METHODS & DESIGNS

Visual receptive fields and clustering

EVANGELIA MICHELI-TZANAKOU

Rutgers University, Piscataway, New Jersey

A pattern recognition technique—clustering—has been used to analyze and evaluate meaningful characteristics of visual receptive fields in the frog tectum. The fields were mapped either by an automatic scanning technique or by a response-feedback method called ALOPEX. The data were then analyzed by the clustering technique, which separates the receptive field into iso-response regions. The latter can be checked on line by stimulating the eye with each cluster and with combinations of clusters. In this way, existing nonlinearities can be checked objectively.

The receptive fields (RFs) of cells in the visual pathway have been studied extensively and have revealed a wealth of information about the functions of the different visual areas. The organization and properties of RFs are well described for different species (Hubel & Wiesel, 1962, 1965, 1968; Keating & Gaze, 1970; Kuffler, 1953; Maturana, Lettvin, McCulloch, & Pitts, 1960). The responsiveness of cells to different light stimuli has been studied by means of microelectrode monitoring of cells in specific areas of the brain. RF shapes are then constructed either qualitatively (Hartline, 1940; Lettvin, Maturana, McCulloch, & Pitts, 1959; Spinelli, 1966) or quantitatively (Sasaki, Bear, & Ervin, 1971).

Depending on the stimulus characteristics to which the cell is most sensitive, the cells in the mammalian visual cortex are classified as simple, complex, or hypercomplex. In the experiments described in this paper, the animal was the frog. The frog's visual RFs do not have the center-surround organization that is present in the mammalian cortex. As Maturana et al. (1960) pointed out, criteria have to be devised in order to determine which points of the mosaic of light on the frog's retina constitute the boundaries between an object and the background.

In this paper, a method is presented by which visual RFs are analyzed by a clustering algorithm, one of the frequently used methods for classification of data. The main purpose is to separate the meaningful features in the RF from the background noise and to classify the responses of a cell as a function of the spatial distribution of brightness of the stimulus. This is well justified from the fact that the RF geometry may be drastically altered as a function of the stimulus luminance (Arden, 1963); both distortion in shape and changes in the nature of the response are encountered. The reason for

the complicated relationship between stimulus and response seems to be clear: Stimulation at any point of the RF results in the activation of both excitatory and inhibitory changes that typically differ in their magnitudes, latencies, and time courses, whereas the recorded outcome reflects some combination of these changes.

The method presented in this paper is designed to take into account these complications and to avoid as much as possible the preconceived notions of the experimenter concerning RF mapping and analysis.

METHOD

Frogs (*Rana pipiens*) measuring between 2.5 and 3 in. are anesthetized by immersing them in a warm water solution of Finquel (tricaine methylsulfonate 1% weight in H₂O). Boosters of 0.05-0.1 cc are used as needed. A flap is cut in the skin above the skull, leaving one edge intact at the caudal margin. The blood vessels are then carefully stripped back from the surface of the skull above the tectum. To expose the tectum, a small circular piece of bone is drilled out, with special care being taken to avoid cutting the large vessels that run below the skull and over the caudolateral margins of the tectum. The dura and the arachnoid membrane are then removed, and the brain is covered with a drop of mineral oil. At all times during surgery and recording, the frog is kept covered with a moist cotton pad. After the surgery, the frog is left to recover completely from the anesthesia. The eyes are carefully examined to determine whether they are healthy. The animal is then given 0.05-0.06 ml of tubocurarine (3 mg/ml Squibb); additional 0.01-ml doses are injected whenever needed. The frog is placed on a horizontal foam pad.

A tungsten microelectrode is positioned at the point above the rostral part of the tectum by using a placing microscope. A silver ball electrode (indifferent) is placed under the frog's body. The recording electrodes have tips of diameter 1-10 microns and impedances of 5-10 Mohms at 1000 Hz. The electrodes are moved by a hydraulic,

The author wishes to thank S. Albin and W. W. Welkowitz for critically reading the manuscript. The research was supported by NSF-ECS-82-05245. The author's mailing address is: Department of Electrical Engineering (Bioengineering), Rutgers University, P.O. Box 909, Piscataway, New Jersey 08854.

remotely operated microdrive. Relative readings of electrode depth are recorded.

The action potentials are recorded extracellularly in the upper layers of the tectum. The retinotopic map of Jacobson (1962) is used as a guide for placing the microelectrode on the surface of the tectum and choosing the CRT position. The CRT used for stimulus presentation is mounted on a custom-made stand so that the azimuthal and polar angles (ϕ, θ) can be read directly. The eye of the frog is at the origin of the spherical coordinate system. The face of the CRT is at all times perpendicular to the line of sight to the center of the screen. The distance r between the eye and the CRT is 20 to 30 cm.

Signals picked up by a single-ended field-effect transistor (FET) probe are between 50 and 300 microV, and the signal-to-noise ratio is about 10:1. After amplification to about 1 V, a discriminator is used to select pulse heights. Most of the accepted pulses are believed to have originated from single units (cells). In general,

47	35	45	20	35	13	28	40	35	25	24	25	24	22	38	13
33	19	20	33	39	27	25	61	52	56	29	34	27	23	27	21
42	38	42	16	44	55	35	55	25	79	45	23	23	6	17	14
36	32	34	47	27	53	56	47	47	27	33	23	30	17	23	17
6	5	29	39	43	71	57	41	46	46	54	31	29	23	23	15
39	26	48	34	45	82	82	40	98	108	55	58	23	25	42	13
33	38	58	30	72	78	64	94	57	59	86	34	44	26	39	12
43	59	61	84	48	82	128	64	71	108	58	37	39	18	13	15
54	13	53	86	107	152	157	97	215	56	82	47	41	20	25	21
43	29	58	93	76	228	217	180	91	85	89	55	40	2	22	5
57	76	85	108	195	208	218	179	104	82	40	56	21	26	23	12
58	72	92	134	218	215	232	136	87	50	33	51	19	31	17	17
79	98	89	117	141	121	111	14	87	45	35	23	35	12	17	16
75	45	101	73	134	68	59	28	77	55	7	40	30	28	15	11
92	76	69	96	61	76	41	21	66	38	29	8	9	17	9	15
77	47	60	37	42	54	58	32	27	29	30	19	25	4	4	10

(a)

(b)

Figure 1. (a) Response matrix of a retinal fiber in the frog tectum (Run 856). The numbers represent the number of spikes elicited when a 2×2 pixel bright spot was flashed ON and OFF on a dark background. The locations of the spot were selected randomly. Each location was covered only a preset number of times. The spot was moved every two pixels. (b) Linear transformation of the response matrix in (a) to an intensity pattern. The gray scale used has 16 levels.

larger RFs are observed when wider discriminator windows are used. This reflects the retinotopic character of tectal organization.

The visual stimuli are computer controlled. The detailed methodology has been described elsewhere (Tzanakou, Michalak, & Harth, 1979). Briefly, dots of variable luminance and size are presented on a CRT screen at randomly chosen locations and in such a way that each location is covered exactly a preset number of times. The responses of the cell are recorded and form a response matrix (Figure 1a). The response matrix is then linearly transformed into an intensity pattern corresponding to the RF of the neuron (Figure 1b). After the RF is mapped, a clustering algorithm, described in the next section, is used to separate the RF in a number of isobrightness regions and, hence, a number of isoresponse areas.

A second method for mapping the RFs is also used. This method, called ALOPEX, is a response-feedback method and is described in detail elsewhere (Harth & Tzanakou, 1974; Tzanakou & Harth, 1977; Tzanakou et al., 1979). It is an iterative method in which a grid of 32×32 picture elements (pixels), consisting of 16 levels of gray scale, is presented to the animal. The response of the neuron to the whole pattern is recorded and used as feedback to change the pattern of the next iteration. As the process continues, an optimal stimulus is created on the CRT screen; the process is guided by the increase of the response every time the intensities of the pixels are changed in the right direction. This optimal stimulus is the RF of the neuron under investigation. This method is most efficient in the case of the nonlinear behavior of the cell and is much more economical in time. Since the whole CRT screen is used as a stimulus, the response is not a function of the dot position; rather, it is a global response to the whole pattern presented to the eye.

RECEPTIVE FIELDS AND CLUSTERING

The relationship of response and stimulus is not linear (Easter, 1968; Stone & Fabian, 1968). This means that neural excitation and inhibition cannot simply be summed algebraically. Their relationship is, rather, a means of enhancing differences and of providing visual neurons with a greater dynamic range of response. Therefore, in the pattern representing the RF of a neuron (as described in the Methods section), every single pixel of intensity is of importance, since with the right combination of luminance and positioning with respect to other pixels, it affects the firing rate of the neuron under consideration. The method presented here uses every bit of information on brightness and spacing between the pixels of the RF to form the optimal stimulus for a particular neuron and to find most of the nonlinear characteristics in its behavior.

In order to classify patterns, key features must be extracted from the pattern. Clustering is one of the frequently used methods in pattern recognition. Many

algorithms have been developed in this direction (Friedman & Rubin, 1967; Ling, 1972; Shepard & Arabie, 1979). Clustering in most cases is done on the spatial distribution of the data. The algorithms search for a "compactness" of a parameter (in our case, the luminance) in the different areas of the pattern. When these regions of luminance are used as stimuli, one finds the functional properties of these regions.

Another technique used in image segmentation is thresholding. Its main usefulness lies in discriminating objects from backgrounds in many classes of scenes, including cloud covers, blood smears, chromosome spreads, and others. It is less appropriate in dividing textured images into meaningful regions, but it still has a variety of applications in situations in which the objects and background are homogeneous and smooth. A number of techniques have been proposed for automatic threshold selection; these include global threshold-selection techniques based on gray-level histograms (Doyle, 1962; Prewitt & Mendelson, 1966) or local thresholding methods (Bartz, 1969; Morrin, 1974; Sklansky, 1968; Ullman, 1974; Wolfe, 1969).

The data collected from frogs in the experiments described here are represented by a matrix, each pixel of which is assigned a brightness level. This matrix of intensities represents the RF of a neuron in the visual pathway. As such, the spatial organization and the brightness of the pixels are both important. The algorithm one uses should not, therefore, alter the RF organization; rather, it should check for similarities between the pixels with respect to brightness. The algorithm described below groups together pixels of approximately the same luminance.

CLUSTERING ALGORITHM

Since the visual input is analyzed by some processor that extracts local features, a pattern can be discriminated from a homogeneously illuminated background by the fact that the light flux of the different areas of the pattern is higher or lower than that of the background.

Consider an $L \times M$ matrix S and luminance ℓ_{ij} for the pixel (i,j) in the matrix (this matrix will henceforth be called the pattern).

Step 1: Find the average luminance $\bar{B}(S)$ of the pattern S , using the equation

$$\bar{B}(S) = \frac{\sum_{ij} \ell_{ij}}{N(S)} \quad i, j = 1, 2, \dots, 32. \quad (1)$$

$N(S)$ = number of pixels in matrix S .

Step 2: Divide the matrix S into two subgroups $S1$ and $S2$ so that $S1$ contains all pixels with $\ell_{ij} > \bar{B}(S)$ and $S2$ contains all pixels with $\ell_{ij} \leq \bar{B}(S)$.

Step 3: For each subgroup $S1$ and $S2$, find the "centers" corresponding to the maximum luminance ℓ_{ij} , $\ell_{\max}(S_n)$, $n = 1, 2, \dots$. Also, find the average luminance $\bar{B}(S_n)$ for each subgroup (cluster) using Equation 1, that is,

$$\bar{B}(S1) = \frac{\sum_{ij} \ell_{ij}(S1)}{N(S1)} \quad \text{and} \quad \bar{B}(S2) = \frac{\sum_{ij} \ell_{ij}(S2)}{N(S2)}.$$

Step 4: The original pattern S is now divided into three subgroups SA , SB , SC such that

$$\text{pixel } (i,j) \in SA \text{ if } \ell_{ij} > \bar{B}(S1)$$

$$\text{pixel } (i,j) \in SB \text{ if } \bar{B}(S2) \leq \ell_{ij} \leq \bar{B}(S1)$$

$$\text{pixel } (i,j) \in SC \text{ if } \ell_{ij} < \bar{B}(S2)$$

Step 5: Repeat Steps 3 and 4 as many times as necessary to find all meaningful features of the pattern and to test their effects on the firing rate of the neuron.

In general, starting with one group of data, namely the RF of the neuron under investigation, we divide it into a number of clusters. For iteration n of the clustering algorithm, one more cluster is added to the number of clusters available in iteration $n-1$. If K is the number of clusters at iteration n , then at iteration $n+1$ we have $K+1$ clusters (Friedman & Rubin, 1967).

EXPERIMENTAL RESULTS

The clustering procedure described above is now routinely used in the Vision Laboratory at Rutgers University as an extension of RF studies in the frog's visual pathway. After the RF of a neuron is mapped, on-line clustering is performed and the neuron is stimulated again, with each one of the clusters found and with combinations of them. A response is recorded for each region and then analyzed in order to establish existing linearities or nonlinearities of spatial summation (Tzanakou, Note 1).

The RFs mapped have, in general, a circular or elliptical shape (Figures 1 and 4). Two or three levels of clustering are usually enough for a complete separation of the data into functionally meaningful regions. This results in a total of no more than 10 clusters and their combinations. Statistical assumptions, appropriate to the clustering method, are used in order to reduce the number of clusters in an objective way. The first two subgroups formed from the RF of Figure 1 are given in Figures 2a and 2b, and further clustering produces those seen in Figures 2c and 2d. It is obvious that, even in Figures 2a and 2b, a gross distinction of the more sensitive areas versus the less sensitive areas has taken

(a)

(d)

(b)

(c)

(e)

Figure 2. Clustering performed on the original data of 1(a) revealed the two subgroups (a) and (b) as the center and the surround of the receptive field. Further clustering can further separate the most sensitive area of the receptive field—the brightest part of the pattern (c). The rest of the pattern if the center is removed (d). A combination of the outer surround and the center, with a dark annulus around the center (e).

Figure 3. Computer facsimile of a small RF (Run 405). One grouping was enough to give meaningful regions. (a) The whole RF mapped with a bright spot on a dark background. (b) The center of the RF after clustering. (c) The surround of the RF. Note that the most and least sensitive regions are well separated. Also, note that there exist some spots way outside the central region which clustering of the closest neighbors did not affect. These spots could very well be secondary centers in the RF.

place. The fine separation of the regions is more evident in Figures 2c and 2d. The most sensitive area (center) corresponds to the brightest part of the pattern.

In some cases, a single clustering is enough to produce two distinct regions. This is particularly true for small-size RFs (Figure 3).

Interactions of the different regions of the RF can be studied using any two or more clusters. This procedure can be thought of as one way of searching for the optimal stimulus for the neuron. The combination that gives the maximum response is the one that uses the center of the RF and its outer surround. This way a dark "annulus" surrounds the most sensitive area of the RF, and the outer surround appears as a brighter "annulus" for the dark area. This combination of clusters (Figure 2e) has given a higher response, a result that is reminiscent of the McIlwain "periphery effect" (McIlwain, 1964).

Some cells with RFs that have multiple centers, as shown in Figure 3, have been encountered. In these cases, usually there is a large center and one or more smaller areas of brightness similar to the large center (Figure 3b). When each one of the centers and their combinations are used as stimuli, the response is maximum when the large center and one of the small ones are presented together—the most effective small center being the most distant from the large center. This process again implies a synergistic effect between the center and the outer periphery. The advantage of clustering versus other methods used for RF analysis is that it does not distort the shape of the RF and uses the same intensities of the composite RF when using only parts of it. The function of the different parts of the RF can be studied in conjunction with other parameters, such as time. If for example, cluster S_k is presented first, cluster S_n can be introduced at a later time and the temporal properties associated with this variable delay time can be studied.

DISCUSSION

Two types of recognition of images have been described in the visual system: recognition of the form of an image and recognition of its spatial and temporal properties.

The method of analyzing receptive fields presented in this paper is capable of furnishing this kind of information. Texture discrimination is strongly dependent on clusters that are formed either by chance or by some ordered way. Any local feature is visually extracted by local detectors tuned to that particular feature.

The perceptual significance of clusters has already been demonstrated by Julesz and his collaborators (Julesz, 1981a, 1981b; Julesz, Gilbert, & Victor, 1978). Julesz, (1971) stated that "there is a cluster detector operating in the visual system, probably at the earliest levels, and only those statistical descriptors that describe cluster formations can have some perceptual significance" (p. 114). The clustering method of analyzing the

(a)

(b)

(c)

(d)

(e)

Figure 4. A large receptive field and its clusters, mapped by a bright spot of 2×2 pixels flashed ON and OFF on a dark background (Run 642). The intensity pattern, (b) After clustering, the most sensitive area of the field is well concentrated in the central area. (c) The immediate surround of (b). (d) The outer surround—least sensitive area of the receptive field. (e) A combination of (b) and (d). The part presented in (c) is missing so that a dark annulus surrounds the most sensitive area of the RF.

RF of single neurons in the visual pathway is one way to find how the neuron perceives intensity information with specific properties such as position, orientation, and size of the stimulus related to the differences of function in the various parts of its receptive field.

The statistical analysis of RFs of neurons is the same analysis that the visual system as a whole is doing when perceiving complex scenes of intensity information. One starts with a global analysis of the scene and then goes on with a more localized inspection of the various parts of the scene, looking for those elements that are of importance. The importance depends largely on the differences evaluated between part A and part B of the scene. As Julesz (1983) pointed out, the best texture segregation corresponding to human texture discrimination will be the outcome of a combination of several filters; this combination of filters would correspond to a "texton detector" (Julesz, 1983). Similarly, the RF of a neuron as defined here is the best stimulus for that neuron. It is composed of a number of clusters, each containing some meaningful feature of the RF. These clusters can be considered a set of filters that further analyze the scene.

It is important to note that the RFs of neurons in close proximity have common characteristics with a lot of overlapping areas. Yet, these areas are not exactly the same, a fact that became apparent through these clustering experiments. They respond to the same general areas of stimulation, with the same general kind of ON or OFF response, and to the same general stimulus characteristics. And yet, when clustering was used for analysis, one particular area of these RFs was not quite the same, in all neighboring neurons, as if one of the filters used, as mentioned before, had changed its characteristics. This is probably the main reason why, in order to perceive a line segment, many neuronal units that have similar orientations but different firing thresholds and tuned to different widths are needed.

In summary then, we have developed a method that forces the neuron to perform in the same way that the visual system performs when perceiving patterns of different complexities. It is our hope that correlation between clusters of RFs of neighboring neurons might provide the missing link between the psychophysical local features called textons (Julesz, 1983) and the features found to be optimal stimuli for the neural units in the striate cortex of the monkey.

REFERENCE NOTE

1. Tzanakou, E. *Some nonlinear characteristics in the frog's visual system*. Manuscript submitted for publication, 1983.

REFERENCES

- ARDEN, G. B. Types of response and organization of simple receptive fields in cells of the rabbit's LGB. *Journal of Physiology* (London), 1963, **166**, 449-467.
- BARTZ, M. R. Optimizing a video processor for OCR. In *Proceedings of the International Joint Conference on Artificial Intelligence*, Bedford, Mass: Mitre Corporation, 1969.
- DOYLE, W. Operations useful for similarity-invariant pattern recognition. *Journal of Association for Computer Machinery*, 1962, **9**, 259-267.
- EASTER, S. S. Excitation in the goldfish retina: Evidence for a non-linear intensity code. *Journal of Physiology* (London), 1968, **195**, 253-271.
- FRIEDMAN, H. P., & RUBIN, J. On some invariant criteria of grouping data. *Journal of the American Statistical Association*, 1967, **62**, 1159.
- HARTH, E., & TZANAKOU, E. Alopex: A stochastic method for determining visual receptive fields. *Vision Research*, 1974, **14**, 1475-1482.
- HARTLINE, H. K. The receptive fields of optic nerve fibers. *American Journal of Physiology*, 1940, **130**, 690-699.
- HUBEL, D. H., & WIESEL, T. N. Receptive fields, binocular interactions and functional architecture in cat's visual cortex. *Journal of Physiology* (London), 1962, **150**, 106-154.
- HUBEL, D. H., & WIESEL, T. N. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, 1965, **28**, 229-289.
- HUBEL, D. H., & WIESEL, T. N. Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology* (London), 1968, **195**, 212-243.
- JACOBSON, M. The representation of the retina on the optic tectum of the frog. Correlation between retino-tectal magnification factor and retinal ganglion cell count. *Quarterly Journal of Experimental Physiology*, 1962, **47**, 170-178.
- JULESZ, B. *Foundations of cyclopean perception*. Chicago: University of Chicago Press, 1971.
- JULESZ, B. Textons, the elements of texture perception and their interactions. *Nature*, 1981, **290**, 91-97. (a)
- JULESZ, B. A theory of preattentive texture discrimination based on first-order statistics of textons. *Biological Cybernetics*, 1981, **41**, 131-138. (b)
- JULESZ, B. Textons, the fundamental elements in preattentive vision and perception of textures. *Bell System Technical Journal*, 1983, **62**(6), 1619-1645.
- JULESZ, B., GILBERT, E. N., & VICTOR, J. D. Visual discrimination of textures with identical third-order statistics. *Biological Cybernetics*, 1978, **31**, 137-140.
- KEATING, M. J., & GAZE, R. M. Observations of the surround properties of the receptive field of frog retinal ganglion cells. *Journal of Experimental Physiology*, 1970, **55**, 129-142.
- KUFFLER, S. W. Discharge patterns and functional organization of the mammalian retina. *Journal of Neurophysiology*, 1953, **16**, 37-68.
- LETTVIN, J. Y., MATURANA, C. R., MCCULLOCH, W. S., & PITTS, W. H. What the frog's eye tells the frog's brain. *Proceedings of the Institute of Radio Engineers*, 1959, **47**, 1940-1951.
- LING, R. F. On the theory and construction of k-clusters. *Computer Journal*, 1972, **15**, 326-332.
- MATURANA, C. R., LETTVIN, J. Y., MCCULLOCH, W. S., & PITTS, W. H. Anatomy and physiology of vision in the frog (*Rana pipiens*). *Journal of General Physiology*, 1960, **43**(Suppl. 2), 129-175.
- MCILLWAIN, J. T. Receptive fields of optic tract axons and lateral geniculate cells: Peripheral extent and barbiturate sensitivity. *Journal of Neurophysiology*, 1964, **27**, 1154-1173.
- MORRIN, T. H. A black-white representation of a gray-scale picture. *IEEE Transactions on Computers*, 1974, **23**, 184-186.
- PREWITT, J. M. S., & MENDELSON, M. L. The analysis of cell images. *Annals of the New York Academy of Science*, 1966, **128**, 1035-1053.
- SASAKI, H., BEAR, D. M., & ERVIN, F. R. Quantitative characterization of the unit response in the visual system. *Experimental Brain Research*, 1971, **13**, 239-255.
- SHEPARD, R. N., & ARABIE, P. Additive clustering: Representation of similarities as combinations of discrete overlapping properties. *Psychological Review*, 1979, **86**, 87-123.

- SKLANSKY, J. Image segmentation and feature extraction. *IEEE Transactions on Systems, Man and Cybernetics*, 1968, **8**, 237-247.
- SPINELLI, D. N. Visual receptive fields on the cat's retina. *Science*, 1966, **152**, 1767-1769.
- STONE, J., & FABIAN, M. Summing properties of the cat's retinal ganglion cells. *Vision Research*, 1968, **8**, 1023-1040.
- TZANAKOU, E., & HARTH, E. Alopex: Self-portrait of a feature detector. In R. P. Dooley (Ed.), *Advances in the psychophysical and visual aspects of image evaluation: A program summary*. Washington, D.C: Society for Photographic Scientists and Engineers, 1977.
- TZANAKOU, E., MICHALAK, R., & HARTH, E. The Alopex process: Visual receptive fields by response feedback. *Biological Cybernetics*, 1979, **35**, 161-174.
- ULLMAN, J. R. Binarization using associative addressing. *Pattern Recognition*, 1974, **6**, 127-135.
- WOLFE, R. N. A dynamic thresholding scheme for quantization of scanned images. In *Proceedings of Automatic Pattern Recognition*. Washington, D.C: National Security Industrial Association, 1969.

(Manuscript received July 7, 1982;
revision accepted for publication September 9, 1983.)