

*Chapter VIII*

RECOGNITION OF VISUAL PATTERNS<sup>1</sup>

D. N. SPINELLI

The problem of how we identify objects, how we can tell books from tables or ashtrays from cigarettes is obviously a central one, not only in understanding perception, but the workings of the brain itself. The interest in object identification or in the recognition of visual patterns has recently been intensively stimulated by research in a variety of fields. Specifically, engineers and computer people have been working very actively on this problem because of the great market potential of reading machines; physiologists and psychologists have also invested a great deal of effort in an attempt to achieve a solution to this problem because of their interest in how real organisms work. It may be useful to point out from the start that at present there is no tenable theory in any field that would allow us to build an artifact capable of identifying objects, *i.e.*, of recognizing visual patterns, under circumstances that offer no problems for the average human subject. This being the case a number of psychologists, not to mention people from other fields, have directed their efforts toward definition of those problems that would have to be answered by such a theory.

Pattern recognition has to do with object identification over a number of transformations. First and simplest are transformations involving color, saturation of color, illumination and the dispersion of brightness. The second group of transformations has to do with movements of the object in visual space. The third set of transformations has to do with objects which are not rigid, for example, a dog or a cat; in this case, a geometrical description of the object and of the transformations of the object becomes improbable.

Much of the work done to date has dealt with the stability of object identification over the second of the three classes of transformations. For the engineers and the computer people this embodies the problem of optical character readers; for the physiologists and the psychologists this is the problem of how two-dimensional patterns are recognized in spite of translations, tilting or changes in the retinal size of the image.

---

<sup>1</sup> This research was supported by National Institute of Mental Health Grant MH 12970.

To put it very succinctly, the problem of pattern recognition is intimately intertwined with that of the perceptual constancies. Why is the world steady when our eyes move about; why doesn't it tilt when we tilt our heads; why don't our hands grow smaller when we move them away from our eyes? It could very well be that in attempting to solve these questions at the physiological level we would succeed in simplifying problems which are now apparently insoluble.

Hubel and Wiesel (6, 7) have suggested the following hypothesis to account for pattern recognition: cells in the visual cortex of cats and monkeys receive an input from cells of the lateral geniculate body which in turn are connected to a linear array of ganglion cells in the retina. These cortical units would therefore be optimally activated by a line with the appropriate orientation, positions and thickness in the visual world. These cells, which they named "simple," in turn would converge onto other cells named "complex." These other cells would be capable of detecting a line only if it had the appropriate orientation, but over a wide range of positions.

Baumgartner and others (2), on the other hand, have stated that as many as 50 per cent of the units that they have analyzed in the visual cortex of the cat did not show such selectivity to lines and their orientation, but had instead concentrically organized receptive fields. In our minds the presence or absence of concentrically organized receptive fields in the visual cortex is a crucial piece of information. If all cells in the visual cortex are line detectors or more complicated feature extractors we would be dealing with a nonlinear, relatively noiseless system of templates or detectors in which the basic building block is the line.

A different view is arrived at when the concentric organization of receptive fields is considered basic. The work of Hartline and Ratliff (4) has shown an essentially linear system to be operative in this case and, therefore, susceptible to linear system analysis. Such a system, of course, would have very different sets of properties from those of a detector mechanism. To all of this it should be added that lesions in brain structures that have no direct afferent anatomical connection with the retinogeniculostriate system, namely, the inferotemporal cortex, cause marked defects in visual pattern recognition (see Chapter IX).

The following work can be divided into two sections. The first part which is still being investigated has to do with two experimental attempts at understanding the physiological basis for size constancy and the constancy of the visual world in the face of head or whole body tilting. The second one has to do with the gathering of quantitative high resolution maps of visual cortex cells' receptive fields; this work is referred to here and is in the process of being published in greater detail (13).

## EXPERIMENTAL APPROACH

The method used in mapping the receptive fields has been described in great detail elsewhere (11). It is therefore only briefly summarized here. Adult cats were used exclusively in this experiment. Surgery was performed under a short acting anesthetic (thiopental sodium) which was injected in small incremental doses into the radial vein. While the animal was anesthetized, a small hole was opened in the skin, skull and dura so that a microelectrode could be placed at anterior zero, lateral 0.5 to 1.5 mm. of Horsely Clark coordinates. An agar gel was used to minimize brain pulsation. The trachea was intubated with a soft cannula through the mouth, and then all incisions and pressure points were infiltrated with a long acting local anesthetic (zyljectin) which when used with human patients has been found to be active for a period of over 1 week. The animal was then placed into a stereotaxic apparatus which leaves the visual field free and several hours were allowed to pass. Contact lenses were used to protect the eyes and to correct for accommodation. The optic disc and area centralis of both eyes were projected on a tangential screen placed at 57 cm. from the cat's eyes. At this distance, 1 cm. is approximately equal to  $1^\circ$  of arc. Mapping of receptive fields was done by moving a  $0.5^\circ$  contrasting spot in a scanning pattern over a  $25 \times 25^\circ$  area of the screen. This area was located in such a way that the area centralis for both eyes would fall on it. The spot would cover the  $25 \times 25^\circ$  area in 50 lines. Data collected during one line of scan was broken into 50 segments. Movement of the spot was controlled by a general purpose computer (PDP-8) which also counted the spikes generated by the unit being analyzed for every half degree of movement. Each map, then, consists of  $50 \times 50$  or 2500 data points. Scanning speed was capable of being varied, but it was found that for most units the most effective speed was between 5 to  $10^\circ$  per second. Most units were therefore mapped with a scanning speed of  $10^\circ$  per second. At the end of a map it was possible to ask the computer to display the matrix of data and to show all of those regions where the unit had fired at the specified level or more. The data is presented in the form of integral contour displays. Figure VIII.1 shows a block diagram of the experimental set-up.

Each unit detected with a microelectrode was first studied by moving bars and edges and different stimuli on the display screen in an attempt to define the most significant parameters capable of activating it. It was then mapped with a small spot either white or black on a contrasting screen with both eyes open using whatever scanning direction was found to be most effective. Each eye alone was then mapped in turn, using mostly four scanning directions, *i.e.*, from left to right, from right to left,

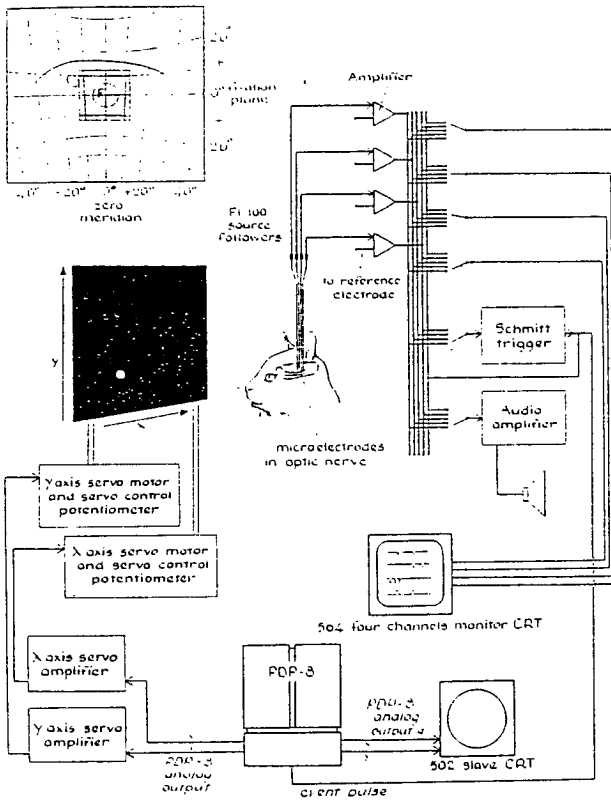


Fig. VIII.1. Flow diagram of the experimental set-up.

from bottom up and from up to bottom. Stimulus and background hues were then reversed and the procedure was repeated. A number of cells were also mapped by using bars instead of the small spot as stimulus. The great majority of cells, over 90 per cent, were found to be binocularly activated and could be divided into at least four classes on the basis of some of their most obvious characteristics. These characteristics were shape of receptive field when the unit was mapped in the direction for which it was most active, level of background activity and direction sensitivity.

### CONSTANCY

Jung (8) and others have described that about 50 per cent of the units of the visual cortex in the cat can be activated by vestibular stimulation. Murata, Kramer and Bach-Y-Rita (10) have also shown that 47 per cent

of the units can be activated by somatic stimulation and about 30 per cent by auditory stimulation. Our own work (12) has shown that highly specific information concerning auditory stimuli reaches visual cortex. It is our subjective experience that, when we tilt our head or when our whole body is tilted, the visual world is not tilted and, as already stated, Jung and others (8) have suggested that the function of the vestibular input to visual cortex is to inform the structure of body displacements. An experiment which immediately suggests itself is to search for a unit with a line-shaped receptive field and then to tilt the stereotaxic apparatus and remap it. If the cortical cell recorded from is wired permanently onto a number of ganglion cells whose receptive fields are all on a line, then the receptive field of the cortical cell should tilt, as much as the stereotaxic. If some other principles are at work, for example, a linear combination of visual and vestibular information, then one might expect that the orientation of the receptive field will be maintained. The result is that the receptive field of most cells with elongated receptive fields tilt just as much as the stereotaxic apparatus, but this did not happen for about 15 per cent of the cells. Figure VIII.2 shows one such unit in which the receptive field was mapped: in *1a* the unit receptive field was mapped monocularly with no tilt; in *2a* with a 30° tilt clockwise of the whole stereotaxic and cat complex. In *3a* it was mapped with no tilt. In *1b* the receptive field was again mapped with a 30° clockwise tilt, in *2b* with a 30° counterclockwise tilt and in *3b* with a 45° clockwise tilt. Size of spot: 0.5°; scanning speed: 10° per second; scanning direction: vertical; illumination: 20 lm. per m.<sup>2</sup>; level of integral contour display: 1 or greater. It can be seen that, although there are some displacements on the receptive field position, the orientation of the receptive field does not change. This was confirmed by linear regression analysis.

Figure VIII.3 shows another attempt at the problem of constancy. In row *1* and *2* the receptive fields of two different cells were mapped. In the *a* column the stimulus display was at the standard distance, 57 cm.; in the *b* column the distance was doubled to 114 cm. At this increase in distance the X and Y axes of the integral contour display covered only 12.5° of visual angle. The expectation is then that the diameter of the receptive field on the map would double, and again this is what happens for most units, but for a few units this is not the case. Row *2* shows the receptive field of one such unit mapped at the standard distance in *a*, and twice the standard distance in *b*. It can be seen that the size of the receptive field remains approximately constant. The tilting experiment could only be tried on the linear elements which were found to be 20 per cent of all units. A classification of the units analyzed follows.

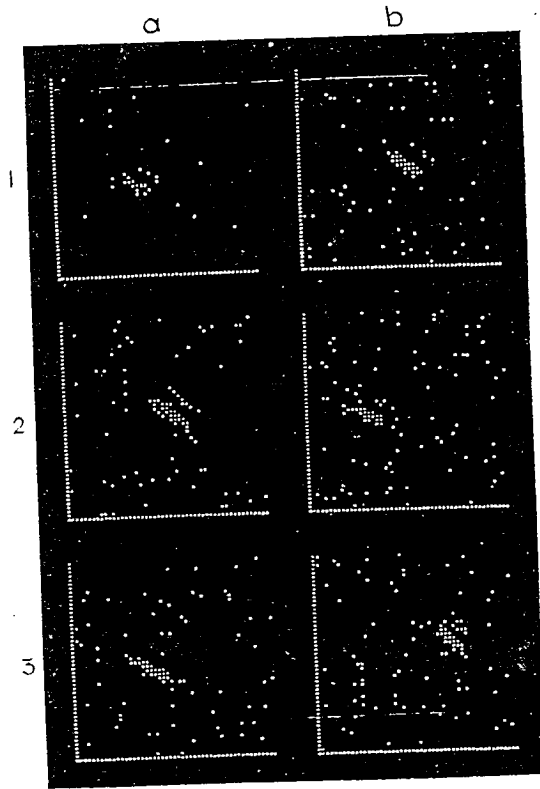
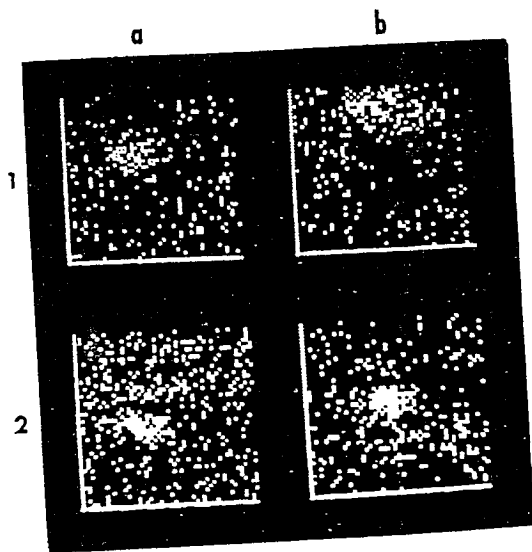


Fig. VIII.2. In *1a* the unit receptive field was mapped monocularly with no tilt, in *2a* with  $30^\circ$  tilt clockwise of the whole stereotaxic and cat complex. In *3a* it was mapped with no tilt. In *1b* the receptive field was again mapped with a  $30^\circ$  clockwise tilt, in *2b* with a  $30^\circ$  counterclockwise tilt and in *3b* with a  $45^\circ$  clockwise tilt. Size of spot:  $0.5^\circ$ ; scanning speed:  $10^\circ$  per second; scanning direction: vertical; illumination: 20 lm. per  $m^2$ ; level of integral contour display: 1 or greater. Notice the almost perfect stability of orientation of the long axis of the receptive field on the face of tilts as great as  $45^\circ$ .

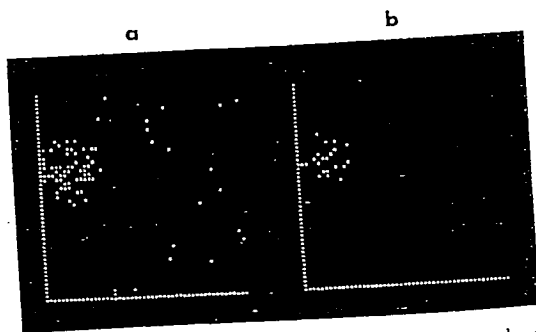
### CLASSIFICATION OF UNITS

*Bars and Edges.* Twenty per cent of the receptive fields of our units had elongated shapes as described by Hubel and Wiesel (6, 7). We have described elsewhere (13) that when mapped with a bar these units show the kind of specificity to direction of movement and orientation of the bar that can be predicted from the shape of the receptive field mapped with a small spot. Figure VIII.4 shows a representative example (see also figure VIII.2).

*Disc-Shaped Receptive Fields.* Forty-four per cent of all units had receptive fields that were classified as disc-shaped. Figure VIII.5 shows one



*Fig. VIII.3.* In *1a* the receptive field of this unit was mapped monocularly at the standard distance of 57 cm., all other conditions being as specified in Figure VIII.2; in *1b* the map was repeated with the display system at 114 cm.: notice that the apparent diameter of the receptive field doubles as expected from the geometry of the situation. In *2a* and *2b* the same procedure is carried out with a different unit and this time there is no increase of the receptive field diameter on the map. See text for further details.



*Fig. VIII.4.* This map is an example of a unit with elongated receptive field. In *a* the level of the integral contour display is 2 or greater. In *b* it is 3 or greater. Notice the oblique excitatory component flanked by two thin inhibitory areas.

such unit. In *a*, the unit was mapped with both eyes open. The two discs of the two receptive fields for the two eyes can be clearly seen; in *b* the left eye only, and in *c* the right eye only was mapped. The work of Bishop (3) and others has shown that the number of binocularly activated cells in the lateral geniculate is extremely small, less than 4 per cent. Hubel and Wiesel (5) did not find any binocularly activated cells in the genicu-

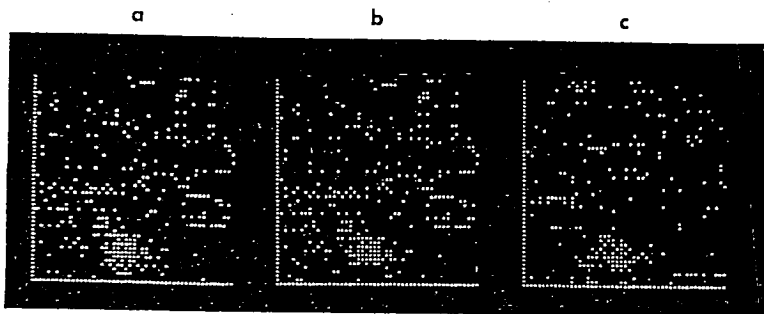


Fig. VIII.5. In *a* the receptive field of this unit was mapped with both eyes open; in *b* the left eye and in *c* the right eye only respectively were open. This unit shows a considerable amount of convergent disparity, as can be noted from the almost complete superposition of the receptive fields from the two eyes in spite of the fact that the eyes of a paralyzed cat diverge about  $5^\circ$ . Integral contour display level: 1 or greater.

late. By the criterion of binocularity and also by the criterion of shape of spikes (3) there is little doubt but that these units were cortical cells and not optic radiation fibers. It is also our experience that recording from fibers is in general extremely difficult and rare, even when we deliberately attempted it from white matter. Figure VIII.6 shows a receptive field of another unit; in this case the left eye was opened and the unit mapped three times. This figure gives an idea of the stability of the receptive fields.

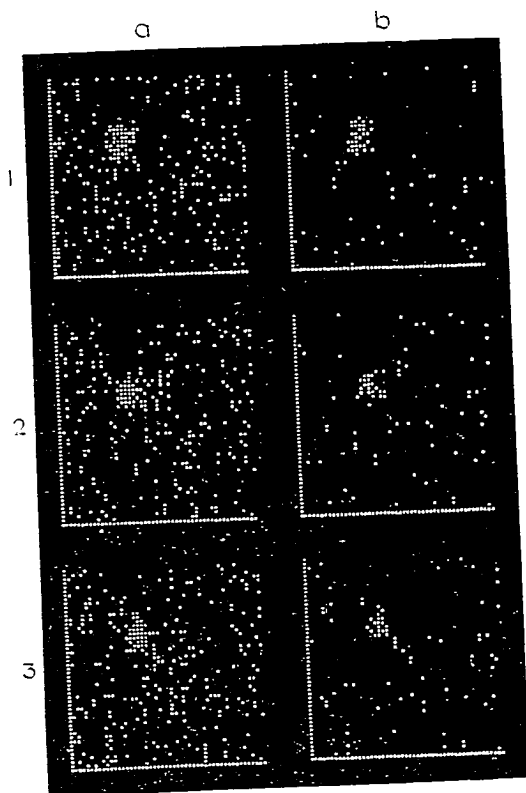
*Diffuse.* About 21 per cent of the units mapped had receptive fields whose center had a diameter of over  $10^\circ$ . The receptive fields of these units are somewhat similar to the disc-shaped ones, but there are some reasons for considering them a separate class. First, even though there was some variation in the sizes of the disc-shaped receptive fields, there seems to be a clear-cut size discontinuity between the disc-shaped and the diffuse ones. Second, the number of diffuse receptive fields which are also direction sensitive is greater for this class than for the disc-shaped ones. Third, the level of background activity usually, even though not always, is greater for this type of unit than for the disc-shaped ones.

*Composite.* A further 12 per cent of the units displayed receptive fields that were comparable in size to the diffuse ones, *i.e.*, over  $10^\circ$  in diameter, but differed from the last class in that the receptive field was formed by a number of subregions. In other words, the diffuse receptive fields are quite smooth regions of excitability, whereas the composite ones show a number of discontinuities within the receptive field.

## DISCUSSION

Hubel and Wiesel (6, 7) have provided us with a very clear theory concerning the initial processing undergone by visual stimuli. It is a theory





*Fig. VIII.6.* The receptive field of this unit was mapped monocularly three times at about 10-minute intervals. Level of integral contour displays: 1 or greater.

that requires a high degree of what, for convenience, we might call binocular coherence. What this means is that cells which are binocularly activated should have the same type of receptive field from both eyes and these receptive fields should be located in homotopic retinal positions. Also there should be no disc-shaped receptive fields. The theory as it stands has, therefore, a number of difficulties. First, the work of Barlow and his associates (1) has shown that, on the horizontal dimension, binocular disparity can reach as much as  $6.3^\circ$ , thus relaxing the condition of homotopicity. Second, although it is true that most of the time the receptive fields from the two eyes have similar properties, they are by no means identical and sometimes they are quite different (13). Third, although there is no question that receptive fields such as the ones in figure VIII.2 and VIII.4 are line-shaped and that bar will activate such units best, we should also not forget that the receptive field had been mapped by using a small spot. In other words these units do not respond exclusively to bars, but they also respond to spots as has already been

pointed out by Hubel and Wiesel (6, 7). They could only be considered to be line detectors, therefore, if they were to converge on other units with a very high threshold so that they would only fire when a line unit is maximally activated. To put it more simply, one should find units in the visual cortex of the cat that do not map to the small spot, but that only respond to lines. This is not the case, as all of the units that we detected could be mapped with the small spot. This does not mean that there are no detector cells in the visual cortex of the cat. It has been pointed out by other authors (9) that for an animal to possess efficient detectors has high survival value, and indeed it is very probable that cells that are sensitive to movement and direction of movement do just that, *i.e.*, alert the animal very quickly and without too much need of computational circuitry that something is moving out there. When it comes to the detection of patterns, this is another question. It is true that we do recognize a pattern, *i.e.*, a capital letter, independently of small translations and rotations or imperfections and even small mutilations. But it is also true that the translations, rotations or mutilations are also recognized. Thus it seems to be a possibility that the visual pathway could be endowed with two systems: a system of detectors very fast and efficient, capable of analyzing only a few parameters of immediate interest, and a system of uncommitted cells who are linear operators capable of conveying each and every nuance of a sensory experience. How would such a system put together the different parts of an image? A system of detectors first breaks an image apart in its components and then converges the feature extractors onto a common master cell. Here is where trouble arises from this kind of scheme. Once the integrity of the pattern has been destroyed it is difficult to know what to put together with what. Thus a tiny scratch on the wall has the same importance as the painting that is hanging on it. On the other hand, if the disc-shaped receptive fields are a part of a linear system, then it is interesting to take a look at the Hartline and Ratliff (4) equations. In computing the activity of one ommatidium, the activity of *all* other ommatidia must be known. Thus there is no need to break the image into its components and then rebuild it. From the very beginning each ommatidium is firing the way that it is firing because of the total pattern. The evidence that lateral inhibition is present in the retina, lateral geniculate and visual cortex of animals is considerable. It seems, therefore, that the same mathematical treatment and its implications could be properly applied to the mammalian visual pathway.

#### REFERENCES

1. BARLOW, H. B., BLAKEMORE, C. AND PETTIGREW, I. D.: The neural mechanisms of binocular depth discrimination. *J. Physiol.*, 195: 327-342, 1967.

2. BAUMGARTNER, G., BROWN, J. L. AND SCHULTZ, A.: Responses of single units of the cat visual system to rectangular stimulus patterns. *J. Neurophysiol.*, 28: 1-16, 1964.
3. BISHOP, P. O., BURKE, W. AND DAVIS, R.: Interpretation of the extra-cellular response of single lateral geniculate cells. *J. Physiol.*, 162: 409-431, 1962.
4. HARTLINE, H. K. AND RATLIFF, F.: Spatial summation of inhibitory influences in the eye of *Limulus*, and the mutual interaction of receptor units. *J. Gen. Physiol.*, 41: 1049-1066, 1958.
5. HUBEL, D. H. AND WIESEL, T. N.: Integrative action in the cat's lateral geniculate body. *J. Physiol.*, 155: 385-398, 1961.
6. HUBEL, D. H. AND WIESEL, T. N.: Receptive fields binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.*, 160: 106-154, 1962.
7. HUBEL, D. H. AND WIESEL, T. N.: Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *J. Neurophysiol.*, 28: 229-288, 1965.
8. JUNG, R. H., KORNHUBER, H. H. AND DA FONSECA, J. S.: Multisensory convergence on cortical neurons. Neuronal effects of visual, acoustic and vestibular stimuli in the superior convolutions of the cat's cortex. In *Progress in Brain Research, Brain Mechanisms*, edited by G. Moruzzi, A. Fessard and H. H. Jasper, pp. 207-234. Elsevier Press, Inc., Amsterdam, 1963.
9. LETVIN, J. Y., MATURANA, H. R., PITTS, W. H. AND McCULLOCH, W. S.: Two remarks on the visual system of the frog. In *Sensory Communication*, edited by W. A. Rosenblith, pp. 757-776. M.I.T. Press and John Wiley & Sons, Inc., Cambridge, Mass., and New York, 1961.
10. MURATA, K., KRAMER, H. AND BACH-Y-RITA, P.: Neuronal convergence of noxious, acoustic and visual stimuli in the visual cortex of the cat. *J. Neurophysiol.*, 28: 1223-1239, 1965.
11. SPINELLI, D. N.: Receptive field organization of ganglion cells in the cat's retina. *Exper. Neurol.*, 19: 291-315, 1967.
12. SPINELLI, D. N.: Auditory specificity in unit recordings from cat's visual cortex. *Exper. Neurol.*, 22: 75-84, 1968.
13. SPINELLI, D. N. AND BARRETT, T. W.: Visual receptive field organization of single units in the cat's visual cortex. Accepted for publication in *Exper. Neurol.*, January, 1969.