Shape recognition and inferior temporal neurons

(visual cortex/pattern recognition/vision/Fourier descriptors/boundary curvature)

ERIC L. SCHWARTZ*, ROBERT DESIMONE[†], THOMAS D. ALBRIGHT[‡], AND CHARLES G. GROSS[‡§]

*Brain Research Laboratory, New York University Medical Center, New York, New York 10016; [†]Laboratory of Neuropsychology, National Institute of Mental Health, Bethesda, Maryland 20205; and [‡]Department of Psychology, Princeton University, Princeton, New Jersey 08544

ABSTRACT   Inferior temporal cortex plays an important role in shape recognition. To study the shape selectivity of single inferior temporal neurons, we recorded their responses to a set of shapes systematically varying in boundary curvature. Many inferior temporal neurons were selective for stimuli of specific boundary curvature and maintained this selectivity over changes in stimulus size or position. The method of describing boundary curvature was that of Fourier descriptors.

What are the neural processes underlying shape recognition? How can we recognize an almost infinite variety of different shapes independent of their size and retinal location? In both man and monkey a likely site for mechanisms of shape recognition is inferior temporal (IT) cortex. Removal of this area impairs the visual recognition of shapes and patterns while leaving basic sensory capabilities, such as acuity, intact (1). Furthermore, many IT neurons are sensitive to the overall shape of objects rather than simply the orientation or location of particular edges (2–4). Finally, IT neurons have large receptive fields that extend into both visual half fields and almost always include the center of gaze (2, 3). These properties of IT neurons depend on the information IT cortex receives from the prestriate visual areas (5, 6).

A common feature of neurons in striate and prestriate cortex is sensitivity to the orientation of local contours or boundaries (7, 8). In this study, we examined how IT cortex might extract information about the overall shape of an object from information about local boundaries. We adopted a method of representing shapes in terms of local boundary orientation that is used in computer pattern recognition systems (9, 10). The method depends, first, on determining the boundary orientation function for the shape—i.e., the orientation (tangent angle) of the shape's boundary measured at regular intervals around the perimeter. Then, the boundary orientation function is expanded in a Fourier series. Each term in the Fourier expansion is associated with a particular frequency, amplitude, and phase and is known as a Fourier descriptor (FD) (9). Equivalently, individual FDs can be extracted by filtering the boundary orientation function with appropriate bandpass filters. Any shape is fully described by its set of FDs, and a smaller set of only the low-frequency terms can often provide the "gestalt" of a shape (9). Thus, the FDs are a powerful and efficient alphabet for representing and classifying shapes. Furthermore, this method of describing shape is independent of both the position and size of the stimulus. Size invariance is achieved by normalizing the perimeter to 2π before calculating the boundary orientation function. Position invariance is achieved by measuring the boundary orientation relative to the orientation of an arbitrary starting point on the perimeter.

Could IT neurons code shape on the basis of global features like FDs? To explore this possibility, we created a set of stimuli from single FDs. The inverse transform of a single FD uniquely determines a closed boundary with a specific number of lobes (frequency), lobe indentation (amplitude), and orientation (phase). We refer to these shapes as "FD stimuli" (Fig. 1). If IT neurons function as "bandpass filters" for shape, one would expect different IT neurons to be tuned to different FD stimuli. Furthermore, the tuning should be independent of the size, contrast, and position of the shape on the retina. The activity of a set of such neurons could specify or code any complex shape.

MATERIALS AND METHODS

We studied 234 visually responsive units in five macaques (*Macaca fascicularis*) immobilized with pancuronium bromide (to eliminate eye movements) and anesthetized with nitrous oxide. All units were tested with slits of light, a set of complex objects, and a series of FD stimuli. The FD stimuli were white on black patterns projected onto a tangent screen at the center of gaze. The stimuli were oscillated 1 deg per sec for a duration of 2.5 sec and were presented in a pseudo-random sequence for a total of 10 trials each. The FDs ranged in frequency from 2 to 64 cycles per perimeter in an octave series, and each unit was initially tested with this frequency series, at constant size (28 deg^2), amplitude, and phase. ("Amplitude" and "phase" are the amplitude and phase of the frequency component in the Fourier expansion of the normalized boundary orientation function; see ref. 9.) For some units, the frequency series was repeated at different sites within the receptive field, with different FD amplitudes, with different stimulus sizes, or with reversed stimulus contrast.

IT neurons were defined as visually responsive to a stimulus if there was an increase or decrease in the firing rate of at least two standard deviations from the spontaneous firing rate.

RESULTS

Fifty-four percent of the visually responsive IT neurons were tuned to the frequency of the FD stimuli, with a mean bandwidth of 2.0 octaves (full-width at half-maximal response). Different cells had different best frequencies, and all frequencies were about equally represented, except for frequency 64, which had a reduced incidence. Typical tuning curves are shown in Fig. 2. Thus, for about half of the cells in IT, the frequency of the FDs appeared to be a relevant stimulus dimension. The remaining neurons either did not respond to the FD stimuli (26%), responded equally well to all of the FD stimuli (13%), or had multimodal tuning curves (7%).

To test whether IT neurons are selective for a particular shape regardless of its size, 31 tuned cells were retested with FD stimuli

Abbreviations: IT, inferior temporal; FD, Fourier descriptor.
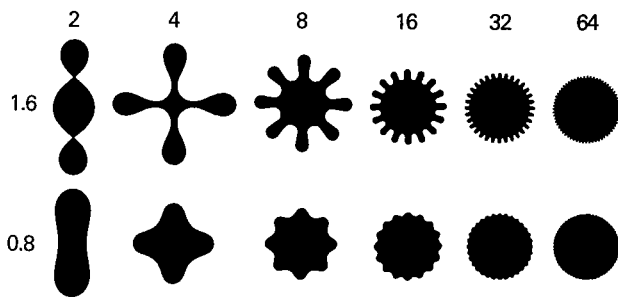§ To whom reprint requests should be addressed.

FIG. 1.   Examples of FD stimuli varying in frequency (2–64 cycles per perimeter) and amplitude (0.8 and 1.6).

of the same frequency, amplitude, and phase as the standard ones but either larger (50 deg^2) or smaller (13 deg^2) in size, or both. For two-thirds of these cells the optimal frequency and the shape of the tuning curve remained similar over changes in stimulus size, although the absolute level of the response varied (Figs. 2 *A* and *C*). Thus, the selectivity of these neurons for shape remained invariant over changes in stimulus size.

To test whether IT neurons were selective for a particular shape independent of its position on the central retina, 18 frequency tuned cells were retested with the FD stimuli repositioned within their receptive field 3–5 deg from the center of gaze. For 72% of these cells, the optimal FD frequency and the shape of the tuning curve remained similar at the more peripheral retinal site, although, as with size retesting, the absolute level of the response varied (responses generally decreasing with eccentricity; see Fig. 2*D*). Thus, these neurons may be described as showing shape constancy over retinal translation (11).

For three of the four cells so tested, reversing the contrast

of the stimuli did not alter the shape of their tuning curves for frequency (Fig. 2*E*).

The finding that some IT cells maintain their selectivity for FD stimuli of a particular frequency over changes in size, position, and contrast cannot be easily explained in terms of selectivity for local features such as the position or orientation of an edge. These cells must be sensitive, in some fashion, to the overall shape of the stimulus.

If IT cortex is coding shape by using a mechanism similar to that of FDs, at least some of its cells should be sensitive to the amplitude of FD stimuli. For 22 FD tuned cells, we repeated the frequency series at a lower FD amplitude. Fifty-nine percent of these cells exhibited tuning curves with the same shape as at the higher amplitude (Fig. 2*A*). Generally, the tuning curves flattened out at lower FD amplitudes, suggesting that the selectivity of the tuned cells for frequency increased with FD amplitude. For another 15 cells we studied the effect of FD amplitude directly. The frequency of the stimuli was held constant and the amplitude varied over five octaves. Eleven of these showed unimodal tuning curves for amplitude, and for 9 of the 11, the response of the cell was an increasing function of amplitude (Fig. 2*B*).

DISCUSSION

The present results support the possibility that boundary curvature is an important stimulus dimension processed by IT cortex. Just as striate cortex analyzes local boundary orientation, IT cortex may be involved in the analysis of global boundary orientation. A possible mechanism by which local edge information in striate and prestriate cortex might be combined to yield a representation of boundary curvature in IT has been described elsewhere (12).

Even if some IT neurons represent or code boundary cur-
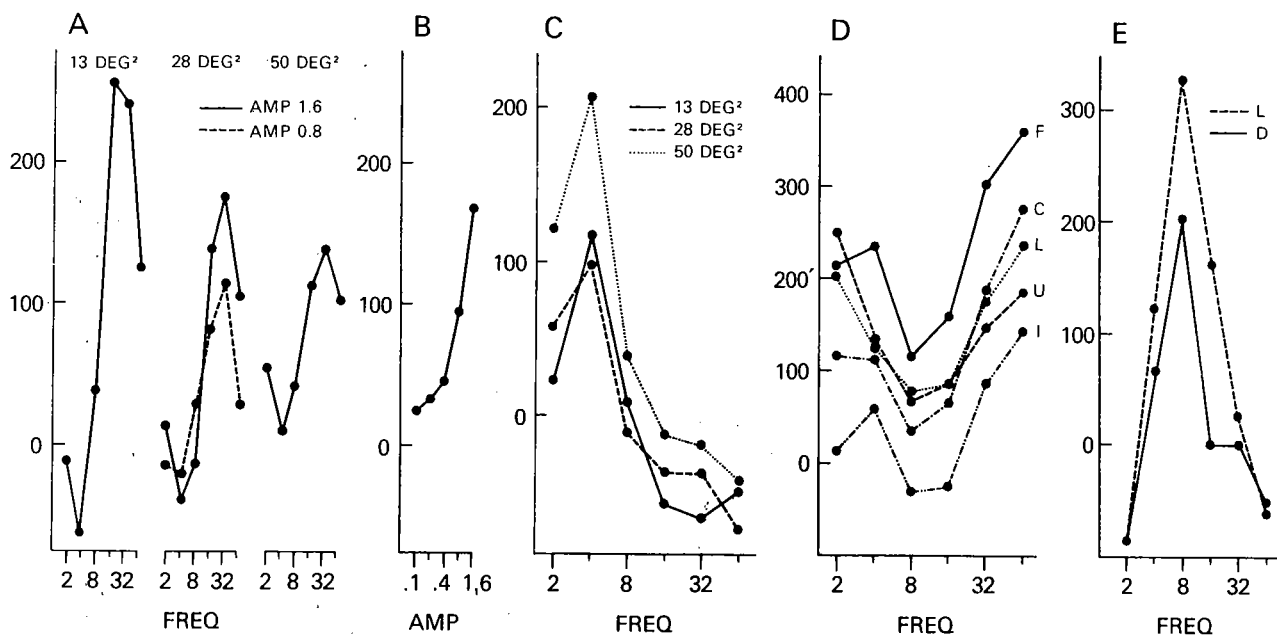

FIG. 2.   Responses of IT neurons to FD stimuli. Responses are plotted as percent change in firing rate over the mean spontaneous rate. Each point represents the mean of 10 presentations. (*A*) Effect of stimulus size and FD amplitude (AMP) on frequency (FREQ) tuning. Note that the shape of the tuning curve remains similar over changes in stimulus size and amplitude. (*B*) Monotonic increase in response with increase in FD amplitude. Stimuli were all 32 cycles per perimeter. (*C*) Effect of stimulus size on frequency tuning. The shape of the tuning curve remains similar at different sizes. (*D*) Effect of location of stimulus on frequency tuning. Stimuli were centered on the fovea (F) or 5 deg into the upper (U), lower (L), contralateral (C), or ipsilateral (I) visual field. Note that the frequencies eliciting the maximal and minimal responses at each location remain the same but the magnitude of the responses is greatest at the fovea and least in the ipsilateral field. (*E*) Effect of contrast on frequency tuning. L, Light pattern on dark background; D, dark pattern on light background. Note that the best frequency is independent of contrast. For *A–E*, unless otherwise indicated, the amplitude of the FD stimuli was 1.6, the area was 28 deg^2, and they were light on a dark background.

vature, this cannot be the exclusive function of IT cortex. Some IT cells were found previously to be selective for color, texture, spatial frequency, three-dimensionality, or even complex stimuli such as faces or hands (2–4, 13–17). These stimulus dimensions are not readily described in terms of boundary curvature. Some, but not all, of a sample of such units tuned to color, texture, or spatial frequency were also found to be tuned to FD stimuli (unpublished data). Thus, selectivity for boundary curvature appears to be neither a characteristic of all IT cells nor the exclusive response property of many.

Whether or not IT cortex employs a mechanism for shape recognition similar to the method of FDs, the use of shapes derived from FDs (*i*) demonstrated that some IT units are selective for overall shape, (*ii*) provided us with a quantitative measure of this selectivity, and (*iii*) showed that selectivity is often maintained over translations of size, retinal locus, and contrast.

1. Gross, C. G. (1973) in *Handbook of Sensory Physiology*, ed. Jung, R. (Springer, Berlin), Vol. 7, Part 3B, pp. 451–482.
2. Gross, C. G., Rocha-Miranda, C. E. & Bender, D. B. (1972) *J. Neurophysiol.* **35**, 96–111.
3. Desimone, R. & Gross, C. G. (1979) *Brain Res.* **178**, 363–380.
4. Desimone, R., Albright, T. D., Gross, C. G. & Bruce, C. (1980) *Neurosci. Abstr.* **6**, 581 (abstr.).
5. Desimone, R., Fleming, J. & Gross, C. G. (1980) *Brain Res.* **184**, 41–55.
6. Rocha-Miranda, C. E., Bender, D. B., Gross, C. G. & Mishkin, M. (1975) *J. Neurophysiol.* **38**, 475–491.
7. Zeki, S. M. (1978) *J. Physiol. (London)* **277**, 273–290.
8. Baker, J. F., Peterson, S. E., Newsome, W. T. & Allman, J. M. (1981) *J. Neurophysiol.* **45**, 387–406.
9. Zahn, C. T. & Roskies, R. Z. (1972) *IEEE Trans. Comput.* **21**, 269–281.
10. Persoon, E. & Fu, K. S. (1977) *IEEE Trans. Syst. Man Cybern.* **7**, 170–179.
11. Gross, C. G. & Mishkin, M. (1977) in *Lateralization in the Nervous System*, eds. Harnad, S., Doty, R., Jaynes, J., Goldstein, L. & Krauthamer, G. (Academic, New York), pp. 109–122.
12. Schwartz, E. L. (1980) *Vision Res.* **20**, 645–669.
13. Perrett, D. I., Rolls, E. T. & Caan, W. (1979) *Neurosci. Lett.* **3**, Suppl., 358 (abstr.).
14. Richmond, B. J. & Wurtz, R. H. (1981) in *Changing Concepts of the Nervous System*, eds. Morrison, A. R. & Strick, P. L. (Academic, New York), pp. 411–422.
15. Sato, T., Kawamura, T. & Iwai, E. (1980) *Exp. Brain Res.* **38**, 313–319.
16. Rolls, E. T., Judge, S. J. & Sanghera, M. K. (1977) *Brain Res.* **130**, 229–238.
17. Fuster, J. M. & Jervey, J. P. (1982) *J. Neurosci.* **2**, 361–375.