. It must be remembered that nerve impulses occurring in the fine fibers tend to decrement in amplitude and speed of conduction thus becoming slow graded potentials. Further, these graded slow potentials or minispikes usually occur in the same unatomical location as the horizontal dendritic inhibitory hyperpolarizations and thus interact with them. In fact, the resulting micro-organization of junctional neural activity (synaptic and ephaptic) could be regarded as a simple summation of graded excitatory (depolarizing) and inhibitory (hyperpolarizing) slow potential processes.

These structural arrangements of slow potentials are especially evident in slicets of neural tissue such as in the retina and cortex. The cerebral cortex. for instance, may be thought of as consisting of columnar units that can be considered more or less independent basic computational elements, each of which is capable of performing a similar computation (Mountenstic, 1957; Hubel & Wiesel, 1968). Inputs to the basic computational elements are processed in a direction essentially perpendicular to the sheet of the cortes, and therefore cortical processing occurs in stages, each stage transforming the activation pattern of the cells in one of the cortical layers to the cells of another cortical layer. Analyses by Kabrishy (1966) and by Werner (1970) show that processing by one basic computational element remains essentially within that element, and therefore the cortex can be considered to consist of a large number of essentially similar parallel processing elements. Furthermore, the processing done by any one of the basic computational elements is itself a parallel process (see, for example, Spinelli, 1970), each layer transforming the pattern of activity that arrived from the previous faver by the process of temporal and spatial summation, the summation of slow hyperand depolarizations in the dendritic microstructure of the cortex. Analyses by Ratliff (1965) and Rodicck (1965) have shown that processing (at least at the sensory fevel) that occurs through successive stages in such a layered neural network can be described by linear equations. Each computationalclearent is thus capable of transforming its inputs through a succession of stages, and each stage produces a linear transformation of the pattern of activity at the previous stage.

MINAMOORAPILY

Inticat Computing

The problem that thus confronts us is essentially this: how can the relationhips between neural activity become distributed and stored (temporarily or note permanently) by a neural network in which such patterns are transnitted and transformed through several successive stages in which processing an essentially linear parallel process?

Fortunately for neurophysiology, physicists have been concerned with such ystems for a long time: optical devices are parallel transmission systems, and during the past 25 years their processing characteristics have been studied ntensively. One property of optical processing initially called it to our attention. As we shall see, records can be produced in which the input becomes listributed throughout the storage medium. This makes the record resistant o damage, and, in fact, loss of all but a small portion does not destroy its mage reconstructing potential.

Most of us are familiar with the image-generating aspects of optical sysems. A camera records on photographic film a copy of the light intensities effected from the objects within the camera's field of view. Each point on the Im stores the Intensity (the square of the amplitude) of the light that arrives com a corresponding point in the field of view, and thus the film's record ooks like the visual field. What have been studied more recently are the roperties of records made when a film does not lie in an image plane of an ptical system. When a piece of film is exposed to coherent light that is reected and scattered by objects in the visual field, there is no ordinary image roduced on the film. In fact, the film becomes so blurred that there is no esemblance whatever between the pattern that is stored on the film and the Isual field itself. However, when properly illuminated, the film reconstructs he wavefronts of light that were present when the exposure was made. As result, if an observer looks toward the film, it appears as if the entire visual zene were present behind H. The reconstructed image appears exactly us it ld during the exposure, complete in every detail and in three dimensions! he light waves from each point of the visual field had interacted to produce n interference pattern at the film, and it is this interference pattern that was forced throughout the film. Interference patterns give rise to the remarkable haracteristics of optical information storage as we shall show.

Even before the practical demonstration of the use of interference pattrus in the reconstruction of images, Oabor (1948) had mathematically proosed a way of producing images from photographic records. Gabor began ith the intent to increase the resolution of electron microphotographs. He aggested that a coherent background wave and the waves refracted by the same could produce interference patterns that would store both amplitude and spatial phase information. Then, in a second step, these stored patterns could be used to reconstruct an image of the original disease. Gabor christened his tim record a hologram because it contains all necessary information to reconstruct the whole (hal-) image. The use of this term for that type of photographic record has since become common.

As the art and science of holography developed, it became clear that a variety of methods described by a number of mathematical procedures could result in holograms. This chapter briefly describes two elementary types—the Fresnel and the Fourier—and provides a network realization of the Fourier holographic process. Other types of holograms have been found useful: they go by names such as Reflection, Volume, Phase, Color, Pulsed-laser, incoherent, and Digital holograms. All are basically similar to the elementary Fourier and Fresnel types but have special properties that make them especially useful in one or another application. The following exposition is therefore meant to provide only a guideline to holographic processing by presenting the requirements necessary for a holographic hypothesis of memory storage in brain function.

I.cusiess Optical IJulography

The problem faced when trying to store a wavefront of light on photographic film is that film does not store the amplitude or phase distributions (patterns) of light, but instead it records only the intensity (amplitude squared) distribution. The image that is stored on the film is a static representation of the dynamic wavefront of light that arrived during the exposure process, but the stored image has no phase information. As a result, it is impossible to recreate the dynamic pattern of light from which the image was made. If olography offers a way to overcome this problem by recording on film the interference pattern formed by two different wavefronts of light. As we shall show, if the intensity pattern formed by the interference (superposition) of two different wavefronts of light is recorded on the film, sufficient information is retained to enable a reconstruction of either wavefront when only the other one is present at a later time.

Let us suppose than an object, say O_A , is placed near a piece of film and illuminated by a coherent source of light. The wavefront produced at the film by light that is scattered from the object will be denoted by A. The ideal situation would result if film could store the wavefront A. However, when the film is exposed and developed, the image recorded by the film is proportional to the intensity pattern $|A|^2 = AA^\circ$, not the desired pattern A. As we mentioned earlier, the wavefront A cannot be reconstructed from AA° alone. Several different holographic techniques have been studied that enable a wavefront of light to be stored, and we shall discuss four of them now.

One technique that has been widely discussed in only of illuminating the object OA with a plane c but at the same time reflecting some of the colier directly loward the film. The second wavefront a wave and will be denoted by R. When the two diff R interact, the superposition (algebraic sum of ch ments) given by A -f- R is formed. Because A and : light source, they have the same frequency. A st formed, and the film records the Intensity of the by $(A - P R)(A - P R)^{\circ}$. If the film is transparent, . gate independently, and we call these output por one of the wavefronts the output patterns or departure patterns.

Now suppose the film is developed and replaced in the optical system. Also suppose that the object OA is removed so that or allowed to Huminate the film. The departure patt of the incoming wavefront attenuated at each polficient of the film at that point. The departure p $R(A + R)(A + R)^{\circ}$, which can be expanded mat

The first term of the expansion describes the refer an amount Mo -f- Ro. The accord term describ the desired wavefront A attenuated by an amount; 40. This wavefront has all the properties of the original wavefront present d As a result, a person looking toward the film wo: cause OA is not present, the reconstructed imagand since the reconstructed wavefront is an exact appears in three dimensions and has all other pr during the exposure. The last term describes nois the system by the holographic process. The film is wavefront A is said to be stored. (Note that if A rectly, the noise term would not have been produ linage could have been formed directly from the

Now let us suppose that the nitror rather than t during the reconstruction process. The departure given by $A(A - - R)(A - - R)^{\circ} = A(|A|^{\circ} - - |R|^{\circ})$. case, the reference wave is reconstructed (the second term), the wavefront A attenuated (the first term), and again noise is prod ed. There is thus a natural symmetry between the two wavefronts in the syst in.

the literature consists not herent wavefront of light, it light by a plane mirror light is culled a reference ent arrival putterns A and tric and magnetic compoare generated by the same ble interference pattern is interference pattern given and R continue to propa-

y the reference wave R is in is given by the product by the transmission coeftern is therefore given by matically to

ice wave R affenuated by · a reconstructed comy of ing the exposure process. I 'sce' the object Oa. Heis called a "ghost image," py of A, the ghost intage perties that could be seen which is introduced into alled a hologram and the ould have been stored died and the reconstructed ored image.)

object had been semoved pattern would have been RAP 4 ROA.A. In this

Recognition Holography

The second type of holography discussed is a slight modification of the first type. Rather than using a plane mirror to produce the reference wave, let us use a spherical mirror that focuses the reflected light onto a point b in front of (on the output side of) the film. When the hologram produced by this system is developed and replaced in the system, two different departure patterns result depending on how the hologram is illuminated. If the object is removed and the hologram is illuminated only by light reflected by the mirror, the wavefront A is reconstructed. Both the focused wavefront and a noise term are also produced. If the object is used to illuminate the hologram and the mirror removed, the focused wavefront is reconstructed, and a bright spot of light is focused at the point b. A light detector placed at b could be used to detect the bright spot, and since a bright spot is produced only when object O_A is present in the system, the detector can be used to 'recognize' the presence of the object O_A . This optical system can be used for the recognition of three-dimensional objects.

Association I lolography

The third lensless system described is similar to the two previous systems. This time, however, the mirror is replaced by a second object, say O_B . See Figure 1. When objects O_A and O_B are illuminated by a coherent light source, two wavefronts A and B are produced. If a film is exposed to the interference pattern produced by A and B and given by A + B, the film will record the static pattern $(A + B)(A + B)^\circ$.

Now let us see what happens if the film is developed and placed exactly where it was during the exposure. Assume that object O_A is removed. If O_B is illuminated, the arrival wavefront that reaches the film is B, and the static distribution of transmission coefficients on the film is given by $(A + B)(A + B)^\circ$. The departure wavefront is given by $B(A + B)(A + B)^\circ = B(|A|^2 + |B|^4) + A|B|^2 + A^\circ B \cdot B$. The first term shows that the wavefront B is transmitted, the second term shows that the wavefront A is also reconstructed, and the third term shows that noise is produced.

If the objects do not cause light to be focused on the film, the intensity distributions $|A|^2$ and $|B|^2$ are nearly uniform across the film even over small (millimeter) distances. However, the interference pattern (A - P - B) gives rise to an intensity pattern that varies considerably over small distances, and the resulting stored pattern (A - P - B)(A - P - B) resembles a very complex diffraction grating. It is precisely for this reason that the departure pattern is a

reconstruction of the wavefront of A that is not present during the reconstruction process.

Figures 1 and 2 illustrate the holographic association process. Figure 2 shows the exposure process. Light is reflected from two objects, a wide block O_A in the background, and a tall block O_B in the foreground. Both objects are illuminated by a single coherent monochromatic plane wave generated by a laser. After exposure, the film is developed and replaced exactly where it was during the exposure process. Now, however, the object O_A is 'seen' so that the film is exposed to light which is reflected only from the tall block O_B . Because the departure wavefront A is reconstructed by the hologram, the object is 'seen' by an observer even though it is no longer present. Again the reconstructed wavefront is an exact copy of the departure pattern that was present during the exposure, so the object appears in three dimensions and has all other properties that could be seen during the exposure.

The three lensless holographic processes described above are among many similar techniques used for producing holograms. Holograms can be produced by exposure to objects that are either near the film (near-field holograms) or far away from the film (far-field holograms). The optical far-field transforms used above are known as Fresnel transforms. However, both the near-field and far-field lensless holograms—i.e., those produced with scuttered wavefronts as reference waves (e.g., the third system)—are usually referred to as Fresnel holograms. Holograms produced with infinitely far-field transforms are called Fourier holograms.

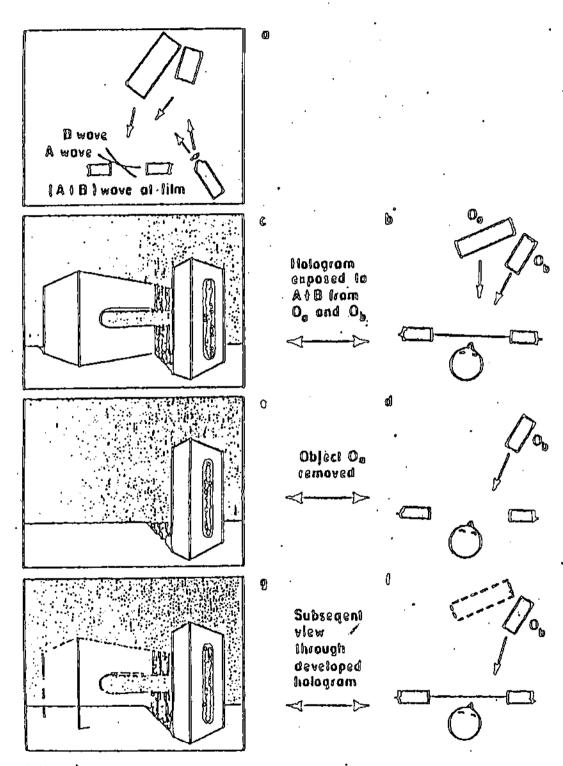
One optical system of the Fourier type, using a plane coherent reference wave produced by lenses, is of special interest to the neurotheorist because a direct analogy can be drawn between it and a layered neural network. Because it is this analogy that we wish to pursue, the mathematical properties of the system are described in somewhat greater detail.

THE FOURIER HOLOGRAPHIC PROCESS

The Lons System

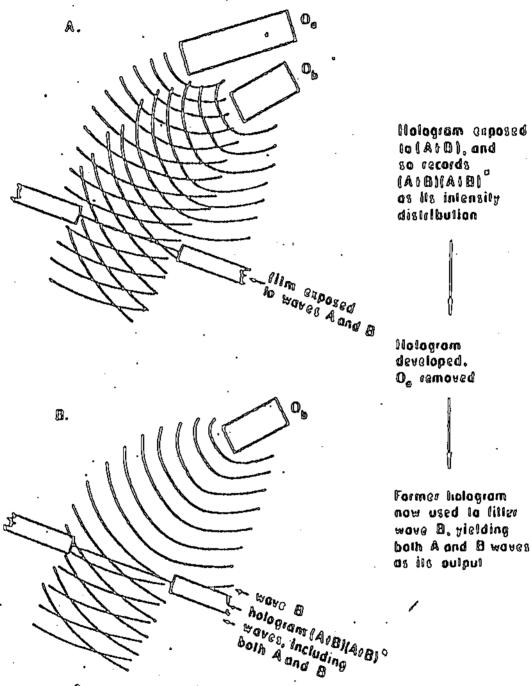
The system of interest consists of two spherical lenses arranged so that the second focal plane of the first lens is coincident with the first focal plane of the second lens. This is shown in Figure 3. The three focal planes of the two lenses are also of interest to this analysis and will be called the input plane, transform or memory plane, and output plane of the system.

It is well known that when a photographic image is placed in one focal plane of a spherical iens and illuminated by a plane, coherent light wave, the Fourier transform of the image is produced in the other focal plane of the lens. (For a nice proof of this result, see Preston, 1963.) For the optical sys-

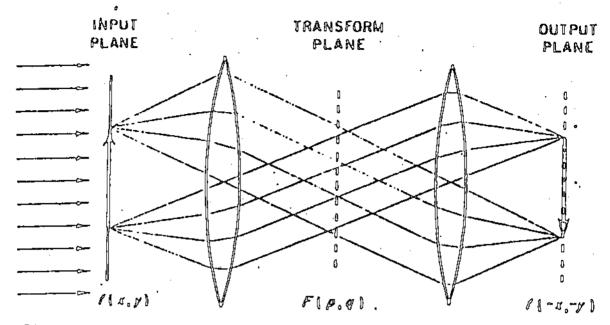


stowns 1.

A hypothetical holographic experiment. The apparatus for a Fresnel hologram is shown in (a). The view as the film is being exposed is shown in (b) and (c); the view after the film and one object are removed in (d) and (e); and the view after the developed hologram is returned, in (f) and (g), demonstrating the ghost unage of the missing object, (Adapted from Collies & Fernington, 1966.)



vigues 2. Diagram of waves in creation (A) and accommendation (B) processes. Note how in (B) the litering of 8 through halogram (A + 8)A + 8)° results in waves including both A and 8.



Plane Coherent Wavelroni

FLOURE 1.
Fourier transforming property of a lens. Note how the second lens effects a second "reverse" transform, causing an image in the output plane like the original object in the input plane.

tem pictured in Figure 3, this transformed image occurs in the transform plane. The transformed image also lies in the first focal plane of the second spherical lens and therefore the Fourier transform of the transformed image appears in the third focal plane or output plane of the system. An elementary theorem of Fourier analysis tells us that the resulting output image of the system is precisely the input image, only it appears upside down and backwards.

This process may be stated more precisely as follows. In an optical system, wavefront patterns are of two distinctly different types: state and dynamic. The static patterns are the stored photographic images that are described in terms of their transmission coefficients for light as a function of position on the photographic film. We will let (x, y) represent the geometric coordinates of position and f(x, y) represent the transmission coefficient at the point (x, y). The function f(x, y) on a piece of film represents a static storage pattern for the optical system. The dynamic or active patterns are the wavefronts of light that are processed and transformed by the optical systems. The specific transformations that occur in an optical system depend on the components (tenses, prisms, etc.), their placement, and the properties of the light waves used. For the optical system being described, Fourier transformation results because of the choice of spherical lenses, their arrangement, had the use of

coherent light. Let F(p,q) be the Fourier transform of f(x,y). Then for the optical system of Figure 3, F(p,q) is imaged in the transform plane. The variables p and q represent the position of the image in the transform plane. If we denote by —— the process of Fourier transformation, then we write

$$f(x, y) \longrightarrow F(p, q)$$

to represent the first stage of processing, and

$$F(p,q) \longrightarrow f(-x,-y)$$

to represent the second stage.

Now consider the process that occurs at the photographic film itself. A light wave of uniform intensity, say A, illuminates the image f(x, y). The intensities of the departure wavefront are attenuated by the presence of the input film. In fact, the intensity at point (x, y) is proportional to the transmission coefficient of the film at that point. The static storage pattern is converted to a dynamic processing pattern with amplitude distribution proportional to the spatial distribution of transmission coefficients of the image. This conversion is multiplicative. Thus, if the amplitude of the arrival pattern at point (x, y) is A, then the amplitude of the departure pattern in front of the image at point (x, y) is Af(x, y).

Now let us suppose that an actual photographic image of the transform F(p,q) could be placed in the transform plane of the optical system. Also suppose that the transform image could be illuminated by a plane coherent wavefront. (A uniform illumination across the transform plane can easily be produced by placing a point source of light in the center of the input plane.) The departure pattern produced by the film would have an intensity distribution F(p,q) and the image f(-x,-y) would be produced in the output plane. We find that both images f(x,y) and F(p,q) contain exactly the same information, the only difference is in the way in which information is coded. In fact, a film record of the transformed image, if made with a plane reference wave, is a Fourier hologram of the input image and from now on we refer to these transform images as holograms. They are the 'memories' of the optical system.

As described earlier, the statle images are intensity distributions whereas the dynamic images are amplitude and phase distributions. Static distributions can be represented by positive real functions whose value is less than unity (light is not produced and the phase of transmitted light is not changed by a piece of film), whereas dynamic patterns are represented by arbitrary complex (in the mathematical sense) quantities. In general, the Fourier transform of a positive real image is a complex quantity and therefore cannot be stored directly as a static pattern. To produce a hologram of the image f(x, y), we cannot simply expose a piece of film in the transform plane. However, by taking the superposition of the desired wavefront and a plane reference wave.

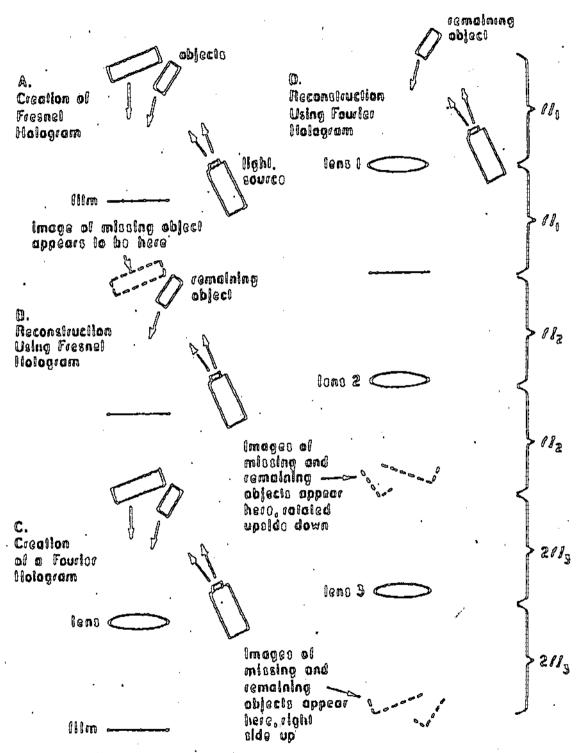
u hologram is produced in which the desired information is not lost. This is illustrated in Figures 4 and 5. The details of this process are similar to the lensless case, which was described earlier using a plane reference wave, and may be pursued in detail elsewhere by the interested reader (Stroke, 1966; Fippett, Berkowitz, Clapp, Koester, & Vanderburgh, 1965).

A second theorem from Fourier theory is important to the discussion that follows. When the Fourier transform F(p,q) of an image f(x,y) is multiplied by the complex conjugate of the Fourier transform $G(p,q)^\circ$ of a second image g(x,y), and the Fourier transform of the product taken, the result is the cross correlation of the-two initial images. This is important because cross correlation is a measure of the similarity of the original two images. A measure of similarity is precisely what is required for recognition.

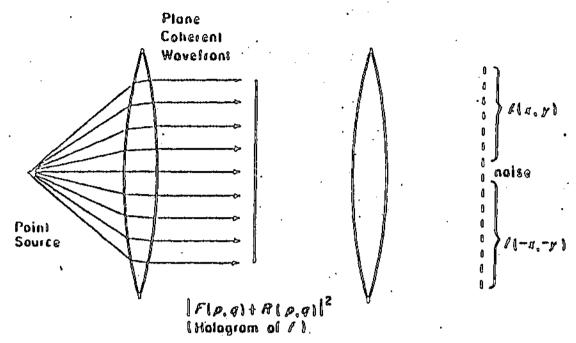
For our optical system, if the hologram (G(p,q) + R(p,q))(G(p,q) +R(p,q))° is placed in the transform plane while the image f(x,y) is placed in the input plane, the departure pattern that results just after the hologram is the product of the two functions F(p, q) and the hologram. The result that follows from the above theorem is that the cross correlation of the functions f(x, y) and g(x, y) is produced in the output plane. See Figure S. If the two images are similar, a bright spot appears in the output plane and the brightness of the spot indicates how similar the two images ore. The system instantaneously cross correlates two spatial patterns. In fact, this technique has been applied successfully to the instantaneous recognition of human faces! (In general, a thresholding light detector is placed in the output plane to determine whether or not a given input image should be "recognized.") The reader should note that the hologram is formed from the interference pattern between the desired transform G(p,q) and a plane coherent wavefront R(p,q). The result is that both the cross correlation and convolution functions of the two images are formed in the output plane. Figure 6 illustrates the formation of these functions by the optical system, and Figure I gives a geometric Interpretation of the cross correlation and convolution functions. Also note the similarity between this system and recognition holography described carlier.

There is another property of the Fourier transform of an image that is of interest to the neurophysiologist. Each point of the transform indicates the presence of specific spatial frequencies that are present in the input image. If, for example, the input image is a simple sinusoidally varying intensity pattern of spatial frequency p, then the Fourier transform at position F(p,0)—or I(0,p) depending on the orientation of the image—would have intensity proportional to the brightness of the image, and the rest of the transform would have zero intensity. By analyzing the Fourier transform of an input image, one can determine its exact spatial components.

The analysis of the optical system is similar to the analyses both of conventional optical holography in which Fourier or Fresnel holograms are pro-



vioual 4. Comparison of apparatus and Images in creation and reconstruction using Frence- and Fourier-type holograms. Note that the third kens in (12) is optional, and simply invests the first reconstructed image.



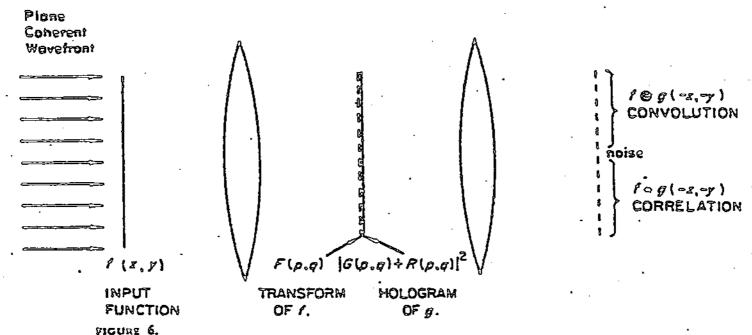
tractus 3. Reconstruction of f(x, y) using Fourier holography. Here the plane wave R(p, q) acts as the second, reference wave in capturing f(p, q) on the hologram. R(p, q) is the Fourier plane image of an off-centered point source. Reconstruction is effected here through use of an on-renter point source. This trick separates f(x, y) from f(-x, -y) in the output plane.

duced and of van Heerden's (1963) method of information storage in solids. There are also analogies between cross correlation in the optical system described here and in recognition (and in fact association) techniques both in van Heerden's system and in conventional holography. The reader is directed for the details to the references cited.

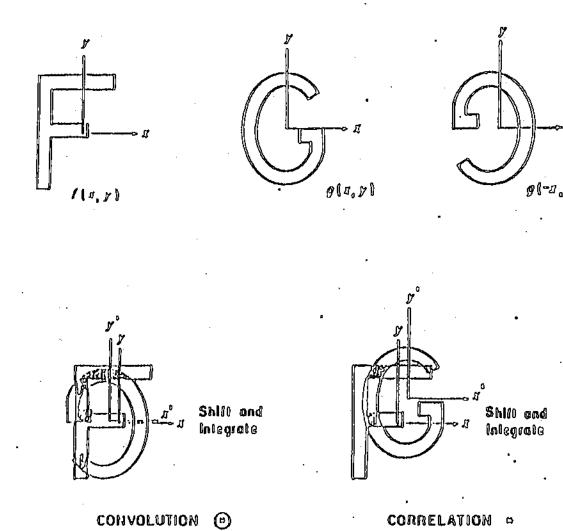
A Mathematical Network Model

through of the formal attributes of the holographic optical systems that we have found to have the essential properties for storage, recognition, and recall. What needs to be done to make holography into a useful metaphor for students of brain function in memory and perception is to see whether the lens system, or even an optical system, is necessary to the accomplishment of the holographic process.

We found earlier that there are both static (storage) and dynamic (processing) patterns in the optical system. We found that optical processing occurs because of the geometry and components of an optical system. We found how static and dynamic information patterns interact, and finally we found that a system in which there are two stages of linear information processing



Results from illuminating a Fourier hologram with a wave, F(p,q), which is not the same as the waves which created the hologram, G(p,q) and R(p,q). Convolution and correlation functions are formed in the output plane.



ricipes 7. Geometric interpretations of the convolution and correlation functions. Note how the convolution involves a rotated version of one function, while the correlation does not.

is capable of storing, recognizing, and recalling visual information. We now draw an analogy between each of these quantities in optical systems and corresponding quantities in computer simulated systems.

There are both static (storage) and dynamic (processing) patterns in our network model. The static patterns are sensitivity values of the junctional contacts between two cells. (For convenience we call our mathematical processing elements 'cells' and the lock of changes of sensitivity values 'junctions.') The junctions are assumed to be distributed throughout the volume of the assembly of cells. The contribution that a unit (not to be confused with a processing unit) of transmissivity has toward the rate of activation of the postjunctional cell may differ depending on the geometric position of its arrival. We therefore define the local sensitivity at point (x, y, z) of a cell to be a measure of the effective contribution toward activation that is produced

of position (x, p, x) from prefunctional cells. Static patterns are described by giving for each cell (p, q) of the network the local sensitivity function $k_{p,q}(x, y, x)$. Local sensitivity values are a function of the junctional microstructure of contacts among the cells.

In our model, dynamic or processing patterns are the patterns of activation in collections of processing cells. We assume for simplicity that the activation of a cell depends only on the average value of local sensitivity throughout its junctional microstructure. (This assumption is made here for expository purposes only. In a more comprehensive model by Baron, 1970, this assumption is not made.) We define the sensitivity of unit (p, q), denoted by K_{pq} , to be the volume integral over the microstructure of the local sensitivity function. That is,

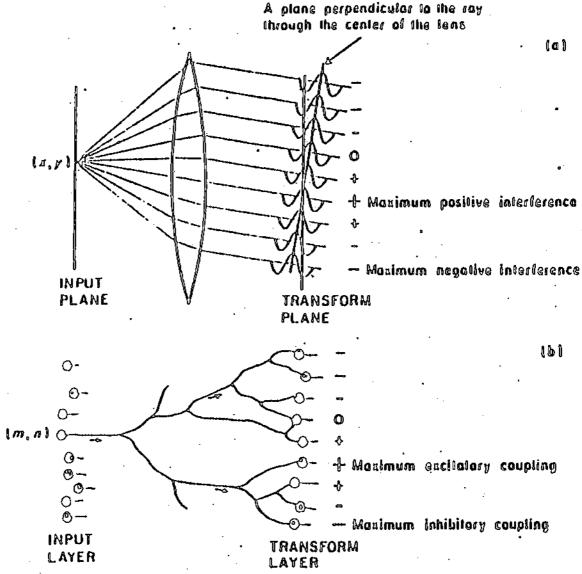
$$K_{00} = \iiint k_0 A x, y, s) dx dy ds.$$

These sensitivity values correspond to the transmission coefficients of the optical systems. Memories, the static stored patterns, are preserved in these constituity values as we now show.

The dynamic patterns process information from one collection of cetts to another. Dynamic patterns are described by giving the rate of activation and inhibition of the junctional contacts of each cell in the collection as a function of time. Dynamic patterns must be real-valued functions of time.

There is also a direct analogy between the transformations of the optical system as specified by its geometry and components, and the transformations of the network model. In the network model, patterns of excitation are transformed as they propagate from one layer of cells to another because of the coupling reefficients between cells. If, for example, one cell excites another at a high rate, the coupling between the two cells is high and positive. If the one cell tends to inhibit another, the coupling is negative. Coupling coefficients are functions of pairs of cells and must be specified for every connected pair of cells of a network.

Static and dynamic patterns interact in the network model in much the same way that they do in optical systems. In an optical system, light travels in a straight line and the 'coupling' is determined geometrically by the precise distance the light has to travel from its origin to its destination. In the optical system described above, the bradiating light wave is coherent, and therefore the phase of light transmitted from each point is the same. If the distance a light ray travels is a multiple of the wavelength, that ray adds; if the distance is an odd multiple of half the wavelength, the ray subtracts. These correspond to maximum exclusory and inhibitory coupling in the network model, respectively. See Figure' 8. Because the distance that the light has to travel varies systematically across the transform plane, the Fourier transform of the input pattern is formed in the transform plane. As noted, in the network



status 8.
Correspondence between optical and neural systems. The couplings in the neural system may be considered with respect to some positive baseline. Thus an 'inhibitory' coupling is in reality just the decrease in excitation below baseline level, and still could be excitatory in physiological terms.

model 'coupling' is determined by the amount of activity occurring in pairs of cells of an ensemble.

In the optical system, if the superposition of all fight waves has amplitude of all point (x, y), and the film at point (x, y) has transmission coefficient f(x, y), then the output pattern has amplitude Af(x, y). In the network, if the spatial summation of all activity results in a net excitatory quantity A at cell (p, q) in the network, the contribution of activity to the next cell is determined

by A and by the sensitivity K_{pq} of the receiving cell. The quantity A is not determined by coherence properties and it is at this point that a geometric analogy with the optical system fails. If the interactions between static and dynamic processes are multiplicative (which we will later propose) the analogy is direct, and the contribution from all prejunctional cells toward the activation of postjunctional cell (p, q) is given by AK_{pq} .

The resulting properties, storage, recognition, and recall depend on specific architecture, coupling coefficients, and sensitivity values, and are now to be studied in detail.

Achieving the Fourter Transformation

Consider the three-layered network illustrated in Figure 9a. We will call the cells in the three layers the input cells, transform or memory cells, and output cells, respectively. The activation patterns in these three collections of cells will correspond directly to the dynamic patterns which occur at the input, transform, and output planes of the optical system of Figure 1.

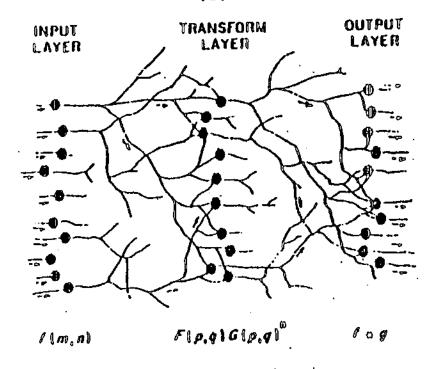
In order to make an analogy between the optical system and the network of Figure 3, we will need to choose a standard labeling scheme for the cells of the three collections. We will begin by assuming that cells are labeled with a two-dimensional system of variables. Thus a particular cell in a collection may be the (4, 6)th cell, or in general, we consider the (m, n)th cell. The activity of the (m, n)th cell at time t will be given by f(m, n)(t). This corresponds to the dynamic input image f(x, y) of the optical system.

For each pair of connected cells we must specify the coupling coefficient between that pair of cells. For the (m, n)th cell in the input layer and the (p, q)th cell in the transform layer we will designate by c_{p0}^{n} the coupling coefficient. (Superscripts refer to prejunctional cells, and subscripts refer to post-junctional cells.) We assume that the coupling coefficients do not depend on the geometric position (x, y, z) in the network. We must also designate a local sensitivity function $k_{p0}(x, y, z)$ for each cell (p, q) of the transform layer. (For the analysis that follows, the only cells for which the local sensitivity values are of interest are the cells of the transform or memory layer. We will therefore assume for simplicity that the local sensitivity values of all other cells have value I and may therefore be omitted from the discussion.)

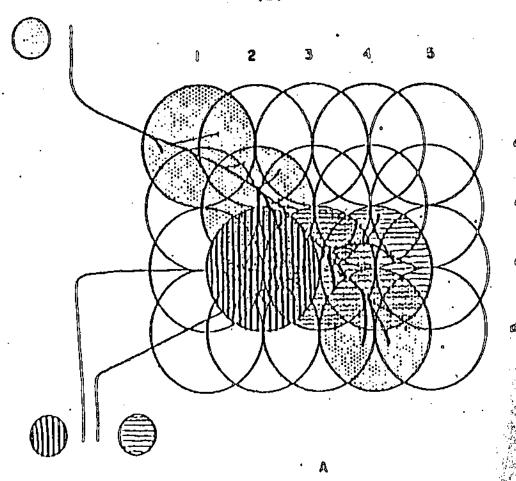
giquat 9.

In part a a neural network is described schematically. The network is analogous to the optical diagrams of Figure 3. Part b represents a more realistic diagram of microstructure of synaptic domains in corten. The ensemble of overlapping circles represented the junctions between branches of input axons and cortical dendrites. (Redrawn after Scheibel & Scheibel in Pribram, 1971.)





(b)



This network model corresponds to one basic computational element: one collection of cells that are closely coupled together. Our assumption is more precisely that every cell of the input layer unkes connections to every cell of the transform layer makes connections to every cell of the transform layer makes connections to every cell of the output layer. We will designate by $f(m, n \times 1)$, $f(p, q \times 1)$, and $f'(m', n' \times 1)$ the dynamic activity patterns in the input, transform, and output layers, respectively at time t. Since we designate by c_m^{-1} the coupling coefficient between the (m, n)th input cell and the (p, q)th transform cell, then the net amount of excitation less the amount of inhibition provided by all input cells at any point in the transform layer of the network is given by

$$\sum_{m=0}^{\infty} c_{n}^{m} f(m,n)(t). \tag{1}$$

The rate of activation of the (p, q)th transform cell is given by

$$F(p,q)(t) = \iiint k_{p_0}(x,y,z) \sum_{(m,n)} c_{p_0}^{-n} f(m,n)(t) dx dy dz, \qquad (1)$$

where integration would be over the volume containing all prejunctional connections of the transform unit. (In each case, if the rate of activation is negative, we assume the cell is inhibited and will not transmit at all.) Decause the coupling coefficients do not depend on position, the integration can be performed and Equation 2 reduces to

$$F(p,q) = K_{p0} \sum_{i=0}^{\infty} c_{po}^{-iq} f(m,n)(i).$$
 (1)

(In a somewhat more complicated model presented by Baron, 1970, the effective coupling between prejunctional and postjunctional cells during storage and recall depends systematically on time and geometric position (x, y, s) in the network. The result is that the local sensitivity values cannot be averaged. The addition of this "timing mechanism" enables the network to store and recall patterns that vary as a function of time. Thus, for example, storage recognition and recall of verbal (auditory) information is possible.)

In the optical system, specific transformations resulted because of the geometry and components of the specific system. Under the analogy, transformations that result in the network depend directly on the coupling coefficients, $c_{rd}^{n,q}$. We assume that the coupling coefficients are chosen so that the network Fourier-transforms the input image. To see how this might be done, we need only to look at the specific equations for Fourier transformation in optical systems. For the optical system, the Fourier transform is given by

$$F(p,q) = \iint \exp \left((px + qy) 2 \sigma t \right) f(x,y) \, dx \, dy. \tag{4}$$

If the system is discrete rather than continuous, Equation 4 becomes

$$F(p,q) = \sum_{n=0}^{\infty} \sum_{n=1}^{\infty} Af^{-n} \exp((px + qy)2\pi i/Af)f(x,y), \qquad (5)$$

where Mo is the number in discrete points of the image.

For our network model, if we could choose the coupling coefficients in the analogous way, that is,

$$c_{pq}^{mq} = M^{-1} \exp\left((mp + nq)2\pi i/M\right), \qquad (6)$$

the network would Fourier-transform the input pattern. However, we are not at liberty to use mathematically complex quantities. (Our early assumption was that dynamic patterns are nonnegative real functions.) In order to build a network model that resembles a biological neural network that will preserve the Fourier transform, we must preserve negative and complex quantities. In all optical systems described, negative and complex quantities were preserved by storing the intensity distribution of the superposition of the desired signal with a second (reference) light wave. Décause coherent light was used as a source of illumination, a stable interference pattern resulted whose intensity distribution enabled recall (reconstruction) of the stored pattern. One possible mechanism for a neural network to encode negative and complex quantities is to allow independent neurons to convey the positive and negative, real and imaginary components of the desired signal. This approach was used by Daron (1970). Another possible method is to have the cells activated at a background rate (which represents an actual value of 0), and to assume that inhibitory effects reduce a cell's activity to below background firing. These slow rates would represent negative quantities. The real and complex components of the transform would still have to be preserved by independent cells. Several researchers have also considered coherent neural activity in direct analogy with the optical system. In each case an internal signal is used to insure coherence. See for example Westlake (1968), Swigers (1967), and Darrett (1969).

The Neural Hologram

We note that the assumed neural processing that underlies this formalism is relatively simple. Chemical transmitters are released at synaptic junctions between presynaptic and postsynaptic cells. These chemical transmitters diffuse across the synaptic cleft and modify the resting potential of the post-synaptic cell. The local modifications to the resting potential near the synaptic junctions propagate to some extent away from the junctions. Excitatory transmitters cause depolarizations; inhibitory transmitters cause hyperpolarizations in the dendritic microstructure of the postsynaptic cell. Excitatory and inhibitory contributions sum by spatial summation, and the result is a small

contribution to the activation of that cell. The contribution to depolarization from any junction is determined by the local sensitivity at that junction that depends in turn on its membrane properties (see below). The local fluctuations, slow potentials, or minispikes interacting with the junctional microstructure, propagate toward the some and cause it to depolarize at a rate given by Equations 2 or 3. The result is a linear transformation of the input depolarization pattern f(m, n)(t) to the transform pattern F(p, q)(t) through one stage of neural processing.

In the optical system, the exposure of an undeveloped piece of film to an optical pattern sensitizes the film so that development at a later time causes points that are exposed to bright light (light with a large amplitude) to become black upon development, and points that are exposed to dim light (or no light at all) to remain transparent. If one then makes a 'positive' image of the film, the result is that the transmission coefficients of points that received light having a high intensity will become high, and the transmission coefficients of points that received no light at all will become low. This is the optical hologram of 'memory trace' of the optical system. By analogy, we suggest that in the model during the 'exposure' period the local sensitivity values become altered in regions that receive a large net amount of activation and become aftered in the opposite direction in regions that received a small not amount of activation. Some neurons store the 'real' part of the transform, other neurons store the 'imaginary' part. In particular, we propose that the local sensitivity values for the 'real' half of the population of junctions become proportional to the net (excitatory less inhibitory) amount of the activation that arrived during the exposure period, while for the 'imaginary' half we propose the opposite, i.e., deactivation (inhibition less excitation) occurs. That is,

$$K_{pq}$$
 is set proportional to $\sum_{(m,n)} (c_{pq}^{mn}) \mathcal{I}(m,n)(t)$, (7)

where f(m, n)(t) is the pattern to be stored. Because the net amount of arriving activation has a spatial distribution (Eq. 1) that is precisely the conjugate Fourier transform of the dynamic input pattern, the resultant distribution of local sensitivity values corresponds to one term of the optical Fourier hologram and it is for this reason that we call these 'memory traces' a neural hologram.

In short, a neural hologram is the pattern of sensitivity values that correspond to one element of an optical Fourier hologram and it is a function of the junctional microstructure of the memory units. The sensitivity values of a neural hologram preserve the conjugate Fourier transform of the patterns of excitation and inhibition that are initiated by the input.

Recognition and Recall

We have shown that Fourier transformation can be performed by a single stage of incoherent neural processing, at least in principle. We now entend our analogy between the optical system shown in Figure I and the three-injected network shown in Figure 9a. We assume that the transformation that occurs between the memory layer and the output layer is also a Fourier transformation. That is, the dynamic activation pattern of the output layer is the Fourier transformation of the dynamic pattern of the memory layer. Once this impoens, there are predictable patterns of activity in the output cells, and this activity takes on two distinctly different forms.

For an arbitrary input pattern (i.e., the encoded form of the sensory stimutus), the output pattern is the cross correlation of the input pattern with the pattern represented by the local sensitivity values. In this case, the system gives a strong signal when the input pattern is similar to the stored pattern. This is the recognition process. In the optical system of Figure I, the recognition signal appears as a bright spot in the output plane of the system. By analogy, the recognition signal for a memory node is the rapid firing of a small group of cells in the output fayer. The sensitivity pattern has a focusing effect on the surrounding activity, and when the input pattern is similar to the stored pattern, the recognition information is gathered at a group of cells. The activity in such cells could easily be monitored by other networks and used for selecting the memory nodes from which to later recall information. This process is completely analogous to recognition holography as described earlier, and to the cross-correlation process of the lens system.

By contrast, the output pattern has a completely different nature when the cells of the transform layer are excited in a uniform way. If the memory cells are uniformly excited, that is, by a single cell whose coupling coefficients are the same to all memory cells, the uniform activation is modified by the sensitivity values of the memory cells and the result is that each memory cell fires at a rate that is proportional to its sensitivity value. In this case, their pattern of depolarization is a reactivation of the stored pattern. It is, in fact, the conjugate Fourier transform of the pattern that is stored. The output pattern is in this case a copy of the original pattern of departure activity that was input when the local sensitivity values were established. The output pattern is a recalled copy of the stored information.

Alternate Models

The Fourier process is not the only process our formulism can ascfully describe. Equation I gives the pattern of activity of the units in the memory layer in terms of the input pattern f(m, n)(t), the coupling coefficients c_{pd}^{eq} .

and the sensitivity values k_{oo} . The coupling coefficients that enabled the network to Fourier-transform the dynamic patterns of activity were an arbitrary choice, and in principle any choice of coupling coefficients could be made. In order for the network to be able to store information, the input transformation should not lose information. It must have an inverse. If the sensitivity values of the memory units preserve the transformed pattern and the transformation between the memory units and the output units produces the inverse transformation, then a uniform excitation of the memory cells will cause the stored pattern to be recalled.

However, if the network is to adequately recognize input information, the recognition signal should be much stronger when the stored pattern arrives than when an arbitrary different pattern arrives. It is well known that for a linear system (Turin, 1960), the best possible recognition signal is achieved when the transfer function of the system is the complex conjugate of the Fourier transform of the pattern to be recognized divided by the frequency spectrum of the moise. This is precisely the result achieved by the holographic model presented in this chapter. (We have assumed for simplicity that there is no systematic noise in the network. If noise is present in the system, the sensitivity values must be divided by the spectrum of the noise.) Thus, although other linear transformations can support the storage and recall processes, they are not as ideally suited for recognition as the Fourier processes.

A very recent neural holographic model was proposed by Cavanagh (1972) In direct analogy with the associative holographic system described earlier. Cavanagh proposes that two independent collections of input units contact the memory units, and that two independent collections of output units carry the departure patterns from each computational element of the system. The initial input patterns, say a and b are converted to the transformed patterns A and B by the first fayer of processing, and the intensity pattern (A + B) (A -1- B)° is stored by the sensitivity values of the memory units. (The patterns A and B correspond to the arrival patterns A and B in the associative holography system described earlier.) The coupling between the memory units and the output units is assumed to have two properties: if no information is stored in the memory units, the departure patterns in the two collections of cells correspond to the arrival patterns. If, however, consulidation had occurred when arrival patterns A and B were present, then at a later time, input pattern a alone will cause the departure pattern b to be reconstructed, and vice versa. Thus, if a is present in the input, a 'ghost image' of b is automatically generated, and this reconstruction process is instantaneous. No search procedure or external control is necessary.

In contrast to our model, Cayanagh has demonstrated explicitly the two collections of input units that correspond to the two wavefronts of light used in optical holography. Also in contrast to our model, Cavanagh proposes that intensity values are stored rather than the real and complex quantities

of the transformed image. The essential point is that many models can be suggested, and each one captures one or more of the essential features of optical holography. At present, each model must be considered as a suitable alternative to the other, and until more direct experimental evidence is at hand, no model can be used to preclude another.

Many models fall into the class of neural holographic storage models, and it is this general class of models we are trying to describe here.

NEURAL PROCESSES

Storage Properties

Network models of the holographic processes have accomplished several aims:

- I. they have shown that optical systems are not required for holographic transformations to be realized; and
- 2. by keeping in mind some simple characteristics of nervous tissue, they have spelled out requirements for (a) temporary or permanent modification of any one layer of neurons by way of changes in the static patterns of sensitivity values within a junctional microstructure, and (b) processing between layers of neurons by specification of the coupling coefficients between them.

Let us begin by citing evidence consonant with the requirements for neural modification (storage, temporary, or permanent) demanded by the model. The model suggests that the arrival patterns to a unit—a neuron and its dendrites—produce a microstructure of slow potentials (depolarizations and hyperpolarizations) in the form of an interference pattern on that neuron's dendritic and somatic membranes. Eccles, Ito, and Szentagothai (1967) have described the mechanism of production of such interference patterns in the cerebellar cortex. At the regions of greatest constructive interference, changes are produced in the membrane's sensitivity to excitation and inhibition. Whenever new arrival patterns similar to the original occur, perturbations of these sensitive areas are produced. Thus the likelihood of depolarization and conduction in that cell are greater than when nonsimilar inputs arrive. It is precisely this, and only this, property that we have shown to be necessary for neural holography to be possible.

There is good reason to believe that a similar process occurs at the cerebral cortex. Benevento, Creutzfeldt, and Kuhnt (1973) have suggested on the basis of intracellular recordings that all input to the cortex results in excitatory (depolarizing) processes while the effects of horizontal interactions are essentially inhibitory (hyperpolarizing). Extracellular recordings, testing the effects of double simultaneous visual stimulation in our laboratory, are most readily

interpreted in the same way (Phelps, 1972). These results give some justification to the emphasis given in our strictly holographic model to a single process by which depolarization and hyperpolarization affect changes in sensitivity values. Neither hyper- nor depolarization per se are therefore considered agents for membrane modification.

That this effect is restricted in locus is not unlikely: a rough inverse square law for the effects of slow graded potentials has been assumed by model builders since Deurle (1956) and has sufficient backing of evidence (e.g., Phelps, 1972) to be taken seriously.

The holographic storage hypotheses require that the modified synaptic sensitivity be proportional either to the input signal strength (our model) or to the square of the input signal strength (Cavanagh). This could be accomplished in the following way. Permanent or some reversible semipermanent change to a membrane would be proportional to the square of the voltage difference between neighboring input patterns of electrical activity, since perturbation in the postsynaptic domain is a function of the differences in distribution of hyper- and depolarizations produced by the arrival of input patterns of presynaptic potentials, resulting in voltage differences parallel to the postsynaptic membrane. A testable physical description of such a mechanism has been developed by Richard Gauthier for the special, though not unusual, case of two synapses from different axons forming adjacent junctions into a dendritic or somatic membrane. The interactions between excitatory and inhibitory processes can be conceived to occur somewhat like this: with no synaptic input to the membrane, there is a resting electric potential across the membrane, with the voltage gradient or electric field lines perpendicular to the membrane surface (Fig. 10a). Suppose that in neighboring terminals the laput to the presynaptic terminals causes the postsynaptic junctions to become locally depolarized and hyperpolarized, respectively (Fig. 10b). The effect is to produce a pair of horizontally oriented electric dipoles at the surfaces of the cell membrane, which superimpose their electric fields on the vertical fields already present (Julesz, 1971; Barrett, 1969; Pribram, 1971). The net effect is to produce significant electric fields or voltage gradient components that are parallel to the surface of the membrane. We propose that these transient horizontal components of electric field trigger structural (e.g., conformational) changes in the membrane that outlast these horizontal fields. The induced structural changes in the membrane which in themselves may be reversible could then set off further biochemical processes leading to long-lasting ion permeability changes.

When either synapse is activated again, these structural or permeability changes can cause the effects (i.e., postsynaptic potentials) of one synaptic input to diffract and mimle the effects of the other, as if the latter were present. Thus the activation of one synapse produces the effects of activating both. The contribution of any such pair of synaptic inputs is small, but when many

PARTS 445 × 448 MISSING

convolutional integrals has been found to adequately describe the transformations that occur between the retinal receptor mosaic and the ganglion cell output from the eye to the brain.

The second quantitative description comes from Enroth-Cugell and Robson (1966). In their account they have demonstrated in the range of ganglion-cell receptive fields a variety of relationships between center and surround, a finding also emphasized by Spincili (1966). Enroth-Cugell, Robson, etc., then showed that they could explain this variety on the assumption of an opponent mechanism—separate excitatory (depolarizing) and inhibitory (hyperpolarizing) retinal processes, each process displaying an essentially Gaussian distribution.

Where do these separate excitatory and inhibitory processes take place? Dowling and Boycott (1965) have shown that, prior to the ganglion cell layer, few if any retinal neurons generate nerve impulses. All interactions are performed by way of slow potentials. These are of two apposing types—depolarizing excitations and hyperpolarizing inhibitory effects. Intracellular recordings (Svactichin, 1967) have suggested that the excitatory potentials are generated along the input transmission paths (bipolar cells) of the retina while the inhibitory potentials are due to the horizontal layers (amacrine and horizontal cells) that cross the transmission channels. This 'lateral' inhibitory process has been studied extensively and made the basis for quantitative descriptions of sensory interaction by Hartline (see Ratliff, 1965) and by von Dekesy (1959) in their treatment of the Mach band phenomenon. The equations they invoke are similar to those used by Rodieck in his description of ganglion-cell receptive fields.

The gist of these experimental analyses is that the retinal mosale becomes decomposed into an opponent process by depolarizing and hyperpolarizing slow potentials and transforms into more or less concentric receptive fields in which center and surround are of opposite sign. Sets of convolutional integrals fully describe this transformation.

The next cell station in the visual pathway is the lateral geniculate nucleus of the thalamus. The receptive field characteristics of the output from neurons of this nucleus are in some respects similar to the more or less concentric organization obtained at the ganglion cell level. Now, however, the concentric organization is more symmetrical, the surround usually has more clear-cut boundaries and is somewhat more extensive (e.g., Spinelli & Pribram, 1967). Furthermore, a second penumbra of the same sign as the center can be shown to be present though its intensity (number of nerve impulses generated) is not nearly so great as that of the center. Occasionally, a third penumbra, again of opposite sign, can be made out beyond the second (Hanunoud, 1972).

Again, a transformation has occurred between the output of the retina and the output of the lateral geniculate nucleus. Each geniculate cell acts as a

peophole "viewing" a part of the retinal image mosaic. This is due to the fact that each geniculate cell has converging upon it some 10,000 ganglion cell fibers. This receptive field peophole of each geniculate cell is made of concentric rings of opposing sign, whose amplitudes full off sharply with distance from the center of the field. In these ways the transformation accomplished is like very near-field optics.

Police, Lee, and Taylor (1971), though supportive of the suggestion that the visual mechanism as a whole may function in a Fourier-like manner, emphasize that the geniculate output is essentially topographic and punctate, is not frequency specific, and does not show translational invariance—i.e., every illuminated point within the receptive field does not produce the same effect. Further, the opponent properties noted at the retinal level of organization are maintained and enhanced at the cost of overall translational invariance. Yet a step toward a discrete transform domain has been taken since the output of an individual element of the retinal mosaic—a rod or cone receptor—is the origin of the signal transformed at the lateral geniculate level.

When the output of lateral geniculate cells reaches the cerebral cortex, further transformations take place. One set of cortical cells, christened 'simpic' by their discoverers (Illubel & Wiesel, 1968), has been suggested to be characterized by a receptive field organization composed by a literally linelike arrangement of the outputs of lateral geniculate cells. This proposal is supported by the fact that the simple-cell receptive field is accompanied by side bands of opposite sign and occasionally by a second side band of the same sign as the central field. Hubel and Wiesel proposed that these simple cells thus serve as line detectors in the first stage of a hierarchical arrangement of pattern detectors, Police et al. (1971) have countered this proposal on the basis that the output from simple cells varies with contrast luminance as well as orientation and that the receptive field is too narrow to show trans-Intional invariance. They argue, therefore, that an ensemble of simple cells would be needed to detect orientation. They suggest that such an ensemble would act much as the strip integrator used by astronomers to cult data from a wide area with instruments of limited topographic capacity (as is found to be the case in lateral geniculate cells). Whether in fact strip integration occurs. the linelike arrangement could be conceived as a preparatory step in Fresnel, Fourier or other frequency-type processing—now a slit rather than a peophole 'views' the retinal mosaic.

But it is not necessary to view simple cells only as way stations in a hier-archy—these cortical units clearly have functions in their own right. A series of ingenious studies by Henry and Bishop (1971) have confirmed that these simple cells are exquisitely tuned to the edges (luminance contrast) of lines in the visual receptive field independent of line width. Some are tuned to the leading, some to the trailing edges. Responses are of two types, excitatory and inhibitory, and very often show apponent properties: i.e., when the edge is

moved in one direction across the receptive field the effect (e.g., excitation) is the converse of that (e.g., inhibition) produced when the edge is moved in the opposite direction. These investigators have shown that this effect is binocularly activated. Only when the excitation zones are in phase is an output signal generated. This occurs exclusively when the image on the two actions superimposes—i.e., when 'objects' are in focus. Thus, simple cells act us gates that let pass only binocularly fused information.

Another class of cortical cells has generated great interest. These cells were christened 'complex' by their discoverers, I lubel and Wiesel, and thought by them (us well as by l'ollen) to be the next step in the pattern recognition blerarchy. Some doubt has been raised (Holfman & Stone, 1971) because of their relatively short latency of response as to whether all complex cells receive their input from simple cells. Whether their input comes directly from the geniculate or by way of simple cell processing, however, the output from complex cells of the visual cortex displays transformations of the retinal input, characteristic of the holographic domain.

A series of elegant experiments by Fergus Campbell and his group have suggested that these complex cortical cells are spatial-frequency sensitive elements. Initially, Campbell showed that the response of the potential evoked in man and cut by repeated flushed exposure to a variety of gratings of certain spacing (spatial frequency), adapted not only to that fundamental frequency but also to any component harmonics present. He concluded therefore that the visual system must be encoding spatial frequency (perhaps in Fourier terms) rather than the intensity values of the grating. He further showed that when a square wave grating was used adaptation was limited to the fundamental and its third harmonic as would be predicted by Fourier theory. Finally, he found neural units in the cut's cortex that behaved as did the gross potential recordings.

Pollen (1973) has evidence that suggests that these spatial-frequency sensitive units are Hubel and Wiesel's complex cells, although both his work and that of Maffei and Fiorentini (1973) have found that simple cells have the properties of spatial frequency filters, in that they are broadly sensitive to a selective band of spatial frequencies. In addition, the latter investigators have found that the simple cells can transmit contrast and spatial phase information in terms of two different parameters of their response; contrast is coded in terms of impulses per second and spatial phase in terms of firing pattern.

The receptive field of complex cells is characterized by the broad extent (when compared with simple cells) over which a line of relatively indeterminate length but a certain orientation will click a response. Pollen demonstrated that the output of complex cells was not invariant to orientation alone—number of lines and their spacing appeared also to influence response. He concluded, therefore, as had Fergus Campbell, that these cells were spatial frequency sensitive and that the spatial frequency domain was fully

uchieved at this level of visual processing. Additional correlating evidence has recently been presented from the Paviov Institute of Physiology in Leningrad by Clieser, Ivanoff, and Tscherbach (1971), who relate their findings on complex (and hypercomplex) receptive fields as Fourier analysess to the dendritic neurostructure of the visual cortex much as we have done here.

There is, however, still another set of problems that must be disposed of before the conception of a spatial frequency transformation of the retinal image by cortex can be accepted. These problems deal with the tuning characteristics of each spatial frequency sensitive element and the extent of retinal image which this element transforms. The evidence to date suggests that each simple cell is rather broadly tuned, but that the receptive fields of complex cells are narrower in their tuning characteristics. Richards and Spitzberg (1972), among others, have suggested the pattern recognition mechanism be considered analogous to that which obtains for color where stepwise recombination of opponent processes sharpens broadly tuned receptor characteristics into a magnificent tool for subtle color perception. We have already detailed the evidence that shows opponent processes to exist at various levels of the visual pattern transformation mechanism. What remains to be done is to show quantitatively how by combinations of opponent processes, sharper tuning characteristics of the spatial frequency mechanism can be achieved.

The problem is not much different for obtaining the greater visual angle over which the spatial frequency mechanism must operate. Some combinational process must occur—the question is, where and how. Evidence pertaining to this point is presented toward the end of this chapter, but first let us take leave from the transform process and look again into the distribution mechanism without which the brain teston effects cannot be explained.

A Limiting Specification

We thus have evidence that neural transformations occur in the visual system that could, given appropriate otorage, result in Fresnel-like (simple cell) and Fourier-like (complex cell) holograms. But all problems are not yet out of the way. Perhaps the most critical current question that is posed by the holographic hypothesis of memory storage is the question of the extent to which input becomes distributed at any one stage of processing. As indicated in an earlier part of this chapter, visual inputs even to the complex cell level of the cortex still represent only a few degrees of visual space. Obviously, input does not become distributed over the whole brain in one pass, if ever, What are the limits on distribution that would yet allow one to use the hologram as a model? This question may not have a single answer and probably depends on the coding and control mechanisms available to the organism for this purpose at any given moment. There is considerable evidence from verbal

learning experiments that rehearsal accomplishes internal distribution of the events rehearsed (Voss, 1969; Trabasso & Bower, 1968). Neisser (1967) points out that retinal translation should destroy the congruence necessary for recognition were a simple point-to-point template involved and it does not. Moyer (1970) on the other hand, in a series of experiments, has shown that recognition at a nonexposed retinal locus is impaired when a complex unfamiliar pattern is presented tuchistoscopically once to a restricted retinal locus. Even a single repetition of the exposure with no change of locus will, however, significantly enhance recognition at a distant locus. Rehearsal is obviously a potent source of distribution of information.

Another way of approaching this question is to ask just how much replication and how much distribution is demanded by the holographic hypothesis. Pollen's data suggest that in the visual cortex only small regions of visual space become encoded. However, Hubel and Wiesel (1968) describe considerable overlap of receptive fields within a single penetration for cells at the same orientation and preferred slit width. How this overlap becomes usefully integrated is an experimental question under present laboratory study (Pribram, 1974) and is discussed to some extent in the final section of the chapter.

The evidence cited above thus gives strong support to the concept that local regions of the cortex are responsible for storing the memories of experience. Does this contradict the results of the ablation studies cited earlier? No, for it is quite possible as already noted, that the proposed system, in response to rehearsal, stores multiple copies of the same experience in remote regions of the cortex, and that each of these records is a complete description of the given event. Evidence that in fact such multiple copies occur has been obtained: small macroelectrodes were implanted over the visual cortex of monkeys and electrical activity recorded in a discrimination experiment. In randomly distributed locations over the visual cortex, localized electrical activity was reliably found to be related to either the stimulus or the response (or reinforcement) events in the experiment (Pribrum, Spinelli, & Kamback, 1967).

Let us suppose that the strictly holographic transformations are a local phenomenon, and that integration of information across the cortea is done either by a hierarchical summing onto the next level of connectivity, or by a parallel processing mechanism via subcortical connections. These alternatives are being explored experimentally at present (see Pribram, 1974, for review). Since the hierarchical alternative is almost universally esponsed, it needs little explanation. Let us for a moment therefore consider the advantages of parallel processing mechanisms that make the experimental investigation worthwhile.

Relation to Control Processes

An essential distinction between a hierarchical serial process and a parallel process is that the latter allows control functions to be exercised before or at the transform or memory plane rather than after transformations and storage have been accomplished. In their operations, control functions can influence several biological memory processes that are independent of the particular storage model, yet each one can crucially affect the way in which the memory store operates. They are the following.

- 1. Permanence: Are the memory traces permanent (they never change once they are made), or are they temporary (the same memory location cun store a different pattern at later times)?
- 2. Aladifiability: Are the memory traces adaptive—do they change slowly after the initial consolidation process? If so, when are the modifications allowed to occur?
- 3. Consolidation: Do the memory traces become permanent upon a single presentation of the input pattern, or do they slowly become established upon repeated presentations of the same input information? If the latter is true, what are the control processes to insure that only the proper memory traces are allowed to consolidate at any given time?
- 4. Interference: If more than a single presentation of the input pattern is necessary, how sensitive is the nascent memory trace to interference?
- 5. Decay: Are the memory traces actually 'permanent' after they are established, or do they slowly degrade as a function of time? If they degrade, what is the rate of degradation? How does degradation affect recognition and recall?

Several memory control functions can operate on these storage functions. These control functions are not particular to the holographic model and in act are not part of the holographic model. They are the following.

- 1. Start-stop: When will a particular memory location begin the 'caposure' process, and when will it finish? This start-stop process may be considered for a particular memory location, or it may be considered for all the memory locations in a memory store.
- 2. Selection: For a memory store, the question may be restated: why with one memory location store the current input information rather than another? The memory may be stored initially on the basis of innate competences of the neural tissue involved and later on the basis of changes in competence produced by experience. Or, memory locations may be arbitrarily (pseudorandomly) selected initially and then input deliberately channeled to compatible sites on the basis of temporal or conceptual contiguity.
- 1. Recognition threshold: An input is judged as "recognized" when it evokes associated memories. The implication is that whenever memories are called

forth from an associative (content addressable) store, then the effects of an earlier similar input must indeed already be in the memory store. How similar must similar be to become recognized by the memory? The simulation of the holographic process developed here proposes that a cross correlation take place in parallel with all stored memory. Then if any region of any cross correlation exceeds some arbitrary threshold, the input is considered recognized. How is that threshold established and aftered?

4. Recall: If storage is truly associative, recall occurs to the eatent that an input evokes the effects of earlier associated inputs. How are the limits placed on such evocution? A control mechanism must be present that is able to decide which of the memory locations that have been addressed are to be used in further processing.

The above questions and comments are typical concerning the relationship between memory and its control functions and some of these have been touched on by Spinelli (1970) in the discussion of "Oceam." What is important to understand is that the general holographic hypothesis is strictly a model of the interaction between storage and input processing and does not address these decisional questions. However, the holographic hypothesis does propose that these decisional properties involving storage, recognition, and recall occur by way of a content addressable parallel-processing mechanism. We therefore consider briefly the evidence that, in fact, the organization and processing of memory occurs in a content addressable parallel-processing system.

There are a growing number of experiments that are designed to determine whether or not the recognition of a climulus and the retrieval of related (associated) information are sequential or parallel processes. Results presented by Sternberg (1969) and more recently by Atkinson and Juoia (1974) suggest in fact that both sequential and parallel processes are involved. It is generally believed true that there is an initial sequence of processes that encode the stimulus information into the form used by the memory stores. There next appears to be an initial addressing of memory that results in a familiarity index' for the stimulus involved. This is the recognition signal or correlation value produced by an associative memory store. The reaction time studies indicate that this is a parallel process—that the encoded stimulus pattern is presented simultaneously to all memory locations (nodes or basic computational elements) of the memory stores.

On the other hand, the experiments indicate that once the familiarity of a stimulus item is obtained, additional information can only be recalled by a secondary sequential search of memory. This secondary search process locates specific items from among those which are related to the stimulus. The sequential search appears to be restricted to those memory locations that contain information associated with the stimulus—memory locations that recognized the encoded stimulus pattern. Furthermore, these experimental

results imply that individual memory locations may be accessed independently and their stored information recalled.

EURIRIARY

Our studies of holographic processes have detailed possible mechanisms for the distributed memory required by the results of experiments on brain function. Several issues became clarified. Holography depends on two separable functions: (a) storage of interference patterns or their equivalents; (b) patterns created by superposition or other Fresnel- or Fourier-like input processing. A network model of one limiting case of holography—the Fourier Hologram—has been accomplished using elements and junctional characteristics plausibly like those in neural networks. Other models have also been touched upon. In addition, evidence has been addited that, in fact, holographic storage and spatial frequency—Fourier-like—processing occurs in the visual system.

Thus the advantages of a holographic memory as a model for brain function in perception can be fruitfully pursued with vigor. Aside from the property of distributed storage, holographic memories show large capacities, parallel processing, and content addressability for rapid recognition, associative storage for perceptual completion and for associative recall. The holographic hypothesis serves therefore not only as guide to neurophysiological experiment, but also as a possible explanatory tool in understanding the mechanisms involved in behaviorally derived problems in the study of memory and perception.

ACKNOWLEDGMENTS

We wish to thank Richard Clauthier, Erich Sutter, Arthur Lange, and Charles Stromeyer for their assistance and critical comments which enlivened considerably the writing of this paper.

REFERENCES

Atkluson, R. C., & Juois, J. F. Search and decision processes in recognition memory. In D. 11. Krantz, R. C. Atkluson, R. D. Luce, and P. Suppes (fids.), Contemporary developments in mathematical psychology. Vol. 1. Sun Francisco: Freeman, 1974.

thuron, R. J. A model for corrical memory. J. of Mathematical Psychology, 1970, 7, 27-59.

United. T. W. The cortex as interferonicies: The transmission of amplitude, frequency and place in cortical structures. Neuropsychologia, 1969, 7, 115-148.

- Benevento, L. A., Creutsfeldt, O. C., & Kulint, U. Significance of Intracortical inhibition in the visual cortex: Data and model. Nature, 1971, in press.
- Deurle, R. I. Properties of a mass of cells capable of regenerating pulses. Philotophical Transactions of the Royal Society of London, Ser. D, 1956, 240, 55-94.
- Cavanagh, J. P. Holographic processes realizable in the neural realin: Prediction of short-term memory and performance. Unpublished doctoral dissertation, Carnegie-Mellon University, 1972.
- Chow, K. L. Dioelectrical activity of isolated cortex, III. Conditioned electrographic responses in chronically isolated cortex. Neuropsychologia, 1964, 2, 175-187.
- Chow, K. L. Visual discrimination after ablation of optic truct and visual cortex in cats. Brain Res., 1968, 9, 361-366.
- Collier, R. J., Burckhardt, C. B., & Lin, L. H. Optical holography. New York: Academic Press, 1971.
- Collies, R. J., & Pennington, K. S. Ghost imaging by holograms formed in the near field. 'App. Phys. Lett., 1966, 8, 44.
- Creutzfeldt, O. D. General physiology of cortical neurons and neuronal information in the visual system. In M. B. A. Brazier (Ed.), Brain and behavior. Washington, D.C.: American Institute of Biological Sciences, 1961.
- Dewson, J. H., HI, Chow, K. L., & Engle, J., It. Bioelectrical activity of isolated cortex, H. Stendy potentials and induced surface-negative cortical responses. Neuropsychologia, 1964, 2, 167-174.
- Dowling, J. E., & Boycott, B. B. Neural connections of the retina: Fine structure of the inner plexiform layer, Quant. Biol., 1965, 30, 393-402.
- Eccles, J. C., Ito, M., & Szentugothal, J. The cerebellum as a neuronal machine. New York: Springer, 1967.
- Enroth-Cugell, C., & Robson, J. G. The contrast selectivity of retinal ganglion cells of the cat. J. Physiol., 1966, 187, 317-352.
- Gabor, D. A new microscopic principle. Nature, 1948, 161, 777-778.
- Galambos, R., Norton, T. T., & Frommer, C. P. Optic tract lesions sparing pattern vision in cats. Experimental Neurology, 1967, 18, 8-23.
- Glezer, V. D., Ivanoff, V. A., & Tscherbach, T. A. Investigation of complex and hypercomplex receptive fields of visual cortex of the cut as spatial frequency filters. Vision Res., 1973, 13, 1873-1904.
- Goldscheider, A. Über die materiellen Veründerungen bei der Assoziationsbildung. Neurol. Zentralblatt, 1906, 25, 146.
- Goodman, J. W. Introduction to Fourier optics. San Francisco: McGraw-Hill, 1968.
 Hammond, P. Spatial organization of receptive fields of LQN neurons. J. Physiol.,
 1972, 222, 53-54.
- Henry, G. H., & Bishop, P. O. Simple cells of the strinte cortex. In W. D. Neff (Ed.), Contributions to sensory physiology. New York: Academic Press, 1971.
- Hoffman, K. P., & Stone, J. Conduction velocity of afferents to cut visual cortex: A correlation with cortical receptive field properties. Brain Res., 1971, 32, 460-466.
- Horton, L. H. Dissecration on the dream problem. Philadelphia: Cartesian Research Society of Philadelphia, 1925.
- Hubel, D. H., & Wiesel, T. N. Receptive fields and functional architecture of monkey striate cortex. J. Physiol., 1968, 195, 215-243.

- Julesz, B. Foundations of cyclopean perception. Chicago: University of Chicago Press, 1971.
- Julesz, B., & Pennington, K. S. Equidistributed information mapping: An analogy to holograms and memory. J. Opt. Soc. Am., 1965, 55, 604.
- Kabrisky, M. A proposed model for visual information processing in the human brain. Urbana: University of Illinois Press, 1966.
- Krast, M. S., Obrist, W. D., & Pribram, K. II. The effect of irritative lesions of the striate cortex on learning of visual discrimination in monkeys. J. Comp. Physiol. Psychol., 1960, 83, 17-22.
- Kuffler, S. W. Discharge patterns and functional organization of mammalian resina.

 J. Neurophysiol., 1951, 16, 37-69.
- Lashley, K. S. The problem of cerebral organization in vision. In, Biological symposia, Vol. 7. Visual mechanisms. Lancaster, Pa.: Jacques Cattell Press, 1942.
- Lashley, K. S. In search of the engram. In, Society for experimental biology (Great Britoin): Physiological mechanisms in unimal behavior. New York: Academic Press, 1950.
- Locb, J. Comparative physiology of the brain and comparative psychology. Science Series. New York: Putman, 1907.
- Mattei, L., & Florentini, A. The visual cortex as a spatial frequency analyzer. Vision Res., 1973, 13, 1255-1267.
- Mountcastle, V. II. Modality and topographic properties of single neurons of cat's somatic sensory corten. J. Neurophysiol., 1957, 20, 408-414.
- Moyer, R. S. On the possibility of localizing visual memory. Unpublished doctoral dissertation, Stanford University, 1970.
- Nelsser, U. Cognitive psychology. New York: Appleton-Century-Cross, 1967.
- Phelps, R. W. Inhibitory interactions in the visual cortex of the cas. Unpublished doctoral dissertation, Stanford University, 1972.
- Pollun, D. A. Strinte corten and the reconstruction of visual space. In. The neuro-sciences study program, III. Cambridge, Mass.: MIT Press, 1973.
- Pollen, D. A., Lee, J. R., & Taylor, J. 18. How does the striate cortex begin the reconstruction of the visual world? Science, 1971, 173, 74-77.
- Preston, K., Ir. Use of the Fourier transformable properties of lenses for signal spectrum analysis. In I. T. Tippett, D. A. Berkowitz, L. C. Clapp, C. I. Koester, and A. Vanderburgh (Eds.), Optical and electro-optical information processing. Cambridge, Mass.: AIIT Press, 1965.
- Pribram, K. 11. Languages of the brain. Englewood Cilifs, N.J.: Prentice-Hall, 1971.
- Pribram, K. II. Why is it that sensing so much we can do so little? In, The neuro-sciences study program, III. Cambridge, Mass.: MIT Press, 1974.
- Pribram, K. H., Spinelli, D. N., & Kamback, M. C. Electrocortical correlates of stimulus response and reinforcement. Science, 1967, 157, 94-96.
- Rutliff, F. Alach bands: Quantitative studies in neural networks in the reting. San Francisco: Holden-Day, 1965.
- Richards, W., & Spitzberg, R. Spatial frequency channels: Blany or few? J. Opt. Soc. Am., 1972, 62, 1394.
- Rodieck, R. W. Quantitative analysis of cat retinal ganglion cell response to visual stimuli. Vishm Res., \$963, 5, 581-601.

- Rescaldate. F. Principles of neurodenamics: Perceptrons and the theory of brain mechanism. Washington, D.C.: Sputten Books, 1962.
- Rosenzweig, Al., Bennett, E., & Diamond, Al. Brain changes in response to experience. Scientific American, 1972, 2, 22.
- Spinelli, D. M. Visual receptive fields in the cat's reting: Complication, Science, 1966, 152, 1768-1769.
- Spinetti, D. M. Oceam, a content addressable memory model for the bruin. In K. II. Bribram and D. Broadbent (Eds.), The Bology of memory. New York: Academic Press, 1970.
- Spinelli, D. N., & Palbram, K. II. Changes in visual recovery function and unit activity produced by frontal and temporal cortex stimulation. *Electrosnooph. Clin. Neurophysiol.*, 1967, 22, 143-149.
- Steinberg, S. Memory-scanning: Mental processes revealed by reaction-time experiments. American Scientist, 1969, 57, 421-457.
- Stroke, G. W. An introduction to coherent optics and holography. New York: Assidenic Press, 8966.
- Svactichin, G. Horizontal and amacrine cells of retina-properties and mechanisms of their control upon bipolar and ganglion cells. Act. Cient., 1967, 10, 234.
- Swigert, C. J. Computational properties of a nerve and nerve net model. Unpublished doctoral dissertation. University of California, Berkeley, 1967.
- Talbut, S. A., & Afarshall, U. 10. Physiological studies on neural mechanisms of visual localization and discrimination. Amer. J. Ophthal., 1941, 24, 1255-1264.
- Tippett, J. T., Berkowltz, D. A., Clapp, L. C., Keister, C. I., & Vanderburgh, A. (Eds.) Optical and electro-optical information processing. Cambridge, Mass.: 811T Press, 1965.
- Frahusso, T., & Hower, G. H. Attention in fearning theory and research. New Yorks Viley, 1968.
- Turin, G. L. An introduction to matched filters. IRE Franzections on Information Theory, 1960, 6, 311-329.
- van Heerden, P. J. A new method of storing and settleving information. Applied Optics, 1963, 2, 387-392.
- von Bekesy, G. Synchronism of neural discharges and their demultiplication in pitch perception on the skin and in hearing. I. of Acoustical Society of America, 1959, 31, 338-349.
- Voss, J. F. Associative learning and thought: The nature of an association and its relation to thought. In J. F. Voss (Ed.), Approaches to thought. Columbus, Ohio: Meredith, 1969.
- Weiss, C. Principles of development. New York: Holt, 1939.
- Werblin, F. S., & Dowling, J. E. Organization of the setlan of the mud puppy, Nectures maculosus, II. Intracellular seconding, J. Neurophysiol., 1969, 32, 339-355.
- Wetner, G. The topology of the body representation in the somatic afferent pathway. In. The neurosciences study program, H. New York: Rockefeller University Press, 1970.
- Westlake, P. R. Towards a theory of brain functioning: A detailed investigation of the possibilities of neural holographic processes. Unpublished doctoral dissertation, University of Culifornia, Los Angeles, 1968.